Australasian Journal of Philosophy

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rajp20

On the Origins of Old Evidence

Benjamin Eva & Stephan Hartmann

To cite this article: Benjamin Eva & Stephan Hartmann (2020) On the Origins of Old Evidence, Australasian Journal of Philosophy, 98:3, 481-494, DOI: 10.1080/00048402.2019.1658210

To link to this article: https://doi.org/10.1080/00048402.2019.1658210

Published online: 14 Oct 2019.

AUSTRALASIAN JOURNAL OF PHILOSOPHY
2020, VOL. 98, NO. 3, 481–494
https://doi.org/10.1080/00048402.2019.1658210

On the Origins of Old Evidence

Benjamin Eva [a] and Stephan Hartmann [b]

[a]University of Konstanz; [b]Ludwig Maximilian University Munich

ABSTRACT
The problem of old evidence, first described by Glymour [1980], is still widely regarded as one of the most pressing foundational challenges to the Bayesian account of scientific reasoning. Many solutions have been proposed, but all of them have drawbacks and none is considered to be definitive. Here, we introduce and defend a new kind of solution, according to which hypotheses are confirmed when we become more confident that they provide the *only* way of accounting for the known evidence.

ARTICLE HISTORY Received 15 May 2018; Revised 14 May 2019

KEYWORDS Bayesianism; philosophy of science; epistemology; probability; old evidence

1. Introduction

The Bayesian Problem of Old Evidence (henceforth, POE) is easy to state. If the probability of the evidence, $P(\mathrm{E})$, is 1, then the likelihood $P(\mathrm{E} \mid \mathrm{H})$ is also 1,[1] and hence, by the definition of conditional probabilities,

$$P(\mathrm{H}|\mathrm{E}) = \frac{P(\mathrm{E}|\mathrm{H})\, P(\mathrm{H})}{P(\mathrm{E})} = P(\mathrm{H}) \tag{1}$$

So, Bayesian confirmation theory has the consequence that known 'old' evidence can never confer confirmation upon scientific hypotheses. As Glymour [1980] pointed out, this observation conflicts with scientific practice. To illustrate, consider the famous example of General Relativity and the anomalous perihelion of Mercury. The anomalous perihelion of Mercury (= E) was known long before Einstein formulated the General Theory of Relativity (= H) in 1915. When he then derived E from H—that is, when he showed that H implies E—he, as well as the whole scientific community, took this to constitute important confirmatory evidence in favour of H (see also Garber [1983]). Thus, it is commonly said that General Relativity was decisively confirmed by the anomalous perihelion of Mercury, which should be impossible by the lights of Bayesian confirmation theory (since E in this case is old evidence).

In the years since POE was first described by Glymour, a plethora of prospective solutions have been defended in the literature (see, for instance, Garber [1983],

[1] To see this, note that the law of total probability implies that $P(\mathrm{E}) = P(\mathrm{E} \mid \mathrm{H})\, P(\mathrm{H}) + P(\mathrm{E} \mid \neg\mathrm{H})\, P(\neg\mathrm{H})$. Hence, $P(\mathrm{E}) = 1$, for all $P(\mathrm{H}) \in (0,1)$ implies that $P(\mathrm{E} \mid \mathrm{H}) = P(\mathrm{E} \mid \neg\mathrm{H}) = 1$.

Howson [1991], Hartmann and Fitelson [2015], and Sprenger [2015]). However, there is still little consensus regarding the proper resolution of the problem. In this paper, we present a novel approach to modelling confirmation by old evidence, and we show that this approach elegantly resolves several major technical and interpretational difficulties with existing solutions to POE. Our approach is characterized by the idea that a hypothesis H is confirmed by old evidence E when it is observed that H provides (or is likely to provide) the *only* available way to account for E. In section 2, we outline and criticise several of the main existing solutions to POE from the literature, focusing especially on the so-called 'Garber-style' and counterfactual solutions. Section 3 presents a new approach to POE, according to which confirmation by old evidence is possible when it is realized that the relevant hypothesis provides the *only* way to account for the known evidence, and illustrates how this new perspective on POE overcomes the limitations of extant solutions.

2. Existing Solutions

2.1 Garber-Style Solutions

Perhaps the most influential response to POE is due to Garber [1983]. His approach is characterized by the introduction of a new class of atomic sentences concerning the deductive relationships between the sentences in the original language. In particular, as well as considering the standard sentences H and E, Garber also considers the sentence $H \vdash E$ (H 'entails' E). He argues that the apparent confirmation of H by old evidence E is actually due to a change in our beliefs about the logical sentence $H \vdash E$. The General Theory of Relativity (GR) was confirmed not by the anomalous perihelion of Mercury, but rather by the realization that the anomalous perihelion of Mercury can be derived from GR. The details of Garber's approach have been the subject of much debate (see, for example, Eells [1990]), and new variants of the main idea are still being developed (see, for example, Hartmann and Fitelson [2015]). Solutions to POE that utilise the learning of atomic logical sentences of the form $H \vdash E$ have come to be referred to as 'Garber-style solutions'.

Now, classical Bayesianism assumes that we should treat rational agents as being *logically omniscient*, meaning that all logical truths should have subjective probability 1 and, for any logically incompatible propositions A and B, we should have $P(A \vee B) = P(A) + P(B)$. Clearly, this assumption precludes the kind of Garber-style solution described above. For, if an agent is already aware of all logical truths, then she can never learn a sentence of the form $H \vdash E$, since any such sentence will already be known. So, in order to get any Garber-style solution off the ground, we need to surrender the assumption of logical omniscience. This move has received a lot of support in the literature, largely because it is psychologically implausible to imagine that a rational agent is ever aware of *all* the logical relations that obtain between the sentences of her language (see, for example, Eells [1990: 209]).

Thus, Garber-style solutions to POE rely on the independently plausible move of rejecting logical omniscience as a universal constraint on the credences of rational agents. But once this assumption has been relaxed and, following Garber, we introduce a new class of logical sentences into our language about which agents can have partial beliefs, we need to specify some constraints on the relationship between the new purely logical sentences and the standard sentences of our original language. Specifically, we need to know how the logical sentence $H \vdash E$ interacts with H and E. Letting $X := H \vdash E$, Garber suggests the following basic constraints:

G1. $P(E \mid H, X) = 1$,

G2. $P(H, E, X) = P(H, X)$.

These constraints are generally taken to be uncontroversial. They tell us that the $\vdash$ relation satisfies a basic form of *modus ponens*. If I am certain of both H and X, then I should also be certain of E. Armed with G1 and G2, Garber was able to prove the following results, where $\mathscr{L}$ is the original scientific language and $\mathscr{L}'$ is the language obtained by adding the additional atomic logical formulae of the form X.

GT1. There exists at least one probability function on $\mathscr{L}'$ such that every non-trivial atomic sentence of the form X has a non-extreme probability value $P(X) \in (0, 1)$.

GT2. For any atomic sentence of the form X, there are infinitely many probability functions satisfying (i) $P(E) = 1$ and (ii) $P(H \mid X) > P(H)$.

GT1 tells us that it is possible to be genuinely uncertain about any set of logical sentences in a coherent way. Crucially, GT2 shows that it is always possible to find a probability function that solves POE, in the sense that the logical proposition X confirms H (that is, conditioning on X increases the probability of H) even though E is old evidence. So, when we learn that the General Theory of Relativity implies the anomalous perihelion of Mercury, we simply condition on the relevant logical sentence and it will generally be possible that doing so confirms the theory.

However, the mere existence of probability functions satisfying the conditions in GT2 is not enough to solve all of our problems here. We also need to know *under what conditions* a probability function will solve POE. Most importantly, we need to verify that the conditions under which such functions exist are plausible and are easily satisfied in realistic scientific reasoning contexts.

This problem has been taken up by Jeffrey [1983], Earman [1992], and Hartmann and Fitelson [2015]. Introducing the additional logical formula $Y := H \vdash \neg E$, Jeffrey considers these five constraints

J1. $P(E) = 1$,

J2. $P(H), P(X), P(Y) \in (0, 1)$,

J3. $P(X, Y) = 0$,

J4. $P(H \mid X \vee Y) \geq P(H)$,

J5. $P(H, \neg E, Y) = P(H, Y)$.

With this, he achieves the following result:

JT. Let P satisfy J1–J5. Then $P(H \mid X) > P(H)$.

So, any probability function satisfying J1–J5 solves POE. It only remains to motivate the constraints J1–J5. J1 reflects the fact that we are dealing with old evidence. J2 encodes the standard assumption that, to begin with, we are uncertain about the logical relationship between the hypothesis and the evidence. J3 is a consistency requirement: that is, we believe that H is consistent. J5 is just an instantiation of the *modus ponens* condition G2. This leaves J4. Intuitively, J4 says that, *ceteris paribus*, we should prefer hypotheses that make definite predictions about the evidence to those that don't. However, as Sprenger [2015: 390] notes, this condition

> conflates an *evidential* virtue of a theory with a *methodological* one. We are well-advised to cherish theories of which we know that they make precise predictions on an interesting subject matter, even if we do not yet know what these predictions look like in detail. This is basically a Popperian rationale for scientific inquiry: go for theories that have high empirical content, that make precise predictions, and develop them further. They are the ones that will finally help to solve urgent scientific problems … But following this rule is very different from arguing that the informativity and empirical content of a theory increases its *plausibility*. Actually, Popper thought the other way round: theories with high empirical content rule out more states of the world and will have low (logical) probability!

In light of this observation, we are forced to surrender J4 as a universal constraint on our credences about logical propositions. Since J4 plays a crucial role in the proof of JT, this undermines Jeffrey's proposed solution to POE.

Earman [1992] offers two alternatives to J4 as the crucial constraint in the solution of POE. The first one is this inequality:

E1. $P(\mathrm{H} \mid \mathrm{X}) > P(\mathrm{H} \mid \neg\mathrm{X}, \neg\mathrm{Y})$.

E1 says that learning that H implies the evidence should be more favourable to our belief in H than learning that H makes no definite predictions about the evidence. But, to quote Sprenger, 'this condition just seems to beg the question.' For we are looking for independently plausible conditions on P that ensure that learning $\mathrm{H} \vdash \mathrm{E}$ will increase $P(\mathrm{H})$. E1 simply imposes something very close to this property without offering any independent motivating arguments. The second alternative constraint is below:

E2. $P(\mathrm{X} \vee \mathrm{Y}) = 1$.

E2 can also be ruled out immediately. For it seems clear (as Earman himself acknowledges) that, in practice, scientists are often uncertain about whether a candidate hypothesis H has any bearing on some body of known evidence. But E2 requires that the scientist should be certain that H will make definite predictions about E as soon as H is formulated. This is clearly an unrealistic and unreasonable assumption.

Hartmann and Fitelson [2015] also identify a set of conditions that are jointly sufficient to guarantee a full solution to POE. Unlike those proposed by Jeffrey and Earman, these conditions are purely qualitative. Before we see the conditions, it is also worth noting that Hartmann and Fitelson suggest an important variation on the interpretation of the new class of atomic sentences introduced by Garber. Specifically, they argue that a purely logical interpretation is unduly restrictive, and they allow for the possibility of a broadly explanatory interpretation of the new variables X and Y. We read [ibid.: 2]:

> we think interpreting 'X' and 'Y' as 'H entails E' and 'H refutes E' is unduly restrictive. It is more plausible to suppose that what is learned in cases of old evidence … may not (always) be a *logical* fact. To be more precise, let 'X' and 'Y' be interpreted as follows:
>
> $\mathrm{X} =_{\mathrm{def}}$ H *adequately explains* (or *accounts for*) E.
>
> $\mathrm{Y} =_{\mathrm{def}}$ H's best competitor (H') *adequately explains* (or *accounts for*) E.

We are sympathetic to this generalization, since it allows for a more flexible approach to dealing with old evidence. There is simply no good reason to suppose that the relevant fact that we learn in cases of old evidence is always logical in character. Furthermore, it is plausible to claim that scientists sometimes take it to be significant when a hypothesis

H is found to provide a good explanation of some evidence E, even though H does not entail E.[2] Finally, it is also plausible to contend that the explanatory virtues of a hypothesis' competitors play a significant role in its evaluation. Part of the reason that General Relativity was so strongly confirmed by the anomalous perihelion of Mercury was surely that no other extant theory was able to explain that fact in a satisfactory manner. If some other competing theory were also capable of explaining Mercury's advanced perihelion, then the degree of confirmation bestowed on GR by Einstein's derivation would surely have been less. Thus, it is natural to loosen the interpretation of the new class of propositions in Garber-style solutions to allow for facts concerning the explanatory power of a hypothesis and its competitors to play a role.

Armed with this new interpretation, Hartmann and Fitelson propose the following qualitative constraints (where it is assumed that $P(\mathrm{E}) = 1$):

HF1. $P(\mathrm{H} \mid \mathrm{X}, \neg\mathrm{Y}) > P(\mathrm{H} \mid \neg\mathrm{X}, \neg\mathrm{Y})$,

HF2. $P(\mathrm{H} \mid \mathrm{X}, \neg\mathrm{Y}) > P(\mathrm{H} \mid \neg\mathrm{X}, \mathrm{Y})$,

HF3. $P(\mathrm{H} \mid \mathrm{X}, \mathrm{Y}) > P(\mathrm{H} \mid \neg\mathrm{X}, \mathrm{Y})$,

HF4. $P(\mathrm{H} \mid \mathrm{X}, \mathrm{Y}) \geq P(\mathrm{H} \mid \neg\mathrm{X}, \neg\mathrm{Y})$.

They then prove the following theorem:

HFT. Let P satisfy HF1–HF4. Then $P(\mathrm{H} \mid \mathrm{X}) > P(\mathrm{H})$.

So, any probability function satisfying HF1–HF4 will allow us to solve POE by modelling the relevant agents as having learned the new atomic sentence X. At first blush, HF1–HF4 all look plausible, given the new interpretation of X and Y posited by Hartmann and Fitelson. HF1 and HF2 essentially require that the 'best' outcome for the probability of H is if H accounts for the old evidence E, but its closest competitor H* doesn't. HF3–HF4 require that both of H and H* accounting for E is no worse for H's probability than neither H nor H* accounting for the evidence. This all sounds natural (see Kao [2017] for a defence of HF1–HF4 in terms of historical case studies).

However, although we agree with Hartmann and Fitelson that (i) it is important to consider not just the hypothesis H itself, but also the hypotheses with which H is competing, (ii) it makes sense to relax the requirement that X and Y are always interpreted in a logical way, and (iii) the constraints HF1–HF4 are overall far more plausible than their problematic predecessors, we still contend that the proposed solution to POE is not the full story. To see this, note that, once we've introduced the proposition Y, it is natural to expect that learning $\mathrm{X} \wedge \neg\mathrm{Y}$ should never decrease the probability of H. If I learn that H is able to account for the evidence but its competitors are not, this should generally make me more confident of the truth of H, and it should certainly never decrease my confidence in H. For, in this situation, I become more confident that H is the *only* way that I can possibly account for the evidence, and I thereby become more confident that H *has to be true*. However, it is easy to see that HF1–HF4 are not sufficient to guarantee that $P(\mathrm{H} \mid \mathrm{X} \wedge \neg\mathrm{Y}) \geq P(\mathrm{H})$. Thus, although introducing the new proposition Y together with the conditions HF1–HF4 allows H to be confirmed by X,

[2] Note that Hartmann and Fitelson remain non-committal regarding which conception of scientific explanation is appealed to in the definition of X and Y. In what follows, we will also remain agnostic, and operate with a folk-theoretic understanding of scientific explanation.

it also introduces a new problem—namely, that of finding extra conditions under which $X \wedge \neg Y$ is guaranteed not to disconfirm H. If it's not the case that the proposition Y behaves in a way that is consistent with its intended interpretation, then the justification of the conditions HF1–HF4 becomes less clear. For these conditions are only compelling to the extent that Y really does represent the proposition 'Some alternative to H accounts for the evidence', and this interpretation is only justified, we contend, if $X \wedge \neg Y$ never disconfirms H. If instead we interpret X and Y as $H \vdash E$ and $H \vdash \neg E$, respectively, Hartmann and Fitelson's conditions will make little sense. Indeed, on this interpretation, HF3 implies that H should have non-zero probability *when we condition on* H *being inconsistent*. Similarly, HF4 says that conditioning on H being inconsistent cannot render H less likely than conditioning on H making no predictions about the evidence. On this reading, HF3 and HF4 are deeply implausible. So, Hartmann and Fitelson really are committed to the interpretation of Y as 'Some alternative to H accounts for the evidence', but they have not yet demonstrated that this interpretation is justified by the way that Y interacts with other propositions in the language. So, although Hartmann and Fitelson's proposal represents significant progress when compared to existing instantiations of Garber's solution to POE, it does not provide the full story.

Overall, then, we've considered three separate attempts at filling in the gaps in the Garber-style solution to POE, and all three have turned out to be unsatisfactory. In the absence of any satisfactory completion, one is compelled to regard the proposed solution with a degree of scepticism. It does us no good to know that there are probability distributions for which POE can be solved, if we have no reason to believe that distributions of this kind are actually representative of the way that scientists and rational agents do or should reason.

2.2 Howson's Counterfactual Solution

Another approach to POE that has received significant attention in the literature is due primarily to Howson (for instance, [1984, 1991]). In contrast to the Garber-style solutions that we've studied thus far, which surrender the usual Bayesian assumption of logical omniscience, Howson proposes an alternative amendment to the theory. In particular, he posits an alternative interpretation of Bayesian confirmation. For Howson, confirmation should not be thought of as positive probabilistic relevance relative to actual degrees of belief. Rather, Howson argues that, in order to gauge the confirmatory relationship between old evidence E and the hypothesis H, we should counterfactually imagine what would happen to our belief in H if we were to remove the old evidence E from our background knowledge K: that is, we imagine how much learning E would increase our belief in H *if we didn't already know that* E *obtained*. This approach seems to provide an elegant solution to POE. However, the interpretational shift is a major one that engenders a plethora of new philosophical and technical challenges. Thus, we read [Eells 1990: 208]:

> there will not necessarily be any *particular* degrees of belief that we can say a person would have had in E, or in H given E, if this person's degree of belief in E had been less than 1. Indeed, it seems plausible that in some cases H *would not even have been formulated* had E not been learned. And surely there also will be cases in which the person's knowledge of E *saved his life* at some time in the past, so that had the individual's degree of belief in E been less than 1, the person would be *dead* now. Also, there are of course the well-known difficulties attending the proper interpretation of counterfactual conditionals that would befall any such modification of Bayesian confirmation theory.

These criticisms are sufficient to cast major doubt on the viability of Howson's counterfactual approach to POE. And even if the advocate of the counterfactual solution is able to resolve these issues, there is still the more general worry that in order to solve POE, we have surrendered a key facet of Bayesianism's philosophical content—namely, the interpretation of probabilities as *actual* subjective degrees of belief. Altogether, we take these considerations to constitute a compelling case for rejecting Howson's counterfactual approach to POE.

It should be noted that Howson [2017] abandons the counterfactual approach to POE outlined here, and argues that POE can be easily resolved in the kind of 'objective Bayesian' setting advocated by Jaynes [1968] and Rosenkrantz [1981]. We remain silent on Howson's claim that POE is rendered a non-problem in an objective Bayesian setting, and we focus rather on the extent to which POE is really a problem for subjective Bayesians. Given that subjective Bayesianism is the dominant approach in the literature, this is clearly a very pressing philosophical problem.[3]

2.3 Sprenger's Hybrid Approach

Sprenger [2015] offers yet another prospective solution to POE. Broadly speaking, his approach can be characterized as a kind of hybrid of the counterfactual and Garber-style solutions described above. For Sprenger, cases of confirmation by old evidence involve both logical learning and the updating of a particular kind of counterfactual probability distribution. The details of his approach differ significantly from the standard versions of the Garber-style and counterfactual solutions, and demand independent examination.

Sprenger begins by positing some constraints on when one should intuitively think of the logical sentence $X := H \vdash E$ as confirming H relative to E:

S1. $P(E \mid H, X) = 1$,

S2. $P(E \mid \neg H, X) = P(E \mid \neg H, \neg X)$.

S1 is another intuitive *modus-ponens*-style condition. If we know with certainty both H and that H entails E, then we also know E with certainty. S2 is only slightly more substantive, and Sprenger motivates it with the observation that 'learning that certain events are predicted by a *refuted hypothesis* is just irrelevant to our assessment of the *plausibility* of those events' [ibid.: 391]. This all sounds plausible, and S2 looks like a natural constraint. However, as Sprenger notes, the condition will always be trivially satisfied if E is old evidence. For in that case $P(E \mid \neg H, X) = 1 = P(E \mid \neg H, \neg X)$ is guaranteed to hold. In light of this consideration, Sprenger proposes that the probability function in S1 and S2 should be reinterpreted counterfactually. In particular, he proposes that we replace the standard Bayesian distribution P with a particular kind of counterfactual distribution $\mathfrak{p}$ [ibid.]:

> Instead of just eliminating E from the background knowledge, $\mathfrak{p}$ should represent the degrees of belief of a scientist who has a sound understanding of theoretical principles and their impact on observational data, but who is an imperfect logical reasoner and lacks full empirical knowledge … In particular, the scientist does not know the entire observational history … How probable

[3] For sustained philosophical criticisms of objective Bayesianism, see, e.g., Seidenfeld [1979]. For a contemporary defence of objective Bayesianism, see Williamson [2010].

> would the actual evidence E be if H were true? How probable would E be if H were false? When H and ¬H are two definite statistical hypotheses, such judgments are immediately given by the corresponding sampling distribution. But even in broader contexts, such judgments may be straightforward, or at least a matter of consensus in the scientific community.

Sprenger also notes that, unlike the counterfactual distributions used by Howson, $\mathfrak{p}$ 'is not bound to a particular credence function'. Presumably then, $\mathfrak{p}$ cannot be thought of as representing the actual belief state of any specific scientist, but only of that of some ideal scientist whose epistemic situation satisfies the constraints listed above. We have a number of reservations about this construal of $\mathfrak{p}$. Most pertinently, it is far from obvious that the belief state of such an ideal scientist is uniquely determined by a specification of their understanding of theoretical principles and the impact of those principles on observational data. Scientists with sound understanding of theoretical principles and equal understanding of the relevant observational data frequently disagree about the respective likelihoods of different theories. For example, there is currently great controversy amongst cosmologists concerning the extent to which the existence of dark matter is rendered probable by the relevant observational data. So, although we agree with Sprenger that the relevant judgments *may* be straightforward in certain cases, we do not think that they are generally so. Thus, the counterfactual interpretation described above strikes us as somewhat imprecise in its current formulation.

But let's put these concerns to one side for the moment. Armed with the new counterfactual distribution, Sprenger proposes the following set of constraints on $\mathfrak{p}$:

S1. $\mathfrak{p}(E \mid H, X) = 1$,

S2. $\mathfrak{p}(E \mid \neg H, X) = \mathfrak{p}(E \mid \neg H, \neg X) > 0$,

S3. $\mathfrak{p}(E \mid H, \neg X) < ((1 - \mathfrak{p}(X \mid \neg H))/\mathfrak{p}(X \mid \neg H)) \cdot (\mathfrak{p}(X \mid H)/(1 - \mathfrak{p}(X \mid H)))$;

and goes on to prove the following theorem:

ST. Let $\mathfrak{p}$ satisfy S1–S3. Then X confirms H for old evidence E: that is, $\mathfrak{p}(H \mid X) > \mathfrak{p}(H \mid E)$.

We have already noted that S1 and S2 are intuitive and relatively uncontroversial. So, the only remaining obstacle for this solution to POE is to justify S3. Clearly, S3 is a much less natural condition than S1 or S2. Sprenger shows that S3 will hold under a number of different circumstances (for example, when Jeffrey's problematic condition J4 holds); but when H and X are negatively relevant to each other, S3 will only hold if $\mathfrak{p}(E \mid \neg X, H)$ is not too close to 1. Sprenger argues that this should generally be guaranteed since, 'given that H is assumed to be true, but that by ¬X, it does not fully account for E, E should not be a matter of course for a rational Bayesian agent' [ibid.: 392]. But this is not a convincing justification. For an agent might be confident (without being certain) of E's truth for reasons *completely independent of* H. I may be very confident that it is sunny outside (E), whilst knowing (or being very confident) both that I am sitting at my desk (H), and that my sitting at my desk has no logical or explanatory bearing on its being sunny outside (¬X). So, I would be perfectly justified in having a credal state P such that $P(E \mid \neg X, H)$ is arbitrarily high. There is simply no good reason to accept S3 as a universal constraint on the credences of rational agents.

Overall then, Sprenger's hybrid solution to POE faces two problems: (i) the specification of the counterfactual distribution $\mathfrak{p}$ is too vague to provide a general account of how old evidence can confirm, and (ii) the constraint S3 is devoid of any compelling independent motivation.

In summary, we have studied three classes of prospective solution to POE (Garber-style, counterfactual, and hybrid) and found all three to be wanting. We saw that, despite several valiant attempts, nobody has yet been able to spell out the necessary details of a Garber-style solution in a satisfactory manner. The problems faced by the counterfactual approach were rather more fundamental. In the absence of any principled and coherent way of determining the counterfactual distribution, the approach seems to raise more problems than it solves. Finally, Sprenger's hybrid approach inherits some of the problematic aspects from each of its predecessors.

We have reached an impasse. We now turn to providing a novel solution to POE that bypasses the problems faced by the approaches considered so far.

3. The Origins of Old Evidence

Of the proposed solutions considered thus far, we view Hartmann and Fitelson's (HF) proposal as the most promising. Unlike other variants of the Garber-style solution, HF's solution allows for the natural possibility that explanatory facts as well as purely logical facts can provide the basis for confirmation by old evidence. It also takes into account the important observation that the plausibility of any hypothesis is affected by the availability of competing hypotheses and their ability to account for the relevant evidence. But, as we noted in section 2.1, HF's story is also incomplete. While HF provides plausible conditions (HF1–HF4) under which the proposition X = 'H *adequately explains* E' confirms H, these conditions still allow for the possibility that the proposition $X \wedge \neg Y$ = 'H *adequately explains* E *and* H*'s best competitor does not adequately explain* E' can disconfirm H. And this is implausible. Learning that H explains E, and that none of H's competitors explain E, should always make us more confident in the truth of H. Thus, the advocate of HF's solution is obliged to identify plausible extra conditions (consistent with HF1–HF4) under which $X \wedge \neg Y$ is guaranteed not to disconfirm H. Fortunately, such conditions do exist. Specifically, consider the following constraint (where it is assumed that $P(E) = 1$):

HF4*. $P(H \mid X, Y) = P(H \mid \neg X, \neg Y)$.

HF4* is a slight strengthening of HF4. It says that H is equally plausible in the case where both it and its competitors account for E and the case where neither it nor its competitors account for E. Intuitively, the idea is that, in both of these cases, there is no difference between H and its competitors (in terms of their ability to account for E), and hence that there is no significant confirmatory distinction between the cases. We take this to be an intuitively compelling assumption, which captures the idea (defended in more detail below) that hypotheses typically only receive significant confirmation from old evidence when it is observed that they explain the relevant evidence more effectively than their competitors. Armed with HF4* we can show this (proof in the Appendix):

HFT*. Let P satisfy HF1–HF3 and HF4*. Then (*i*) $P(H \mid X) > P(H)$, (*ii*) $P(H \mid \neg Y) > P(H)$, and (*iii*) $P(H \mid X \wedge \neg Y) > P(H)$.

So, when we assume that HF1–HF3 and HF4*, we get the desired result. H is confirmed both by the realization that it accounts for the known evidence, and also by the realization that it, but none of its competitors, is able to account for the evidence. From a technical perspective, this looks like a perfectly adequate response to POE. Both X and Y behave exactly as one would expect given the proposed interpretations, the proposed conditions look perfectly plausible, and the desired confirmatory relations hold.

And this leads to a more general philosophical observation—namely, that, in real cases of confirmation by old evidence, what typically matters is not simply that the given hypothesis accounts for the old evidence, but rather that the hypothesis is the *only* available hypothesis that accounts for the evidence. In the case of General Relativity (GR) and the anomalous perihelion of Mercury (M), the crucial fact in virtue of which GR received such strong confirmation was that it accounted for M *and none of its competitors accounted for* M. If there had been serious competing theories that were also capable of adequately explaining M in a plausible way, then the degree of confirmation conferred on GR by M would intuitively have been far weaker (and possibly negligible). Of course, there were some competing hypotheses. Le Verrier hypothesized the existence of an unobserved planet 'Vulcan' whose gravitational attraction would have explained Mercury's advanced perihelion, and Von Seeliger postulated a ring of particulate matter around the sun (see, for example, Crelinsten [2013]) for the same reason. But these hypotheses were not serious competitors to GR at the time of Einstein's derivation. It was the absence of any satisfactory explanation of M that led scientists to regard the phenomenon as a mystery that needed solving, and it was in virtue of the mystery surrounding M that GR's explanation of M carried such confirmatory significance. Furthermore, it is clear that if, in establishing that a hypothesis H explains some old evidence E, we also demonstrate that a host of alternative and independently viable hypotheses also explain E, then H will not generally receive significant confirmation. For example, imagine that H belongs to a large and heterogeneous class of mutually incompatible theories S, and I prove that all of the theories in S adequately explain some old evidence E. My proof will not do much to increase our confidence in H in this case, since it does nothing to distinguish H from its competitors. Overall, then, what really matters in cases of confirmation by old evidence is the realization that H does a better job than its competitors do of explaining the existing evidence.

This point has been overlooked by all of the existing Garber-style solutions, which assume that confirmation by old evidence is based only on the realization that the relevant hypothesis adequately explains the old evidence. We contend that this realization only has meaningful evidential import when it is combined with the realization that the hypothesis' competitors do *not* adequately explain the old evidence. Theory selection is, by its very nature, a competitive game. And the explanatory virtues of a theory are largely irrelevant if they do not serve to distinguish that theory from its competitors.

The preceding observations suggest that it may be more fitting to think of scientists as being concerned primarily with propositions of the form $A =_{def}$ 'H is the *only* available hypothesis that adequately accounts for E.' What leads a scientist to increase their credence in a hypothesis H is the realization that A is more plausible than they had previously thought. Thus, we can dispense with HF's propositions X and Y altogether, and model confirmation by old evidence purely in terms of changes in the probability of A. When Einstein derived M from GR, the scientific community became more confident that GR was the *only* way to account for M, and this is what led to GR's confirmation. In order to model this interpretation of the story, we need only one minimal constraint—namely (assuming that $P(\mathrm{E}) = 1$),

A1. $P(\mathrm{H} \mid \mathrm{A}) > P(\mathrm{H} \mid \neg\mathrm{A})$.

A1 stipulates that H is more likely to be true on the assumption that it is the only available hypothesis that accounts for the old evidence E than on the assumption that it is *not* the only available hypothesis that accounts for E. The intuition behind A1 is both simple and powerful. If we additionally assume that E must be explained by *some* hypothesis, then we can plausibly motivate significantly stronger conditions, but A1 is all that is needed for present purposes.

Now, the idea is to model cases of confirmation by old evidence as cases in which the agent becomes more confident that A is true without necessarily becoming certain of the truth of any individual proposition. Following Bayesian orthodoxy, we can model this by assuming that the agent Jeffrey-conditionalizes to increase the probability of A, that is, they replace their prior credence function P with the new credence function P^* defined by

$$P*(\mathrm{H}) \;=\; P(\mathrm{H} \mid \mathrm{A})P*(\mathrm{A}) + P(\mathrm{H} \mid \neg\mathrm{A})P*(\neg\mathrm{A}) \tag{2}$$

So, the scientist considers only the propositions corresponding to the relevant hypotheses and items of evidence, and the propositions of the form A (which they treat as atomic). Over time, they become more or less confident in A, and they update their credences accordingly (via Jeffrey-conditionalization). From A1, it straightforwardly follows that, when the scientist become more confident in A, H is thereby confirmed: that is, $P^*(\mathrm{H}) > P(\mathrm{H})$.[4] This is an extremely simple model of confirmation by old evidence that relies only on the basic premise that scientists are concerned primarily with what distinguishes a hypothesis from its competitors.

It should be noted that the model proposed here is perfectly continuous with the amended model given by HFT*. That model also has the implication that learning the proposition 'H is the only available hypothesis that adequately explains E' confirms H: that is, $P(\mathrm{H} \mid \mathrm{X} \wedge \neg\mathrm{Y}) > P(\mathrm{H})$.[5] The main difference between the models is that, in our model, this is the *only* confirmatory relation that is assumed. In a sense, one can think of our model as a coarse-graining of the amended HF model in which we model only those aspects of the scientist's cognitive state that are necessary to explicate the origin of confirmation by old evidence. Since, in many cases, all that matters is whether H is the *only* available explanation of the old evidence, we can treat the corresponding proposition A as atomic and 'throw away' the finer detail given by the propositions X and Y. Of course, one could equally well model confirmation by old evidence in terms of the amended HF model, which, as we have argued, is also based on plausible conditions. But we contend that, in the salient examples that have been considered in the literature, the extra structure imposed by the HF model is redundant and can be disregarded without loss.

Furthermore, it should be reiterated that one could reasonably take issue with the claim that HF's individual propositions X and ¬Y should always confirm H. As we

---

[4] *Proof.* Using the law of total probability, we note that $P(\mathrm{H}) = P(\mathrm{H} \mid \mathrm{A})\, P(\mathrm{A}) + P(\mathrm{H} \mid \neg\mathrm{A})\, P(\neg\mathrm{A})$. Next, using A1 and eq. (1), one then obtains $P^*(\mathrm{H}) - P(\mathrm{H}) = (P(\mathrm{H} \mid \mathrm{A}) - P(\mathrm{H} \mid \neg\mathrm{A})) \cdot (P^*(\mathrm{A}) - P(\mathrm{A})) > 0$.∎

[5] Note that this inequality is equivalent to A1 because $P(\mathrm{H} \mid \mathrm{X}, \neg\mathrm{Y}) > P(\mathrm{H})$ is equivalent to $P(\mathrm{H} \mid \mathrm{A}) > P(\mathrm{H})$ (as $\mathrm{A} \equiv \mathrm{X} \wedge \neg\mathrm{Y}$), which is equivalent to $P(\mathrm{H} \mid \mathrm{A}) > P(\mathrm{H} \mid \neg\mathrm{A})$— i.e. to A1.

stressed above, the realization that H accounts for E might have no confirmatory significance if we are in a situation in which E has already been well-explained by all of H's serious competitors. Likewise, learning that none of H's competitors can explain E might have no confirmatory significance if we already know that H also fails to explain E. If we want to capture these kinds of situations faithfully, then the amended HF model cannot be correct, since, by HFT*, it ensures that both X and ¬Y always confirm H. Our model assumes only the much weaker condition A1, which does not imply that X and ¬Y individually confirm H, but only that their conjunction does so.

The solution to POE defended here is closely related to recent work on the so-called 'no alternatives argument' (NAA). Roughly, the NAA has the following schematic form:

Premise 1. Hypothesis H has some desirable features F.

Premise 2. Despite significant effort, the scientific community has been unable to find any alternatives to H that share those desirable features F.

Conclusion. Hence, we have at least one good reason in favour of H.

Dawid, Hartmann, and Sprenger [2015] provide a Bayesian analysis of the NAA and identify conditions under which the premises of the argument (especially Premise 2) can provide legitimate confirmatory support for H. Note that the conjunction of the premises states (roughly) that H is the only available hypothesis with the relevant desirable features F. So, in the special case in which F refers to the ability to adequately explain some existing body of old evidence, the conjunction of the premises corresponds to the proposition A. Thus, our assumption A1 (according to which A confirms H) can be seen as a particular instantiation of the conclusion of the NAA. Indeed, we have argued here that all of the paradigmatic examples of confirmation by old evidence discussed in the literature can typically be understood in this way.

4. Conclusion

In summary, we've presented a novel solution to POE, based on the idea that a hypothesis is confirmed when it provides the *only* available explanation of old evidence. We've seen that this approach has a number of major advantages over its competitors: Unlike previous attempts to fill in the gaps of Garber's conditionalization model, our model does not rely on dubious constraints on the relationship between the logical and empirical parts of the language. And, unlike counterfactual style approaches, we do not require any radical revisions to the classical subjective Bayesian interpretation of probability. Furthermore, the model presented here is extremely simple, and it models only those aspects of a scientist's doxastic state that are necessary to plausibly account for confirmation by old evidence.

It should be noted that our model depends crucially on the assumption that scientists believe the proposition A with a certain explicit probability and that they change this probability over time as they alter their assessments of the theoretical landscape (for example, by learning that H accounts for E, or that one of its competitors does not). This is an empirical assumption, and it has to be investigated in detail whether it is true or false for concrete cases.[6]

[6] The authors would like to thank two anonymous referees for their helpful and constructive input on previous drafts of this article. We would also like to thank Branden Fitelson, Colin Howson, Jan Sprenger, and Reuben Stern for stimulating conversations on the problem of old evidence.

Disclosure Statement

No potential conflict of interest was reported by the author(s).

Funding

This research was generously supported by the Alexander von-Humboldt Foundation and the LMU Centre for Advanced Studies.

ORCID

Benjamin Eva http://orcid.org/0000-0002-3764-3398

References

Crelinsten, J. 2013. *Einstein's Jury: The Race to Test Relativity*, Princeton: Princeton University Press.

Dawid, R., S. Hartmann, and J. Sprenger 2015. The No Alternatives Argument, *The British Journal for the Philosophy of Science* 66/1: 213–34.

Earman, J. 1992. *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*, Cambridge, MA: The MIT Press.

Eells, E. 1990. Bayesian Problems of Old Evidence, in *Scientific Theories (Minnesota Studies in the Philosophy of Science, Vol. XIV)*, ed. C.W. Savage, Minneapolis: University of Minnesota Press: 205–23.

Garber, D. 1983. Old Evidence and Logical Omniscience in Bayesian Confirmation Theory, in *Testing Scientific Theories (Minnesota Studies in the Philosophy of Science, Vol. X)*, ed. J. Earman, Minneapolis: University of Minnesota Press: 99–131.

Glymour, C. 1980. Why I Am Not a Bayesian, in his *Theory and Evidence*. Princeton: Princeton University Press: 63–93.

Hartmann, S. and B. Fitelson 2015. A New Garber-Style Solution to the Problem of Old Evidence, *Philosophy of Science* 82/4: 712–17.

Howson, C. 1991. The 'Old Evidence Problem', *The British Journal for the Philosophy of Science* 42/4: 547–55.

Howson, C. 2017. Putting on the Garber-Style? Better Not, *Philosophy of Science* 84/4: 659–76.

Jaynes, E. 1968. Prior Probabilities, *Institute of Electrical and Electronic Engineers Transactions on Systems Science and Cybernetics* 4/3: 227–41.

Jeffrey, R. 1983. Bayesianism with a Human Face, in *Testing Scientific Theories (Minnesota Studies in the Philosophy of Science, Vol. X)*, ed. J. Earman, Minneapolis: University of Minnesota Press: 133–56.

Kao, M. 2018. Old Evidence in the Development of Quantum Theory, *Philosophy of Science* 85/1: 126–43.

Rosenkrantz, R. 1981. *Foundations and Applications of Inductive Probability*, Atascadero, CA: Ridgeview Publishing.

Schwan, B. and R. Stern 2017. A Causal Understanding of When and When Not to Jeffrey Conditionalize, *Philosopher's Imprint* 17/8: 1–21.

Seidenfeld, T. 1979. Why I Am Not an Objective Bayesian: Some Reflections Prompted by Rosenkrantz, *Theory and Decision* 11/4: 413–40.

Sprenger, J. 2015. A Novel Solution to the Problem of Old Evidence, *Philosophy of Science* 82/3: 383–401.

Williamson, J. 2010. *In Defence of Objective Bayesianism*, Oxford: Oxford University Press.

Appendix: Proof of Theorem HFT*

We parametrize the probability distribution P over the binary propositional variables H, X, Y in the following way: $x := P(\mathrm{X})$, $p := P(\mathrm{Y} \mid \mathrm{X})$, $q := P(\mathrm{Y} \mid \neg\mathrm{X})$, $\alpha := P(\mathrm{H} \mid \mathrm{X}, \mathrm{Y})$, $\beta := P(\mathrm{H} \mid \mathrm{X}, \mathrm{Y})$, $\gamma := P(\mathrm{H} \mid \mathrm{X}, \mathrm{Y})$, and $\delta := P(\mathrm{H} \mid \mathrm{X}, \mathrm{Y})$. Hartmann and Fitelson [2015] have shown that HF1: $\beta > \delta$, HF2: $\beta > \gamma$, HF3: $\alpha > \gamma$, and HF4: $\alpha \geq \delta$ jointly imply that $P(\mathrm{H} \mid \mathrm{X}) > P(\mathrm{H})$. Proceeding in the same way, one can show that HF5: $\beta > \alpha$, HF6: $\delta > \gamma$, HF7: $\beta > \gamma$, and HF8: $\delta \geq \alpha$ jointly imply that $P(\mathrm{H} \mid \neg\mathrm{Y}) > P(\mathrm{H})$. Comparing the two sets of conditions, one sees that HF7 ≡ HF2 and that HF4 and HF8 are only compatible if one replaces them by HF4*: $\alpha = \delta$. Then HF5 ≡ HF1 and HF6 ≡ HF3. Hence, HF1–HF3 and HF4* imply $P(\mathrm{H} \mid \mathrm{X}) > P(\mathrm{H})$ and $P(\mathrm{H} \mid \neg\mathrm{Y}) > P(\mathrm{H})$.

Finally, we show that HF1–HF3 and HF4* also jointly imply that $P(\mathrm{H} \mid \mathrm{X}, \neg\mathrm{Y}) > P(\mathrm{H})$—that is, that $\beta > \alpha x p + \beta x \bar{p} + \gamma \bar{x} q + \delta \bar{x} \bar{q}$. Setting $\alpha = \delta$ (HF4*), this inequality is equivalent to the inequality $(\beta - \delta) p \cdot x + (\beta - \gamma q - \delta \bar{q}) \cdot \bar{x} > 0$, which holds as $\beta > \delta$ (HF1) and $\beta > \gamma q + \delta \bar{q}$ because $\beta > \gamma, \delta$ (HF1, HF2).∎