Philosophy of Science (2022), **89**, 252–267
doi:10.1017/psa.2021.28

PSA PHILOSOPHY OF SCIENCE ASSOCIATION

PHILOSOPHY OF SCIENCE

ARTICLE

Against Probabilistic Measures of Explanatory Quality

Marc Lange

Department of Philosophy, The University of North Carolina at Chapel Hill, Chapel Hill, NC, US.
Email: mlange@email.unc.edu

(Received 30 May 2020; revised 24 August 2020; accepted 26 February 2021; first published online 10 February 2022)

Abstract

Several philosophers propose probabilistic measures of how well a potential scientific explanation would explain the given evidence. These measures could elaborate "best" in "inference to the best explanation". This paper argues that none of these measures (and no other measure built exclusively from such probabilities) succeeds. The paper considers the various rival explanations that scientists proposed for the parallelogram of forces. Scientists regarded various features of these proposals as making them more or less "lovely" (in Lipton's sense). None of these probabilistic measures of loveliness can reflect these features. The paper concludes by considering the kinds of probabilities that could reflect these features.

I. Introduction

"Inference to the Best Explanation" (IBE) has long been recognized as concerned not solely with *inferences* from evidence to hypotheses, but more broadly with explanatory considerations making hypotheses more or less credible. In particular, IBE is the view that "hypotheses are to be assessed at least partly on the basis of their explanatory virtues" (Douven 2017, 7), that is, on how *well* or *badly* they explain the facts that constitute our evidence.

In referring to a hypothesis's "explanatory goodness" (Douven 2017, 11), IBE's advocates do not intend to presume that the hypothesis does in fact explain why the evidence obtains. The hypothesis may turn out not even to be true, and a falsehood explains nothing (except as an idealization, approximation, explanation sketch, or the like). By the hypothesis's "explanatory goodness", IBE's friends mean how well the putative explanation supplied by the hypothesis explains the given evidence *if* the "potential explanation" (Lipton 2001, 97; Schupbach and Sprenger 2011, 107) appealing to the hypothesis turns out to be a genuine explanation.[1] IBE is thus the idea that a hypothesis's explanatory goodness—what Lipton (2004, 59) calls its "loveliness"—can (and often does) contribute to the hypothesis's credibility.

[1] For simplicity, I will (as others do) sometimes say "explanation" where I mean "potential explanation." The slogan "inference to the best explanation" obviously employs this shorthand.

Lipton (2001, 93-94,105) emphasizes that the hypothesis's loveliness must be distinguished from its likeliness (i.e., its all-things-considered plausibility), since otherwise IBE would be the triviality that our epistemic justification for our degree of confidence in a given hypothesis can be its overall plausibility.

IBE's friends (such as Lipton 2001, 105) readily acknowledge that they owe us an account of the "B" in "IBE", that is, of loveliness as "a measure of how good a potential explanation is" (Lipton 2001, 119). There is currently an active research program to elaborate, in purely probabilistic terms, what makes one potential explanation better than another. For example, Lipton (2004), McGrew (2003), and Okasha (2000) have proposed that h_1's being lovelier than h_2 in its potential explanation of e entails that either $\mathrm{pr}(h_1) > \mathrm{pr}(h_2)$ or $\mathrm{pr}(e|h_1) > \mathrm{pr}(e|h_2)$. Okasha (2000, 73) writes: "The correct way of representing IBE, I suggest, views the goodness of explanation of a hypothesis *vis-à-vis* a piece of data as reflected in the prior probability of the hypothesis (P(H)) and the probability of the data given the hypothesis (P(e|H)). The better the explanation, the higher is one or both of these probabilities."[2] Some philosophers have proposed more precise measures (in strictly probabilistic terms) of how well h explains e. Among those most frequently discussed are

i. $\dfrac{\mathrm{pr}(h|e) - \mathrm{pr}(h| \sim e)}{\mathrm{pr}(h|e) + \mathrm{pr}(h| \sim e)}$ (Schupbach and Sprenger 2011).

ii. $\dfrac{\mathrm{pr}(e|h) - \mathrm{pr}(e)}{1 - \mathrm{pr}(e)}$ if $\mathrm{pr}(e|h) \geq \mathrm{pr}(e)$, $\dfrac{\mathrm{pr}(e|h) - \mathrm{pr}(e)}{\mathrm{pr}(e)}$ otherwise (Crupi and Tentori 2012).

iii. $\ln\left[\dfrac{\mathrm{pr}(e|h)}{\mathrm{pr}(e)}\right]$ (Good 1960; McGrew 2003).

Arguments concerning the adequacy of one or another of these measures have been the subject of considerable discussion. (Along with the references already cited, see also Cohen 2016, 2018; Eva and Stern 2019; Schupbach 2017; Douven and Schupbach 2015a, 2015b; and Sprenger and Hartmann 2019, 185-205; among others.)[3]

The purpose of this paper is to argue that none of these measures—indeed, no measure built exclusively from probabilistic elements such as $\mathrm{pr}(h|e)$, $\mathrm{pr}(e)$, and $\mathrm{pr}(h)$—can capture explanatory quality in the sense of "B" in "IBE". Scientific practice recognizes various factors contributing to (or detrimental to) explanatory quality that cannot be expressed in combinations of such probabilities. My argument for this negative conclusion will appeal to an example of an extended controversy in the history of science over the explanation of a given fact: the parallelogram law for the composition of forces. Various potential explanations of this law have been

[2] Lange (forthcoming) examines and criticizes this approach to explanatory quality.

[3] The purpose of any of these measures is not to determine whether some proposal is explanatory or not. As all of these authors (e.g., Schupbach and Sprenger 2011, 107) emphasize, it would be no objection to a given measure that a hypothesis scores high on it but is not explanatory. (For instance, e cannot explain e, so it is no objection to Schupbach and Sprenger's measure that it is maximized when e is inserted as h.) The purpose of these measures is not to say whether a given hypothesis (if true) explains e. Rather, the purpose of these measures is to measure *how well* a given hypothesis explains e, *if* it does in fact explain e. That is, the purpose of these measures is not to reveal the nature of explanation, but rather to capture the quality of the explanation that a hypothesis would supply if it did supply an explanation.

https://doi.org/10.1017/psa.2021.28 Published online by Cambridge University Press

proposed, and various scientists have helpfully identified various features of these potential explanations as affecting their quality. But none of these features makes any difference to the probabilities figuring in the proposed measures of explanatory goodness. Rather, according to any of the above measures, all of these potential explanations automatically receive the same score. No measure in purely probabilistic terms is sensitive to the features that have been widely regarded as making these potential explanations more or less lovely.

In section 2, I will distinguish several ways in which the proposed measures of loveliness might be thought to function. The measures are more plausibly asked to play some of these roles than others, but whichever role they are called upon to play, they will encounter the problems that I will describe in sections 3 and 4. In section 3, I will specify the features that scientists have regarded as enhancing or as detracting from the loveliness of various potential explanations of the parallelogram law. In section 4, I will show that none of the proposed measures of loveliness reflects these features and that no algebraic combination of probabilities like pr($h|e$), pr(e), and pr(h) can do so. I will conclude in section 5 by arguing that a *richer* set of probabilities may enable us to give a necessary condition for a factor to enhance a potential explanation's loveliness. Those probabilities would include the probability that a given argument is *explanatory* given that its premises and conclusion are true.

2. Interpreting measures of loveliness

The various proposed measures of loveliness that I am critiquing could be interpreted in various ways. A measure could be proposed as capturing loveliness all-things-considered or instead as capturing only one contribution to loveliness. (Loveliness all-things-considered may be higher or lower than loveliness as influenced only by the single contribution being measured.) In addition, a measure could be proposed as applying (i.e., as capturing whatever it is supposed to capture) in all cases or instead as applying only under certain conditions. Perhaps different authors have (or the same author on different occasions has) different aims in proposing measures of loveliness. Let us briefly consider some of these options.

The view that I quoted Okasha (2000) as defending seems intended to capture loveliness all-things-considered in all cases. The view's motivation seems to be that in order for loveliness to have a confirmatory impact within a strictly Bayesian account of confirmation, h's loveliness in explaining e must have an impact either on pr(h) or on pr($e|h$); there is no other factor involving h in the formula for Bayesian updating, so there is nowhere else for h's loveliness to have an impact. Okasha concludes that if h_1 is lovelier than h_2 in its explanation of e, then either $\text{pr}(h_1) > \text{pr}(h_2)$ or $\text{pr}(e|h_1) > \text{pr}(e|h_2)$.

However, this conclusion does not follow from Okasha's premises. Even if pr(h) and pr($e|h$) are the only places where loveliness can have an impact on confirmation, loveliness may not be the only consideration having an impact. If pr(h) [or pr($e|h$)] reflects not only h's loveliness but other considerations as well, then h_1 can have greater loveliness than h_2 without pr(h_1) exceeding pr(h_2) [or pr($e|h_1$) exceeding pr($e|h_2$)]. Loveliness may contribute toward raising pr(h_1) [or pr($e|h_1$)] over pr(h_2) [or pr($e|h_2$)], but its contribution may be outweighed by other considerations pushing in the opposite direction.

https://doi.org/10.1017/psa.2021.28 Published online by Cambridge University Press

Generally, a given measure of loveliness is portrayed by its advocates as capturing what makes one potential explanation better than another in the sense employed by IBE. In that case, even if different measures apply in different cases, a given measure must (when it applies) capture loveliness all-things-considered. Douven (2017), for instance, thinks that there may well be several correct measures of explanatory goodness, each applying in its own separate range of cases. However, Douven (2017, 11, 13–15) also maintains that when a measure (like the one proposed by Schupbach and Sprenger 2011) is applicable, it captures "B" in "IBE"; it measures whether one hypothesis (if it were an explanation) would explain *e* better than another hypothesis would (if it were an explanation). So when the measure applies, it must capture explanatory goodness all-things-considered, not merely a single contributor to explanatory goodness that must be combined with (and could be outweighed by) others to yield total explanatory goodness.

Schupbach and Sprenger (2011) present their measure as measuring "B" in "IBE" and so, when it applies, as capturing loveliness all-things-considered. Referring to the "explanatory power" that their measure is supposed to capture, Schupbach and Sprenger (2011, 106) write: "Humans regularly make judgments of explanatory power and then use those judgments to develop preferences for hypotheses or even to infer outright to the truth of certain hypotheses. Much of human reasoning ... makes use of judgments of explanatory power".[4] Schupbach and Sprenger (2011, 109) present their measure as expressing the idea that "a hypothesis offers a powerful explanation [of some fact] ... to the extent that it makes that [fact] less surprising." They (2011, 107) "take no position on whether [their] analysis captures the notion of explanatory power generally" (although their paper is entitled "*The* Logic of Explanatory Power"). So apparently, they allow for the possibility that their measure, though capturing loveliness all-things-considered when it is applicable, is applicable only in a special range of cases.

However, they also emphasize that their measure applies in a wide range of cases: "our account captures at least one familiar and epistemically compelling sense of explanatory power that is common to human reasoning" (2011, 107). They mathematically derive their measure from some "conditions of adequacy" for a measure of explanatory power, and they present the apparent plausibility of those "adequacy conditions" as showing that the measure applies widely: "in the wide range of cases in which our conditions of adequacy are rationally compelling and intuitively applicable, one ought to think of explanatory power in accord with" their measure (2011, 117-18; cf. Schupbach 2017, 43). I will argue that there are many sources of loveliness other than the single dimension to which Schupbach and Sprenger's measure responds and that these other sources commonly come into play in scientific practice. I will argue further that these other dimensions of loveliness cannot be captured by any measure that restricts itself to probabilities such as pr(e), pr($h|e$), and pr($h|\sim e$). The first "adequacy condition" from which Schupbach and Sprenger (2011, 109) derive their measure is that an adequate measure must be capable of being

[4] Perhaps, however, this passage is not intended to suggest that their measure captures loveliness all-things-considered whenever it applies, but rather that it captures one loveliness-enhancing consideration that is "regularly", but not always, the sole factor determining loveliness. I will turn to this interpretation below.

https://doi.org/10.1017/psa.2021.28 Published online by Cambridge University Press

represented as a function of these probabilities. (Sprenger and Hartmann [2019, 193] impose the same condition.) Thus, their "adequacy conditions" preclude any "adequate" measure from being sensitive to (for example) precisely those features of the parallelogram law's potential explanations that scientists have regarded as making those potential explanations more or less lovely.

Schupbach (2017, 48) shows that on Schupbach and Sprenger's measure, h_1's loveliness exceeds h_2's with respect to some common e if and only if $\mathrm{pr}(e|h_1) > \mathrm{pr}(e|h_2)$. Therefore, by Bayesian conditionalization, as long as h_1's prior probability is no less than h_2's, h_1's posterior probability will exceed h_2's if h_1's measure exceeds h_2's. But this consequence seems too strong if Schupbach and Sprenger's measure applies only in certain conditions, since in conditions where that measure does *not* capture loveliness all-things-considered, the result that Schupbach has derived still holds. That is, h_1's posterior probability still exceeds h_2's (as long as h_1's prior is no less than h_2's) if h_1's measure exceeds h_2's, even where the given measure *fails* to capture loveliness all-things-considered. This seems too strong.

Alternatively, suppose instead that Schupbach and Sprenger's measure captures only one contribution to the potential explanation's loveliness (reflecting the degree to which the putative explainer would make the explained fact less surprising), where this contribution may be accompanied by other contributions to loveliness or other theoretical virtues besides loveliness (Schupbach 2017, 41-42). Then once again, Schupbach's derivation seems to prove too much. It should not turn out that h_1's posterior exceeds h_2's (as long as h_1's prior is no less than h_2's) as long as h_1 possesses more of a single dimension of loveliness than h_2 does, regardless of any other dimensions of loveliness and any other theoretical virtues that make themselves felt through $\mathrm{pr}(e|h)$ rather than through the prior. Room should be left for these other considerations (which may favor h_2 over h_1) to influence the posteriors, perhaps even outweighing loveliness.

In the next section, I will discuss a particular example from the history of science where scientists explicitly discussed some of the factors that contribute to or detract from certain potential explanations' loveliness. I will then argue that these factors cannot be captured by any measure of explanatory quality like those we have been examining. Of course, advocates for those measures could try to find a non-*ad-hoc* way of carving out a range of applicability for a given measure that excludes cases like the one I will discuss below.[5] But that sort of case is not uncommon in science. That the

[5] Schupbach and Sprenger (2011) focus on explaining contingent events by contingent events, whereas the example on which I will focus in the next section involves the explanation of a law by laws. But a measure of explanatory quality that is inapplicable to explanations of laws by laws is a narrow measure indeed. Schupbach and Sprenger (2011, 109) restrict their proposed measure to contingent propositions h and e. But this restriction does not motivate restricting their measure so as not to apply when h and e are natural laws. The motivation for their restriction to contingent h and e seems to be to avoid running into undefined probability values (as would happen if e were a necessity and $\mathrm{pr}(h|{\sim}e)$ appeared in the measure, since if e is a necessity, then $\mathrm{pr}({\sim}e) = 0$). This rationale does not motivate restricting their proposed measure so as not to apply when h and e are natural laws. Although the laws of nature are generally thought to possess a variety of necessity ("physical necessity"), a probability function is not required to assign unity to a law of nature, unlike the requirement that it assign unity to a logicomathematical truth. Rational agents whose degrees of belief are represented by probability functions are *logically* omniscient, not *nomologically* omniscient.

https://doi.org/10.1017/psa.2021.28 Published online by Cambridge University Press

factors I have identified make a difference to loveliness in these cases and cannot be captured by the measures I have discussed shows that those measures apply more narrowly than they may initially appear to do and that the "adequacy conditions" motivating these measures are less "compelling conditions" (Schupbach 2017, 43) than they may at first appear to be.

3. What makes the parallelogram law's potential explanations more (or less) lovely?

In classical physics, a force applied at a point can be represented by an arrow starting from that point, pointing in the force's direction, and having a length proportional to the force's magnitude. The resultant of forces *F* and *G* acting together at a point is the force represented by the arrow extending from that point to form the diagonal of a parallelogram whose adjacent sides represent *F* and *G*. Accordingly, this principle is frequently called the "parallelogram law" for the composition of forces.

This law was introduced in 1586 by Simon Stevin. It seems to have been widely recognized by Newton's day, since both Pierre Varignon and Bernard Lamy stated it in the same year (1687) as Newton did in the *Principia*.[6] But long after the parallelogram law's *truth* had become uncontroversial, considerable dispute remained over *why* it holds. Rival proposed explanations were developed and criticized by many notable scientists over the course of the eighteenth and nineteenth centuries. My concern will be some of the considerations that have been widely regarded as contributing to or detracting from the loveliness of one or another of these potential explanations.

Potential explanations of the parallelogram of forces fall into three main classes.[7] First, there is the dynamical approach that is commonly attributed to Newton. This approach applies Newton's second law of motion (force = mass x acceleration) to the component accelerations and net acceleration produced individually and collectively (respectively) by the two forces being composed. As a matter of geometry, component displacements compose parallelogramwise, and from this fact, it follows that component velocities and component accelerations also do so. Newton's second law links each component force to the component acceleration that it causes. Since (by Newton's second law) the resultant force is in the direction of and proportional to the resultant acceleration and since the component accelerations compose parallelogramwise, the component forces must do so too.

This approach remained popular throughout the eighteenth and nineteenth centuries. As with all of the parallelogram law's potential explanations that I will be describing, the soundness of this derivation of the parallelogram law was completely uncontroversial. The controversy concerned whether or not this derivation is an explanation. The considerations that scientists regarded as making it (or its rivals) more or less lovely are my concern.

One consideration that some scientists regarded as helping to make the dynamical approach lovelier is that it purports to explain the parallelogram law by deriving it

[6] For historical background, see Dugas (1988) and Duhem (1905-6/1991).

[7] This summary of the most widely endorsed potential explanations draws upon Lange (2010, 2016), which not only gives many references to notable scientists endorsing and criticizing various potential explanations, but also gives the precise steps of the three derivations that I will be describing.

https://doi.org/10.1017/psa.2021.28 Published online by Cambridge University Press

from the forces' power to cause accelerations (as given by Newton's second law of motion). Critics of the dynamical approach, however, emphasized that Newton's second law introduces mass into the law's derivation. All of the "m"s thereby introduced end up eventually cancelling one another out so that ultimately, of course, none figures in the parallelogram law. This feature of the derivation was widely cited as detracting from this potential explanation's loveliness. That mass enters the derivation only to be eliminated later was regarded by the dynamical approach's critics as evidence that mass (the constant of proportionality between force and its effect on motion) has no place in the parallelogram law's explanation – that the law does not arise from dynamical considerations, but rather from statics alone. That is, the law is explained by what force is required to balance a pair of component forces, not by what motion an unbalanced force would cause.

This became a point of heated controversy. For instance, whereas critics such as William Whewell (1858, 226) called dynamical considerations (such as masses) "extraneous" and John Robison (1822, 64) called them "gratuitous" to the parallelogram law, one proponent (A.H. 1848, 107) of the dynamical explanation (after criticizing Whewell's proposed explanation (just below) as "forced and unnatural") praised the dynamical explanation for unifying the parallelogram of forces with the parallelogram of velocities.[8] My aim here, of course, is not to settle this once-lively scientific controversy. Rather, it is to identify the features of various potential explanations that were widely regarded as enhancing or diminishing their loveliness and then to examine (in the next section) whether the proposed measures of explanatory quality (discussed in the previous sections) are sensitive to these features.

Neither of the other two main approaches to explaining the parallelogram law appeals to the connection between force and motion. The approach most commonly advocated in the mid-nineteenth century originated with Duchayla (1804; see Lange 2010, 404–8). It exploits the "principle of the transmissibility of force": that when a force acts on a body, the result is the same whatever the point, rigidly connected to the body, to which the force is applied, provided that the line through that point and the force's actual point of application lies along the force's direction. From this principle, Duchayla derives the parallelogram law for the resultant force's direction. Then he uses that result, in turn, to derive the parallelogram law for the resultant force's magnitude.

Its advocates regarded this putative explanation as lovely ("very simple and beautiful", in the words of one textbook [Mitchell, Young, and Imray 1860, 47; cf. Lange 2010, 407]) partly by virtue of making no appeal to force's causal powers. By contrast, critics of this approach regarded it as unlovely partly because it does not give the same explanation for the parallelogram law's holding for direction as it does for the parallelogram law's holding for magnitude. Although it derives both from the same ultimate premises, it arrives first at the parallelogram law's holding for direction and then uses that result to derive that the parallelogram law also holds for magnitude. Its detractors saw the derivation as "essentially artificial" (Goodwin 1849, 273) by virtue of its failing to treat these two parts of the parallelogram law alike (Lange 2010, 407–8). It fails to derive them together through the same steps.

[8] For the full passages just quoted and many others along similar lines, see Lange 2010, 402–4.

https://doi.org/10.1017/psa.2021.28 Published online by Cambridge University Press

The third general approach was Poisson's from 1811 (Lange 2010, 408–14), which elaborated a strategy pursued earlier by Foncenex, d'Alembert, and Daniel Bernoulli among others. Instead of using the transmissibility principle, Poisson appeals to symmetries (such as that the composition law must be invariant under rotation), dimensional considerations, and that two forces must have a unique resultant determined entirely by their magnitudes and directions. Its defenders regarded Poisson's potential explanation as lovelier than Duchayla's by virtue of according the same treatment to both magnitude and direction, deriving them together rather than separately.

Its advocates also regarded Poisson's approach as lovelier than Newton's partly by virtue of the fact that the same sort of derivation as Poisson's regarding the composition of forces could also be used to explain the parallelogram laws that hold for various other quantities (such as energy flux densities, heat flows, water flux densities through soils, as well as velocities and accelerations), since they all have the same features to which Poisson's derivation appeals, such as the rotational invariance of their composition. As Maxwell puts it, a Poisson-style derivation "is applicable to the composition of any quantities such that turning them end for end is equivalent to a reversal of their signs" (cited by Lange 2010, 414). These potential explanations do not identify a common explainer of two quantities' both composing parallelogram-wise; one quantity's doing so might be explained by facts about heat, whereas the other's is explained by facts about forces. But the two explanations proceed from analogous premises by analogous steps, which scientists regarded as enhancing their loveliness.

4. These measures of loveliness are inadequate

Are the loveliness-enhancing and loveliness-detracting features of the parallelogram law's potential explanations reflected in the various proposed measures of loveliness mentioned in section 1? Two obstacles block their being so reflected.

The first obstacle is that when Duchayla and Poisson proposed their potential explanations, all of the potential explainers h that they cited (as well as the parallelogram law e) had already been discovered. Therefore, the probabilities $\mathrm{pr}(h|e)$, $\mathrm{pr}(e)$, and so forth that figure in the various measures are all automatically extremal (i.e., 0 or 1). No opportunity remains for the considerations that scientists have widely regarded as affecting these explanations' loveliness to have any impact on these measures.

However, it may well be uncharitable to emphasize this obstacle. Bayesian confirmation theory notoriously encounters the "problem of old evidence" (Glymour 1980, 85–93): if e is already known when h is first proposed, then $\mathrm{pr}(e) = 1$ and so $\mathrm{pr}(h|e) = \mathrm{pr}(h)$; hence, Bayesianism deems e powerless to confirm h, contrary to many episodes where "old evidence" e did confirm h. Somehow, Bayesianism must avoid this result. Accordingly, suppose we grant that Bayesianism can find a rationale for appealing to some probability function that assigns old evidence some non-unitary probability for the purposes of assessing its confirmatory significance. Bayesianism is thereby granted the means of assigning non-unitary probability to the parallelogram law and to all of its potential explainers cited by Duchayla and

https://doi.org/10.1017/psa.2021.28 Published online by Cambridge University Press

Poisson. The various probabilities in the measures will then not automatically be extremal merely because h and e are "old."[9]

It is more difficult to discount the second obstacle keeping the proposed measures of loveliness from reflecting the considerations that scientists have regarded as making lovelier (or less lovely) the parallelogram law's potential explanations. As I noted in the previous section, all of these potential explanations are deductively valid derivations of the parallelogram law e from various other laws h. Hence, for any probability function (even one assigning non-extremal values to the probabilities of "old evidence" e and h), $\mathrm{pr}(e|h) = 1$ and $\mathrm{pr}(\sim e|h) = 0$. Therefore, all three of the rival potential explanations discussed in the previous section are deemed equally lovely under any one of the proposed measures of loveliness mentioned earlier, despite the many important respects in which these explanations differ – respects that scientists have recognized as affecting their loveliness.

For example, under Schupbach and Sprenger's measure, the loveliness of any of these potential explanations equals $\frac{\mathrm{pr}(h|e) - \mathrm{pr}(h|\sim e)}{\mathrm{pr}(h|e) + \mathrm{pr}(h|\sim e)}$, which (by Bayes's theorem) equals

$$\frac{\mathrm{pr}(h)\left[\frac{\mathrm{pr}(e|h)}{\mathrm{pr}(e)} - \frac{\mathrm{pr}(\sim e|h)}{\mathrm{pr}(\sim e)}\right]}{pr(h)\left[\frac{\mathrm{pr}(e|h)}{\mathrm{pr}(e)} + \frac{\mathrm{pr}(\sim e|h)}{\mathrm{pr}(\sim e)}\right]},$$

which (canceling the $\mathrm{pr}(h)$'s and using $\mathrm{pr}(e|h) = 1$ and $\mathrm{pr}(\sim e|h) = 0$, since h entails e) equals $\frac{1/\mathrm{pr}(e)}{1/\mathrm{pr}(e)} = 1$. So all three potential explanations are measured to have equal (indeed, maximal) loveliness. The same holds under Crupi and Tentori's measure, where the loveliness of any of these potential explanations equals $\frac{\mathrm{pr}(e|h) - \mathrm{pr}(e)}{1 - \mathrm{pr}(e)}$ (since $\mathrm{pr}(e|h) \geq \mathrm{pr(e)}$ because $\mathrm{pr}(e|h) = 1$), and so (since $\mathrm{pr}(e|h) = 1$) equals $\frac{1-\mathrm{pr}(e)}{1-\mathrm{pr}(e)} = 1$. All three explanations again have equal loveliness under Good and McGrew's measure, since

$$\ln\left[\frac{\mathrm{pr}(e|h)}{\mathrm{pr}(e)}\right] = \ln[1/\mathrm{pr}(e)] = -\ln \mathrm{pr}(e).$$

None of these measures is sensitive to the features that scientists have emphasized in assessing the quality of these three potential explanations.

Okasha's approach yields the same result. Admittedly, his approach considers not only $\mathrm{pr}(e|h)$, which equals 1 for all three potential explanations, but also $\mathrm{pr}(h)$, which is not extremal for the parallelogram law's potential explainers h (presuming that the problem of old evidence has somehow been circumvented). But I do not see how $\mathrm{pr}(h)$ can reflect the kinds of considerations that scientists have regarded as affecting the loveliness of the parallelogram law's potential explanations. Those considerations go beyond the mere fact that h entails e (which is all that $\mathrm{pr}(e|h)$ reflects) to concern the

[9] Okasha (2000, 705), for example, seems comfortable simply acknowledging that his approach to capturing loveliness presumes that the problem of old evidence has been circumvented so that $\mathrm{pr}(e|h)$ can differ from 1 despite e's being old.

https://doi.org/10.1017/psa.2021.28 Published online by Cambridge University Press

route by which *h* entails *e*. For example, Duchayla's derivation takes a different route for the resultant's magnitude than for its direction, detracting from how well this proposal explains, whereas Poisson's derivation takes the same route for both, enhancing its explanatory power. Its loveliness is also enhanced by its using a route by which (from different but analogous premises) a diverse range of other physical quantities can be shown to obey analogous parallelogram laws. Furthermore, the Newtonian, dynamical explanation takes a route that traces the individual component forces' causal influences, enhancing its loveliness. But the route thereby introduces mass only to have mass ultimately cancel out, making its introduction "gratuitous" and thereby detracting from how well the argument explains.

All of these features concern the route taken to the parallelogram law. These features, therefore, cannot be captured by pr(*h*). After all, the same *h* might be the start of a loveliness-enhancing deductive route to some e_1 and a loveliness-detracting deductive route to some e_2. (Indeed, these could even be different routes from *h* to the same conclusion; e_1 could be e_2.) The same pr(*h*) cannot be high (to capture the former's loveliness) and low (to capture the latter's unloveliness).

No algebraic combination solely of probabilities like pr(*h*) and pr(*e*|*h*) can reflect the features of a given inferential route from *h* to *e*; the same probability values could accompany a route with the loveliness-enhancing features that we have seen or a route with the loveliness-detracting ones. Therefore, no proposal appealing exclusively to such probabilities can measure loveliness. Recall that the first adequacy condition that Schupbach and Sprenger (2011, 109) impose on any measure is that it be a function of pr(*e*), pr(*h*|*e*), and pr(*h*|~*e*). By my reckoning, any "adequate" proposal is doomed to fail.

The loveliness-enhancing and loveliness-detracting features of the parallelogram law's potential explanations call to mind Kitcher's (1989) account of scientific explanation. Admittedly, Kitcher's aim is to identify what makes an argument explanatory, whereas my concern is what makes a potential explanation better. Kitcher maintains that an argument is explanatory by virtue of its argument pattern earning admission into the "explanatory store": the collection of argument patterns possessing the optimal combination of broad coverage of facts, few argument patterns, and stringent constraints on the arguments fitting those patterns. By contrast, I have offered no comprehensive analysis of what makes a potential explanation lovelier.[10] But beyond these differences, there is a fundamental and important similarity. Kitcher's account of explanation (unlike covering-law and statistical-relevance accounts, for example) focuses not merely on the existence of some deductive or probabilistic relation between the explainer and explained, but primarily on the *route* by which the explanatory inference proceeds: the argument's pattern (Kitcher 1989, 430–1). Likewise, I maintain that many of the factors that scientists have widely recognized as affecting the loveliness of the parallelogram law's potential explanations concern the routes taken by those derivations.[11]

[10] I pursue this question in Lange (forthcoming).

[11] Glymour (2015) offers a sharply worded critique of some of the same purported measures of explanatory quality that I criticize. However, Glymour does not make any of the criticisms that I have made. In particular, he does not argue that these purported measures (and any others that are built exclusively from probabilistic elements such as pr(*h*|*e*), pr(*e*), and pr(*h*)) cannot capture the particular sorts of

https://doi.org/10.1017/psa.2021.28 Published online by Cambridge University Press

5. Conclusion

The controversy over the parallelogram law's explanation is not at all unusual. Scientists frequently investigate which facts among those that have already been discovered help to explain a given fact (also already known). A notable example is the longstanding controversy over various rival potential explanations of the Lorentz contraction, time dilation, and other relativistic phenomena (see Lange 2016, 96–149). Initially, Einstein (1905/1989) derived these phenomena from the "principle of relativity" (that the fundamental laws of nature take the same form in all inertial frames) and the "light postulate" (that there is an inertial reference frame where the speed of electromagnetic radiation is independent of the motion of its source). Even though these principles are now well-established, the explanation of these relativistic phenomena remains contested. Some (e.g., Mermin 2009, 185; Brown 2005) have argued that these relativistic phenomena have dynamical explanations appealing to the microforces inside rods and clocks. Others (e.g., Berzi and Gorini 1969; Pal 2003) have instead defended an explanation appealing neither to dynamics nor to electromagnetism, but rather to spacetime symmetries (including the principle of relativity) and the invariance of the spacetime interval. All of these potential explanations are valid deductions of the Lorentz contraction entirely from known laws of nature. Scientists' credences in these potential explanations have been justly influenced by how well (in their view) a given potential explanation would explain. Thus, this example possesses the same crucial features as the parallelogram-law case.

Admittedly, in the philosophical literature, IBE is most often associated with cases where various rival potential explainers are *not* already known to be true – where, for example, we investigate whether or not there is a mouse in the wainscoting by considering the best explanation of scratching sounds and the disappearance of cheese (van Fraassen 1980, 19-20). These cases might suggest that scientists were not using IBE in assessing the merits of the parallelogram law's rival potential

explanatory virtues and vices that scientists have used to argue for or against various potential explanations of the parallelogram law. More generally, Glymour does not identify *any* particular explanatory virtues and vices and then argue that they cannot be captured by the purported measures. So Glymour's critique takes a very different form from mine.

Some of Glymour's criticisms of these purported measures focus on causal explanations of one event by another (Glymour 2015, 595–7). These criticisms do not render mine superfluous, because I have focused instead on explanations of laws by other laws. (These explanations may be causal or non-causal; see Lange [2016].)

One of Glymour's main criticisms of these proposed measures seems to me not as compelling as he takes it to be. Glymour (2015, 594-95) argues that if we apply these proposals to measure the potential explanatory quality of a hypothesis h that we currently believe false, then these proposed measures will unfortunately yield values that are either zero (because $\mathrm{pr}(h) = 0$) or undefined (because $\mathrm{pr}(e|h)$ is undefined). However, as I mentioned in section 2, a given measure of loveliness is generally portrayed by its advocates as capturing what makes one potential explanation better than another in the sense employed by IBE. One important motivation for the project of measuring explanatory quality is to identify the way that IBE is supposed to judge among rival *live* hypotheses. This aim perhaps provides some reason to think that these proposed measures should not be applied to hypotheses that the given epistemic agent regards as already having been ruled out, in which case Glymour's criticism misses the target; it does not matter what happens to these measures when $\mathrm{pr}(h) = 0$. My critique of these measures is not subject to this objection, since my critique does not exploit what happens to these measures when $\mathrm{pr}(h) = 0$.

https://doi.org/10.1017/psa.2021.28 Published online by Cambridge University Press

explanations, since in that case, all of the potential explainers were already known to be true.

I contend, however, that those scientists were indeed using IBE; their credences in the rival potential explanations were guided partly by the loveliness-enhancing and loveliness-detracting features of those explanations.[12] In fact, that these scientists were assessing loveliness is *more* evident precisely because the potential explainers were already known to be true. There is no danger of our conflating loveliness with likeliness by taking the scientists' credences as having been guided by the likeliness that the potential explainers are true rather than by the potential explanations' loveliness. All that could have been at issue among the scientists is which potential explanation would explain best. Advocates (and critics) argued for (or against) a given proposed explanation of the parallelogram law not by arguing for (or against) the truth of the proposed explainers (since the proposed explainers, such as Newton's second law, were already accepted as true by all parties to the dispute). Rather, advocates (and critics) argued for (or against) a given proposed explanation by arguing that a given feature of the proposed explanation would enhance (or detract from) the quality of the proposed explanation, and that the feature thereby makes the proposal more (or less) likely to be a genuine explanation. In using explanatory quality to assess which potential explanation is more likely to be the parallelogram law's genuine explanation, scientists were using IBE.

Cases like the parallelogram law are helpful in that they allow loveliness to be isolated from other grounds for thinking that a given hypothesis genuinely explains a given fact. A given consideration should generally have the same impact on our judgments of a potential explanation's loveliness no matter how likely or unlikely we believe it to be that the potential explanation's explainers actually hold – in particular, should have the same impact when those potential explainers have already been ascertained to hold as when they have not yet been ascertained to hold. (Of course, some potential explanations contain some explainers that have already been ascertained to hold and other explainers that have not yet been ascertained to hold but are considered more or less likely.) Our degree of confidence that the potential explainers hold should make no difference to our judgments of loveliness because loveliness *brackets* likeliness: judgments of loveliness concern the potential explanation's explanatory quality *if it turns out to be an explanation*. IBE requires that loveliness be distinct from likeliness; as Lipton (2001, 97; 2004, 63) says, IBE is the idea that one of the considerations that ought to guide our confidence in a given hypothesis is the quality of the potential explanations it *would* supply. As I mentioned at the start of the paper, Lipton (2001, 93–4,105; 2004, 60,140) emphasizes that IBE would tell us

[12] It is important that this be "credences in the rival potential *explanations*," not "credences in the rival potential *explainers*," since (as I have been emphasizing) all of the potential explainers here (such as Newton's second law on the dynamical approach, the principle of the transmissibility of forces on Duchayla's approach, and so on) were already accepted as true by all parties to this dispute. What the parties disagreed about was which of these *explained* the parallelogram law. Note that in section 1, I said that by a hypothesis's "explanatory goodness" with regard to *e*, IBE's friends mean how well the putative explanation of *e* supplied by the hypothesis explains *e if* the "potential explanation" appealing to the hypothesis turns out to be *a genuine explanation*. I did not say instead "if the 'potential explanation' appealing to the hypothesis turns out to be *true*", since in some cases, a potential explanation could be true (that is, could appeal exclusively to truths) without, in fact, explaining *e*.

https://doi.org/10.1017/psa.2021.28 Published online by Cambridge University Press

little about confirmation if it were the view that our credence in various hypotheses should be guided by how likely we believe those hypotheses are. For IBE not to be empty, loveliness must not be affected by likeliness, and since loveliness is not affected by likeliness, the same factors enhance (or detract from) loveliness whether the hypothesis has already been accepted as true, is deemed likely to be true, or is considered to be a long shot. A conspiracy theory, Lipton (2004, 60) says, illustrates that the potential explanations supplied by some theory can be very lovely while being very unlikely. (A conspiracy theory is lovely by virtue of "showing that many apparently unrelated events flow from a single source and many apparent coincidences are really related," but because there is typically a great deal of other evidence against it, such a potential explanation is typically "very unlikely, accepted only by those whose ability to weigh evidence has been compromised by paranoia.")

What an IBE argument enables some evidence to confirm by some increment (or what an IBE entitles us to conclude, when the evidence is strong enough) is that a given potential explanation is an actual explanation, which requires, in turn, that the relevant potential explainers obtain. Confirmation that it is an actual explanation often yields confirmation of the potential explainers' truth (as in the case of the mouse suspected to reside in the wainscoting). But when those potential explainers have already been accepted as true, then confirmation that they explain is not accompanied by confirmation that they are true.[13] That IBE yields confirmation that the potential explainer *explains*, and only sometimes yields confirmation of the potential explainer's *truth*, is an important but easily overlooked aspect of IBE. As Lipton (2004, 58, my emphasis) says, "According to Inference to the Best Explanation, ... we infer that the best of the available potential explanations *is an actual explanation*." In the case of a causal explanation of some event E, for instance, we could know that a given event C occurred but not yet know whether it helped to cause E. An IBE might confirm that C was a cause of E without confirming that C occurred.[14]

I have argued that frequently there are contributions to the loveliness of some h's putative explanation of e that cannot be captured by any measure restricted to probabilities like $\mathrm{pr}(e)$, $\mathrm{pr}(h|e)$, and $\mathrm{pr}(h|{\sim}e)$. However, a *richer* set of probabilities may enable us to give a necessary condition for a factor to influence a potential explanation's loveliness. Instead of considering the credence of some potential explainer h, let us consider (for example) the credence of x: that Duchayla's 1804 paper gives an actual explanation of the parallelogram law. Let t be that the

[13] That p entails q (as when p is that h explains some accepted fact and q is that h is true) does not require that p's incremental confirmation be accompanied by q's incremental confirmation. That is, Hempel's "special consequence condition" does not hold for incremental confirmation (see Salmon 1975).

[14] In the case of the parallelogram of forces, the controversy was always over what *explains* the parallelogram law, not over the *truth* of the potential explainers. For instance, in arguing against a dynamical explanation of the parallelogram law, Whewell wrote that Newton's second law of motion is "extraneous" (1858, 226) to the parallelogram law and that the parallelogram law "cannot be dependent on the laws of the motions which take place when the forces do not balance" (1832, 88), i.e., Newton's second law. In referring to what the parallelogram law is "dependent" on, Whewell means what *explains* the parallelogram law; according to Whewell, Newton's second law holds but is "extraneous" to the parallelogram law's explanation. Of course, that one argument explains a given fact does not entail that it is that fact's *only* explanation; a fact can have multiple actual explanations. By arguing for his own proposal for explaining the parallelogram law, Whewell was not automatically arguing against the dynamical proposal; he had to argue expressly against it, as he did.

https://doi.org/10.1017/psa.2021.28 Published online by Cambridge University Press

parallelogram law is true and that the potential explainers in Duchayla's (1804) potential explanation of the parallelogram law are also true. Suppose that *f* attributes some feature to the route taken by Duchayla's potential explanation, e.g., that it treats magnitude differently from direction, or that it could be used to derive parallelogram laws for other physical quantities, or that it exploits the formula for the area of a triangle. (Presumably, the first would be loveliness-detracting; the second would be loveliness-enhancing, and the third would have no impact on loveliness.) Then a necessary condition for that feature to enhance (or detract from) the explanation's loveliness is that $pr(x|t,f)$ is greater (less) than $pr(x|t)$. Roughly, the idea behind this necessary condition is that someone who regards a given feature as loveliness-enhancing, knows that some Jones has offered a potential explanation (appealing exclusively to truths), but has not yet ascertained whether Jones's argument possesses this feature, would raise her confidence that Jones's argument is a genuine explanation upon learning that Jones's proposal possesses this feature.[15]

The probabilities figuring in this *necessary* condition are not automatically rendered extremal by the potential explainers being "old evidence." They are also not rendered unitary by the deductive validity of Duchayla's potential explanation. This *necessary* condition is obviously not *sufficient* for *f* to enhance (or detract from) loveliness: if *f* is that an expert testifies to you that a genuine explanation of the parallelogram law takes the route of Duchayla's argument, then $pr(x|t,f) > pr(x|t)$, but the expert's testimony does not contribute to the potential explanation's loveliness.

In a case where the premises of a potential explanation are not already known to be true, some *f* can satisfy this necessary condition despite disconfirming the truth of those explainers. That is because the condition concerns the probability of the potential explainers explaining *given* that they are true. The condition thus nicely respects the distinction between loveliness and likeliness. By appealing to the probability of an *x* that refers to *explanation* over and above truth, this condition clearly avoids a worry that McGrew (2003, 565) expects to be directed by "members of the explanationist camp" at the measures of loveliness that I discussed earlier: "that there is nothing distinctively explanatory left by the time we have flattened out the virtues on a probabilistic plane."[16]

This probabilistic necessary condition for a feature to enhance a potential explanation's loveliness does not aim to do the same job as the measures of loveliness that I have critiqued were supposed to perform. Firstly, those measures (under one interpretation that I discussed in section 2) aimed to capture a potential explanation's loveliness all-things-considered, whereas the probabilistic necessary condition aims to deal only with one factor at a time, not with how those factors should be combined to yield an overall measure. Secondly, the measures that I critiqued were intended to supply necessary and sufficient conditions for one potential explanation to be lovelier (if it is an explanation) than another (if it is an explanation) – or, at least, necessary

[15] Here, any sentences concerning Duchayla's 1804 paper, such as *x*, should be read *de dicto*. Thus, logical omniscience (built into representations of credences as probabilities) does not require, for example, that cr(the potential explanation of the parallelogram law given by Duchayla [1804] is deductively valid) = 1.

[16] For more on how to construe IBE so as to ensure that there is something "distinctively explanatory" left once IBE has been expressed in Bayesian terms, see Lange (forthcoming).

https://doi.org/10.1017/psa.2021.28 Published online by Cambridge University Press

and sufficient conditions for the first to be lovelier in one particular respect than the second. By contrast, the probabilistic condition that I have just given aims only to be necessary, not sufficient, for some feature of a potential explanation to make it lovelier (if it is an explanation). Thirdly, the probabilities in the measures that I critiqued made no reference to explanation and so could be estimated independent of any expressly explanatory considerations. This is obviously not the case with the necessary condition that I have just offered, which involves the probability that some proposal is actually an explanation (given that it possesses certain features and appeals exclusively to truths).

Even if this condition is correct, it does little to reveal *which* features make a potential explanation more (or less) lovely. However, perhaps we should not expect such information from a general probabilistic framework for loveliness. Perhaps loveliness tends to be enhanced by different considerations for scientific theories concerning different subjects. Which attributes help to make a potential explanation lovely in the social sciences (e.g., that it depicts a disparate variety of factors as subtly combining to produce a result qualitatively unlike what any single factor in isolation produces) may well differ sharply from the simplicity that helps to make a potential explanation lovelier in cosmology or elementary particle physics.[17] A probabilistic framework for loveliness would have to accommodate such diversity.

References

A.H. 1848. "Whewell's Mechanics – Last Edition." *The Mechanics' Magazine* 48:103–7.

Berzi, Vittorio, and Vittorio Gorini. 1969. "Reciprocity Principle and Lorentz Transformations." *Journal of Mathematical Physics* 10:1518–24.

Brown, Harvey. 2005. *Physical Relativity*. Oxford: Clarendon.

Cohen, Michael. 2016. "On Three Measures of Explanatory Power with Axiomatic Representations." *The British Journal for the Philosophy of Science* 67:1077–89.

Cohen, Michael. 2018. "Explanatory Justice: The Case of Disjunctive Explanations." *Philosophy of Science* 85:442–54.

Crupi, Vincenzo, and Katya Tentori. 2012. "A Second Look at the Logic of Explanatory Power (With Two Novel Representation Theorems)." *Philosophy of Science* 79:365–85.

Douven, Igor. 2017. "Inference to the Best Explanation: What Is It? And Why Should We Care?" *In Best Explanations: New Essays on Inference to the Best Explanation*, edited by Kevin McCain and Ted Poston, 7–24. Oxford: Oxford University Press.

Douven, Igor, and Jonah Schupbach. 2015a. "The Role of Explanatory Considerations in Updating." *Cognition* 142:299–311.

Douven, Igor, and Jonah Schupbach. 2015b. "Probabilistic Alternatives to Bayesianism: The Case of Explanationism." *Frontiers in Psychology* 6:459.

Duchayla, C.D.M.B. 1804. "Demonstration du Parallèlogramme des Forces." *Bulletin des Sciences par la Société Philomathique de Paris* 4:242–3.

Dugas, René. 1988. *A History of Mechanics*. New York: Dover.

Duhem, Pierre. 1905-6/1991. *The Origins of Statics*. Dordrecht: Kluwer.

[17] It might seem a bizarre use of "lovely" for "loveliness" to be associated with multifactorial complexity. But bear in mind that "loveliness" is not a term from scientific practice; rather, it is Lipton's term for explanatory quality. "Good explanation" is a term from scientific practice, and it is realistic to expect an anthropologist or sociologist to say of a proposal that it would make a "good explanation" of some phenomenon partly because it does not seem like an oversimplification, but rather is multifactorial.

https://doi.org/10.1017/psa.2021.28 Published online by Cambridge University Press

Einstein, Albert. 1905/1989. "On the Electrodynamics of Moving Bodies." In *Collected Papers of Albert Einstein*, vol. 2, trans. Anna Beck, 140–71. Princeton: Princeton University Press.
Eva, Benjamin, and Reuben Stern. 2019. "Causal Explanatory Power." *The British Journal for the Philosophy of Science* 70:1029–50.
Glymour, Clark. 2015. "Probability and the Explanatory Virtues." *The British Journal for the Philosophy of Science* 66:591–604.
Good, I.J. 1960. "Weight of Evidence, Corroboration, Explanatory Power, Information and the Utility of Experiments." *Journal of the Royal Statistical Society* 22 (2):319–31.
Goodwin, Harvey. 1849. "On the Connection between the Sciences of Mechanics and Geometry." *Transactions of the Cambridge Philosophical Society* 8:269–77.
Kitcher, Philip. 1989. "Explanatory Unification and the Causal Structure of the World." In *Minnesota Studies in the Philosophy of Science*, vol. 13, edited by Philip Kitcher and Wesley Salmon, 410–505. Minneapolis: University of Minnesota Press.
Lange, Marc. 2010. "A Tale of Two Vectors." *Dialectica* 63:397–431.
Lange, Marc. 2016. *Because Without Cause: Non-Causal Explanation in Science and Mathematics*. New York: Oxford University Press.
Lange, Marc. Forthcoming. "Putting Explanation Back Into 'Inference to the Best Explanation.'" *Noûs*. https://doi-org.libproxy.lib.unc.edu/10.1111/nous.12349.
Lipton, Peter. 2001. "Is Explanation a Guide to inference?" In *Explanation: Theoretical Approaches and Applications*, edited by Giara Hon and Sam Rakover, 93–120. Dordrecht: Kluwer.
Lipton, Peter. 2004. *Inference to the Best Explanation*, 2nd ed. Abingdon: Routledge.
McGrew, Timothy. 2003. "Confirmation, Heuristics, and Explanatory Reasoning." *The British Journal for the Philosophy of Science* 54:553–67.
Mermin, N. David. 2009. *It's About Time: Understanding Einstein's Relativity*. Princeton: Princeton University Press.
Mitchell, W., J.R. Young, and J. Imray. 1860. *The Circle of the Sciences, vol. IX*. London: Griffin.
Okasha, Samir. 2000. "Van Fraassen's Critique of Inference to the Best Explanation." *Studies in History and Philosophy of Science* 31:691–710.
Pal, Palash. 2003. "Nothing But Relativity." *European Journal of Physics* 24:315–19.
Robison, John. 1822. *A System of Mechanical Philosophy, vol. 1*. Edinburgh: Murray.
Salmon, Wesley. 1975. "Confirmation and Relevance." In *Minnesota Studies in the Philosophy of Science*, vol. 6, edited by Grover Maxwell and R.M. Anderson, 3–36. Minneapolis: University of Minnesota Press.
Schupbach, Jonah N. 2017. "Inference to the Best Explanation, Cleaned Up and Made Respectable." In *Best Explanations: New Essays on Inference to the Best Explanation*, ed. Kevin McCain and Ted Poston, 39–61. Oxford: Oxford University Press.
Schupbach, Jonah N., and Jan Sprenger. 2011. "The Logic of Explanatory Power." *Philosophy of Science* 78:105–27.
Sprenger, Jan, and Stephan Hartmann. 2019. *Bayesian Philosophy of Science*. Oxford: Oxford University Press.
van Fraassen, Bas. 1980. *The Scientific Image*. Oxford: Clarendon.
Whewell, William. 1832. *The First Principles of Mechanics*. Cambridge: L. and L.L. Deighton.
Whewell, William. 1858. *History of Scientific Ideas, vol.* 1, 3rd ed. London: Parker.

Cite this article: Lange, Marc. 2022. "Against Probabilistic Measures of Explanatory Quality." *Philosophy of Science* 89 (2):252–267. https://doi.org/10.1017/psa.2021.28