Noname manuscript No.
(will be inserted by the editor)

Inference, Explanation, and Asymmetry

Received: date / Accepted: date

Abstract Explanation is asymmetric: if A explains B, then B does not explain A. Traditionally, the asymmetry of explanation was thought to favor causal accounts of explanation over their rivals, such as those that take explanations to be inferences. In this paper, we first develop a new inferential approach to explanation, and argue that it outperforms causal approaches in accounting for the asymmetry of explanation. Explanation is asymmetric: if A explains B, then B does not explain A. Traditionally, the asymmetry of explanation was thought to favor causal accounts of explanation over their rivals, such as those that take explanations to be inferences. In this paper, we first develop a new inferential approach to explanation, and argue that it outperforms causal approaches in accounting for the asymmetry of explanation.

1 Introduction

A surefire way to embarrass a theory of explanation is to show that it fails to respect the commonsense idea that explanation is an asymmetric relation. Give or take some rare exceptions, if A explains B, then B does not explain A. In the lore of philosophical accounts of explanation, the origin myth almost always includes reference to a flagpole and its shadow.

The symmetry problem serves as an expedient way to disqualify a position that we call *Explanation-As-Inference* (hereafter: EAI). Bromberger (1965) used symmetry to critique Hempel's covering-law model of explanation. Kitcher's (Kitcher 1989) unificationist theory purported to restore explanation's asymmetry, but he faced other, searching symmetry counterexamples (see Barnes 1992). By contrast, the symmetry problem has been a great advertisement for causal approaches to explanation. According to these views, explanation's asymmetry follows effortlessly in the wake of causation's asymmetry. Given the long shadow that the symmetry problem casts, it is no wonder that causal approaches to explanation seem to enjoy a privileged status in contemporary philosophy of science (Strevens 2008; Woodward 2003).

Despite their prominence, causal theories of explanation face their own challenges with respect to the asymmetry of explanation. A growing body literature shows that some scientific explanations are noncausal **?**. While the mere existence of noncausal explanations challenge causal theories, they also raise a further, hitherto unnoticed, asymmetry problem. Noncausal explanations exhibit asymmetries. This suggests that the ultimate source of explanatory asymmetry may not be causal, and it undermines an important dialectical motivation for adopting a causal theory. Hence, the holy grail would be an analysis that accommodates causal and noncausal explanations, and accounts for the asymmetries of both.

In this essay, we shall offer a new set of necessary conditions for explanation that provides a unifying framework for handling asymmetries. While we consider ourselves members of the EAI family, we make bold departures from our predecessors. Section 2 specifies the contours of the symmetry problem when both its causal and noncausal variants are taken on board. Section 3 then presents our new approach, which we call the *defeasibility* model of explanation. Section 4 then shows how the defeasibility model solves some of these symmetry problems without adverting to causal concepts. Section 5 then considers a tougher challenge to the defeasibility model, in which causal concepts must be invoked in order to capture the relevant asymmetries. We argue that even though causation figures prominently in solving this last class of symmetry problems, it does so in a way that supports, rather than diminishes, the thesis that inference is the ultimate locus of explanation.

2 The Symmetry Problem

Historically, the symmetry problem was a pitched as a challenge to Hempel's deductive-nomological (DN) model of explanation. The DN Model was one of three versions of Hempel's covering law theory, along with inductive-statistical (IS), and deductive-statistical (DS) explanations. In all three, the explanandum is inferred from premises that include at least one law of nature. A DN explanation of E is a sound deductive argument of the form $C_1, \ldots, C_k, L_1, \ldots, L_r \vdash E$, where $L_1, \ldots, L_r$ are universal laws of nature and $C_1, \ldots, C_k$ are initial conditions. All premises must be indispensable to the argument's validity.

While most discussions of the symmetry problem involve a flagpole and its shadow, Sylvain Bromberger's original example is somewhat more colorful:

> There is a point on Fifth Avenue, M feet away from the base of the Empire State Building, at which a ray of light coming from the tip of the building makes an angle of θ degrees with a line to the base

> of the building. From the laws of geometric optics, together with the "antecedent" condition that the distance is M feet, the angle θ degrees, it is possible to deduce that the Empire State Building has a height of H feet. Any high-school student could set up the deduction given actual numerical values. By doing so, he would not, however, have explained why the Empire State Building has a height of H feet... (Bromberger 1966, p.92).

At the risk of belaboring what any high-school student could do, two inferences are involved in Bromberger's example:

$$\tan\theta = \frac{H}{M},\ \theta = 60^{\circ},\ H = 1{,}454\,\mathrm{ft}\ \vdash\ M = 839.5\,\mathrm{ft} \quad (\textsc{Classic Tower})$$

$$\tan\theta = \frac{H}{M},\ \theta = 60^{\circ},\ M = 839.5\,\mathrm{ft}\ \vdash\ H = 1{,}454\,\mathrm{ft} \quad (\textsc{Classic Shadow})$$

Since both inferences are deductions from a law and initial conditions, both count as explanations according to Hempel's criteria. But as Bromberger points out, only the first is plausibly an explanation. The problem easily generalizes. Many physical laws are expressed by equations treating one variable as a function of others. These laws permit the value of any one variable to be deduced from the values of the others. Like the building and its shadow, only some of these inferences are explanations.

Subsequent proponents of EAI tried to rule out the asymmetries by further restricting the inferences that count as explanations. For instance, where Hempel required only that explanandum be validly deduced from the explanans, Kitcher's unificationist[1] model required that the explanandum be derived in a particular way. Each explanatory derivation is an instance of a more abstract argument pattern that specifies the kinds of premises and inferences rules that may be used to derive the explanandum. A *systemization* of the corpus of accepted statements K is any set of general argument patterns that derive some members of K from others. Explanation consists of using instances of the "best" systemization, *E(K)*, as measured according to the following criteria:

1. *Acceptability*: Each step of each instance of a general argument pattern in *E(K)* must be deductively valid, and acceptable relative to K.
2. *Scope*: Unification increases in proportion to the size of the conclusion set of the number of acceptable instances of *E(K)*.
3. *Stringency*: Unification increases in proportion to the strictness of the argument patterns in *E(K)*.
4. *Number of patterns*: Unification decreases in proportion to the number of general argument patterns in E(K).

[1] Other unificationists who endorse EAI include **?**Friedman (1974); Schurz and Lambert (1994); Schurz (1999). Space prohibits extensive discussion of their views on symmetry.

Kitcher's view accounts for the asymmetry in Bromberger's example. Kitcher proposes that our present explanatory store contains an "origin and development" (OD) argument pattern, according to which spatial dimensions of any physical object are derived from its origin and subsequent physical changes. Kitcher then invites us to consider a "shadow" (S) argument pattern, wherein the spatial dimensions of physical objects are derived from the length of their shadows. CLASSIC SHADOW would be an instance of S. Simply adding S to our explanatory store runs afoul of the fourth criterion of unification, above. Hence it will only be explanatory if it fares better along one of the other dimensions of unification. However, it appears that the acceptable conclusions that S generates are a proper subset of those that OD generates. Nor does S appear any more stringent than OD. Hence, Kitcher's unificationism is not susceptible to shadowy symmetries because the non-explanatory inferences do not unify a scientific domain.

However, Eric Barnes' Barnes (1992) version of the symmetry problem threatens Kitcher's version of EAI. Barnes imagines a closed system of "Newtonian particles." Given a complete description of this system at a time, the state of the system at any later time can be determined through Newton's laws. Barnes argues that the deduction of future states of the system from the present state satisfies Kitcher's criteria for explanation. Newtonian mechanics, after all, is a paradigm of scientific unification. But Newton's laws also permit the deduction of past states from present states. So, since Newtonian mechanics fits Kitcher's criteria, both the forward and backward calculations must count as explanatory. But clearly, retrodicting past states from the present does not count as an explanation.

The most prominent diagnosis of explanatory asymmetries holds that inferences such as CLASSIC TOWER are explanatory because they track causal relationships, while the non-explanatory CLASSIC SHADOW does not. Since nothing in the form of inference marks a causal relationship, many causal theorists of explanation argue that the inferences are superfluous, or, at the very least, are subservient to the more basic explanatory task of tracking causes. On such a view, all explanations represent causal relationships, and explanation is asymmetrical because causes are. EAI appears to have hit a dead end.

We should be suspicious of any diagnosis that identifies the asymmetry of explanation with the asymmetry of causation. Begin with the observation that some explanations are noncausal. Consider, for example, this explanation:

> The players selected for the [Home Run] Derby are typically among the best home-run producers of the first half, though they may not necessarily be among the best power hitters in baseball. Uncharacteristic performances help players get selected for the Home Run Derby, and the decline in their numbers in the second half is more

likely to be due to natural regression than their participation in the event (Braunstein and Woolums 2014).

For those unfamiliar with baseball, the Home Run Derby invites the best home run hitters in baseball to show their skills in the middle of the season. Braunstein and Woolums are explaining why players selected for the derby tend to do worse in the second half of the season. The fall-off in home run production is sometimes attributed to participation in the Derby itself; it might be something about the psychology of participation, or perhaps the Derby changes the player's swing. Such *causal* hypotheses, however, are generally not good explanations, according to Braunstein and Woolums' analysis. They suggest that the players' slump in the second half of the season is regression toward the mean. That is, if a series of performances are significantly above or below average, subsequent performances will be closer to the average.[2]

Braunstein and Woolums' explanation exhibits an explanatory asymmetry. To see it clearly, consider their explanation as applied to a particular player. They discuss Chris Davis who, in the first half of the 2013 season, hit well above his average at this stage of his career. Braunstein and Woolums explain Davis' poorer second-half performance as regression toward the mean. However, the converse is no explanation: the phenomenon of regression toward the mean is not explained by anything about Chris Davis' performance. Since explaining Davis' second half slump in terms of expected values, standard deviations, and regression is clearly not causal, this explanatory asymmetry cannot be a causal asymmetry. A failure of inferences to track causes is not the source of the explanatory asymmetry.

It appears that there are a variety of explanatory asymmetries and they are susceptible to different kinds of solution. Some symmetry problems, such as Bromberger's, appear to be solvable by either inferential or causal means. Others, such as Barnes' example of the Newtonian particles, appear to favor causal approaches over EAI. Still others, such as the asymmetry involving regression towards the mean, appear to resist any causal analysis. Thus, contrary to the received wisdom, it is far from clear that causal accounts have gotten to the root of explanatory asymmetries. We take this opening as an opportunity to reenvision EAI. The net result will be one in which all three kinds of asymmetries can be seen to spring from a common inferential fountainhead.

3 The Defeasibility Model of Explanation

Suppose, like us, that you are sympathetic to EAI, and want to solve these symmetry problems. Where did others go wrong? Earlier proponents of EAI focused almost exclusively on *classical* logic and probability. Yet

[2]Both Lange (2016) and Lipton (2004) discuss regression to mean as an example of noncausal explanation. Lange provides further examples of "really statistical" explanations in population genetics. For further examples, see Ariew et al. (2014); Walsh (2015).

the classical consequence relation is especially permissive, allowing for inferences such as *Ex Falso Quodlibet* and antecedent strengthening. Earlier propnents of EAI responded by excluding non-explanatory inferences. They imposed requirements such as soundness, syntactic and semantic constraints on laws, probability thresholds—not to mention Kitcher's appeals to acceptability, scope, stringency, and number of patterns. Instead, we will adopt a *nonclassical* logic, which provides alternative consequence relations that behave much more like the explanatory relation.[3]

A further hint for solving symmetry problems is found in Kitcher's solution to Bromberger's problem. Kitcher shows *comparative* failings disqualify them as explanations. In particular, it seems as if the proper explanation of the tower's height—say an architect's design—will succeed where the shadow "explanation" fails. Similarly, many *causal* approaches to explanation hold that CLASSIC SHADOW is not explanatory because, when we hold the architect's design fixed, the tower's height would still be 1,454 feet, even if the shadow's length were not 839.5 feet. As we have seen, gaps remain in both approaches. This suggests that they may be deploying the wrong basis of comparison.

By using nonclassical consequence relations, we provide a new way of explicating the insight that a proper explanation succeeds where its competitors fail. Roughly, our defeasibility model of explanation holds that A explains B only if:

1. A and B are (approximately) true,
2. B is a nontrivial consequence of A, and
3. the inference from A to B succeeds under conditions where all others fail.

We will say that inferences satisfying the last two conditions are "sturdy." To provide a better sense of the defeasibility model, we discuss each of the three conditions (Sections 3.1, 3.2, and 3.3) in turn.

Before doing so, two preliminary clarifications are in order. First, in this paper, we only provide necessary conditions for "A explains B." Hence, our analysis is only partial. This will not matter in what follows, since we will have provided enough of an analysis to solve the symmetry problem within an EAI framework. In future work, we intend to complete this analysis. Second, for many explananda B, there are multiple propositions $A_1, \dots A_n$ such that it is natural to say that A_1 explains B and that A_2 explains B, etc., and that these explanantia are not in competition. While the arguments are outside of the scope of this essay, it is a consequence of our view that when A explains B, A is the exhaustive explanation, and that it encompasses $A_1, \dots A_n$. The pragmatics of explanation permit an element of the exhaustive explanation, A_i, to be treated as "the" explanation of B. For ease of exposition, we will not enumerate all of the different components of an exhaustive explanation

[3]In AUTHOR CITATION, we develop a formal system that offers more precise characterizations of the ideas in this paper. The informal gloss of these ideas suffices for our current purposes.

when we provide examples of sturdy inferences. What matters most for this paper is that our view entails that for any A and B that raises a symmetry problem (*i.e.* such that A explains B, but B does not explain A), B can be shown not to be part of a sturdy inference that has A as its conclusion, and hence can be excluded from the exhaustive explanation of A.

3.1 Explanation and Truth

The least remarkable of our requirements for explanation is that the explanans and explanandum must be *true*. This conforms to common usage, where a false proposition is not the actual explanation. Presumably, several alternative accounts of "quality control" on the explanans and explanandum—*e.g.* involving different theories of truth, or appealing to significantly different semantic or epistemic properties than truth—can be wedded to Conditions 2 and 3 (sturdiness), and still furnish similar solutions to the symmetry problem, so we will mostly take this requirement for granted in what follows.

We have parenthetically added that the explananans and explanandum may be *approximately* true. For many scientific explanations, the premises are known, strictly speaking, to be false Cartwright (1983). For instance, it would be miraculous if the Empire State Building stood at a *perfect* 90° angle to 5^{th} Avenue, *exactly* at $1,454$ feet, etc. Nonetheless, the building's height explains the shadow's length for roughly the reasons implied by CLASSIC TOWER. In Section 3.2, we actually show that approximation and defeasibility pair naturally with each other.

3.2 What is a Nontrivial Consequence?

For our purposes, nontrivial consequences have three key features. Each of them maps on to properties of explanation. First, nontrivial inferences are *irreflexive*. This accords with the idea that explanation is also an irreflexive relationship, e.g., that the shadow being 839.5 feet long does not explain why it is 839.5 feet long.

Second, nontrivial inferences are *premise consistent*. Since a contradiction explains nothing, the classically valid inference pattern *Ex Falso Quodlibet* (where a contradiction entails any proposition) cannot be an explanation. As its name suggests, premise consistency requires that the premises of a nontrivial consequence must be consistent.

Third, nontrivial consequences are *defeasible*. Defeasible consequence relations are disrupted when certain additional propositions—called *defeaters*—are considered. Scientific explanations are replete with them. For instance, a person's disease explains her symptoms, but a person's disease in conjunction with her taking an effective treatment does not. We will represent defeasible inferences this way:

$$\begin{aligned} &\text{The patient is infected with the } \textit{Varicella zoster} \text{ virus} \\ \mid\!\sim_{\Theta}\ &\text{The patient's skin is covered in red spots} \end{aligned}$$

Here, "$\mid\!\sim_{\Theta}$" denotes a defeasible consequence relationship. "Θ" denotes a set of defeaters, such as "The patient has received effective treatment." A good defeasible inference of this sort will turn bad if members of Θ are true.[4]

The turn to defeasible inference is an important departure from earlier proponents of EAI. The consequence relation of classical logic is not defeasible in the sense articulated here. If the sequent $A \vdash B$ is classically valid, no new information will disrupt the inference from A to B. By contrast, new data ought to be able to undermine any candidate for scientific explanation (at least in principle). Hence, if the explanatory relation is inferential, then it is almost certainly defeasible.

One might object that at least some explanations are not defeasible. In particular, explanations involving mathematically formulated physical laws walk and talk like classically valid inferences. Explanations in mathematized sciences, one might argue, are indefeasible through and through. This objection overlooks the way that modeling practices—such as idealization, abstraction, *ceteris paribus* clauses, and approximation—"hide" the defeasibility of explanation. Once this point is appreciated, treating explanations involving mathematically formulated laws as defeasible inferences is advantageous. Consider once again the approximations involved in the explanation of the shadow. If the building is too far from being perpendicular, the inference will not go through. This is the kind of *ceteris paribus* consideration that the "defeater set", Θ, captures. Moreover, the defeater set includes many other limitations on the model, such as that the law of geometric optics involves an idealization (light behaves as a ray) that breaks down under certain conditions (*e.g.*, in quantum systems.)

For these reasons, we will represent the inferences in Bromberger's example as:

$$\tan\theta = \frac{H}{M},\ \theta = 60^{\circ},\ H = 1,454\,\mathrm{ft} \mid\!\sim_{\Theta}\ M = 839.5\,\mathrm{ft} \qquad (\textsc{Tower})$$

[4]Roughly put, an inference is defeated whenever its premises contain a sentence that is logically equivalent to a subset of the defeater set. Our precise definition of defeat departs from this characterization in some respects, only two of which are noteworthy. (1) a disjunction in the premises defeats an inference if both the disjuncts (or their logical equivalents) belong to the defeater set, while (2) a conjunction in the premises defeats the inference if at least one of the conjuncts (or their logical equivalents) belongs to the defeater set. Although we refrain from providing it here, all of our informal references to *defeat* in the present text conform to the precise definition given in AUTHOR CITATION.

$$\tan\theta = \frac{H}{M},\ \theta = 60^{\circ},\ M = 839.5\,\text{ft} \mid_{\overline{\Theta'}}\ H = 1{,}454\,\text{ft} \qquad (\textsc{Shadow})$$

Hereafter, we assume that for *us* to solve Bromberger's symmetry problem, we must show that TOWER is explanatory, but SHADOW is not.

3.3 What Is Inferential Success And Failure?

To solve the symmetry problem, we must introduce a new basis for comparing explanations. In slogan form, it is that only inferences that succeed where all others fail can be candidates for explanation. But what is meant by "success" and "failure" in such a slogan? The defeasibility of explanatory arguments provides an important clue. To say that one inference succeeds when another fails, we can imagine explanations being evaluated according to the following comparative procedure:[5]

Step 1: Line up all of the nontrivial inferences that have the explanandum, B, as their conclusion. For each of these inferences, all other nontrivial inferences leading to B are its "competitors."

Step 2: For each A that has B as a nontrivial consequence, suppose that all of A's competitors' premises are false.

Step 3: If the falsehood of any of these competitors defeats the inference from A to B, then the latter is not sturdy; otherwise, it is sturdy.

Failure is defeat under these conditions; success is the absence of failure.

For example, assume that TOWER is a correct explanation. Step 1 requires us to consider other ways of inferring the shadow's length. For instance, one competitor might be that there is a taller building behind the Empire State Building. This building is in line with the sun and casts the shadow. (Call this the TALLER TOWER explanation.) Step 2 then requires that we suppose the following: there is no such building casting a shadow. Step 3 then asks whether the claim that no taller building is casting the shadow defeats the claim that the Empire State Building is casting the shadow. But, of course, the absence of a taller tower will not disrupt the Empire State Building's ability to cast the shadow. Assuming that this could be done with all other competitors, TOWER is sturdy.

However, this raises an apparent problem. Suppose we flip the script, such that TOWER's premises are assumed to be false and considered as potential defeaters to TALLER TOWER. TOWER does not defeat TALLER TOWER, since the shadow's height could still be inferred from the height

[5]To be clear: we are not claiming that explanations must be the products of this procedure. Indeed, we make no claims about the "production" of explanations whatsoever. Rather, this three-step process is simply a useful heuristic for the reader to identify the relevant inferential properties that distinguish explanations from other nontrivial inferences.

of the taller building. Hence, Taller Tower is also sturdy. The sturdiness of both inferences is no cause for alarm. While it is sturdy, Taller Tower has false premises. Hence, it will not count as an explanation. Intuitively, it is merely a "potential explanation." Furthermore, we can make our account of explanation more realistic by recruiting the idea that defeasible inference happens against a background of conditions that are held fixed. In effect, this background is the mirror image of a defeater, i.e. if C is held fixed in the background, then $\neg C$ is a defeater. Hence, we are likely to hold the non-existence of this taller tower fixed in such a way that it is a "dead option" as a potential explanation (or nontrivial inference) of the shadow's length.

As we shall argue in Section 4, sturdiness is what distinguishes explanations from their symmetry-mongering counterparts. However, before doing so, we should explain why we take sturdiness to be a characteristic feature of the inferences that constitute explanations. First, the ideas underwriting sturdiness capture the comparative evaluation that characterizes much explanatory reasoning. A paradigmatic way of evaluating candidate explanations is to see whether the explanandum still holds when one of the competing explanantia is false while the other is true. For instance, to determine whether Chemical X or Chemical Y caused a reaction, we would hold all other conditions fixed. If the reaction occurred when X was present and Y was absent, but not *vice versa*, then X would be a better explanation of why the reaction occurred than Y. Upon iteration, we would then arrive at the best explanation. Indeed, reasoning such as this animates, *inter alia*, controlled experiments, Mill's Method of Difference, and Lipton's (2004) well-known approach to Inference to the Best Explanation.

Second, sturdiness is a form of "stability" that many philosophers take to be a central feature of explanations. While different philosophical discussions of explanation and laws use different terminologies, in its most general form, X is said to be stable if X remains unchanged as other conditions C change. Call X the *stability-bearer*, and C the set of *stability-conditions*. For instance, Hempel's stability-bearers are so-called "lawlike generalizations," and his stability-conditions include spatiotemporal changes among other things. Similarly, Woodward's Woodward (2003) notion of "invariance" is a kind of stability, its bearer being generalizations and its conditions being various kinds of interventions.[6] Our brand of stability, sturdiness, attaches first and foremost to *inferences*, and its stability-conditions are, at root, *competitors*. Our view is compatible with there being a derivative sense in which generalizations and laws are stability-bearers, and in which spatiotemporal changes and interventions are stability-conditions.[7]

For these reasons, sturdiness is the trademark feature of the defeasibility model. Sturdiness is tied to explanation in two ways: it dovetails with the ways in which scientists compare explanations, and it exhibits a kind of stability that is characteristic of good explanations.

4 Back to the Symmetry Problem

The defeasibility model's animating idea is that an explanation is a nontrivial inference that succeeds where its competitors fail. With this idea in hand, let us return to the symmetry problem. Our argument will proceed as follows:

1. An inference is an explanation only if it is sturdy.
2. The inferences that beget the symmetry problem are not sturdy.

∴ The inferences that beget the symmetry problem are not explanations.

Section 3 motivates the first premise of this argument. The second premise is carrying most of the dialectical load, for our goal in this essay is to show how our version of EAI can resolve the symmetry problem(s). The remainder of this essay supports the second premise by applying our defeasibility model to each of the three symmetry problems discussed in Section 2. For each problem, we will briefly sketch how the correct explanation is a viable candidate for sturdiness, and show that the symmetry-mongering inference is not sturdy. In this section we address Bromberger's classic symmetry problem and the noncausal asymmetry problem, involving regression to the mean. Since it requires a more involved discussion, we postpone our solution to Barnes' symmetry problem until Section 5.

4.1 The Classic Symmetry Problem

To show that the shadow's length does not explain the tower's height, we must find some alternative, nontrivial way of inferring the tower's height that succeeds where Shadow fails. One obvious candidate is:

$$\begin{array}{l} \text{The Empire State Building was designed to be 1,454 feet} \\ \mid\!\sim_{\Theta} H = 1,454\,\text{ft} \end{array} \qquad (\textsc{Design})$$

If the negated premises of just one would-be competitor defeat Shadow, then the latter is not sturdy. Hence, we can treat Design and Shadow as the only nontrivial inferences that have the tower's height as their conclusion. Thus, for our purposes, Step 1 of 3 in our comparative procedure is complete.

[6]Other accounts of stability include Lange (2009); Mitchell (2003); Skyrms (1980).

[7]Indeed, while we will not argue for it here, Woodward's account of interventions can be seen as exhibiting the kind of inferential sturdiness we describe. Section 5 provides some clues as to how this argument would proceed.

Turning to Step 2, suppose that the Empire State Building was designed to be a different height. That is, we suppose that the premise of DESIGN is false. Such a supposition defeats SHADOW in two ways. First, a different design will imply that, when given the same angle of incidence (θ), the shadow's length will be different (i.e. $M \neq 839.5$ feet). However, this contradicts one of SHADOW's premises, and, as discussed above, inconsistent premises are *verboten*. Second, a different design implies that the height of the tower will be different (i.e. $H \neq 1,454$ feet), so when the different design is added to the premises, we do not get the desired conclusion—also a mark of defeat. Per Step 3, both of these considerations imply that SHADOW is not sturdy. Hence, according to the defeasibility model, the shadow's length does not explain the tower's height. The symmetry is blocked.

4.2 Noncausal Explanatory Asymmetries

As argued above, causal theories of explanation cannot do justice to the asymmetry of noncausal explanation. Hence, it will be a victory if the defeasibility model of explanation can outperform the causal theory on this front. Specifically, we will now show that our model captures the sturdiness of Braunstein and Woolums' explanation of Chris Davis' 2013 season, and that it rules out the converse as non-explanatory.

In their discussion, Braunstein and Woolums measure a batter's home run productivity with his home-run-per-fly-ball ratio (HR/FB). Using this statistic, they explain Chris Davis' second half performance in terms of regression toward the mean:

Performance after the Home Run Derby (as measured by HR/FB) regresses toward the mean.

Chris Davis' performance in the first half of 2013 (35.6 HR/FB) was well above his career average (22.3 HR/FB)

$\mid\!\sim_{\Theta}$ Chris Davis' performance in the second half of 2013 (21.3 HR/FB) was significantly worse than the first half.

(REGRESSION)

As noted in Section 2, noncausal explanations exhibit asymmetries. Suppose we turn Braunstein and Woolums's explanation on its head:

Chris Davis' performance in the second half of 2013 (21.3 HR/FB) was significantly worse than the first half.
Chris Davis' performance in the first half of 2013 (35.6 HR/FB) was well above his career average (22.3 HR/FB).

$\vert\!\overline{}_{\Theta'}$ Performance after the Home Run Derby (as measured by HR/FB) regresses toward the mean.

(PERFORMANCE)

Chris Davis' performance should not explain the regression phenomenon any more than the length of the shadow explains the tower. Hence, we must show that PERFORMANCE is not a sturdy inference, and hence not an explanation.

The competitor to PERFORMANCE would be the proof that probability distributions with certain features exhibit regression toward the mean, plus the fact that seasonal home run hitting (as measured by HR/FB) has those features. Changes in Chris Davis' batting would not be defeat such an argument. On the other hand, the denial of this competitor's premise, e.g. the assertion that home run performance (HR/FB) does not regress toward the mean, would defeat PERFORMANCE. Thus, PERFORMANCE is not sturdy. According to the defeasibility model, this means that it is not a candidate for explanation. Hence, our account of explanation blocks the problematic inference that gives rise to noncausal symmetries.

Since it will prove useful below, let us also discuss why the correct explanation, REGRESSION, does not meet a similarly ignominious fate. This inference is nontrivial in the sense characterized by Section 3.2. It is irreflexive, premise consistent, and defeasible. To determine its sturdiness, we need to compare it with other inferences that have the same conclusion. It so happens that Chris Davis broke a blister on his hand during the Home Run Derby, and some blamed this for his fall-off in performance. So, one competitor to REGRESSION would then be:

Chris Davis broke a blister during the Home Run Derby.

$\vert\!\overline{}_{\Theta''}$ Chris Davis' performance in the second half of 2013 (21.3 HR/FB) was significantly worse than the first half.

(BLISTER)

The second step of the sturdiness test requires us to negate the premise of this argument. The third step is to determine whether negating the premise of BLISTER defeats REGRESSION. Combined, this is to say that if Davis had not been mildly injured, his performance in the second half of the season still would have been significantly worse than it was in the first half.[8] This is plausible, because popping a blister is a relatively mild injury. If the injury were more severe—say a broken hand—then it would be a defeater and REGRESSION would not be sturdy. Thus, under specific empirical conditions, REGRESSION is a good candidate for explanation.[9] Moreover, it clearly defeats BLISTER, for if seasonal home run performance did not regress toward the mean, then the inference from Davis' mild injury to the significant fall-off in performance would not go through.

Note that BLISTER is a causal explanation. While the inference is not sturdy, there is certainly some causal relationship between the blister and Davis' second half performance. It probably affected his batting to some degree, but we are supposing that the injury is too mild to account for all of the difference. This illustrates an interesting result: inferences that do not track causes can sometimes trump inferences that do, and whether BLISTER is superior to REGRESSION is an empirical matter. Hence, we have raised two problems with the causal approach to the symmetry problem. First, some explanatory asymmetries, such as the difference between REGRESSION and PERFORMANCE, do not appear to rest on causal facts. Second, under certain empirical conditions, some causal explanations, such as BLISTER, are inferior to noncausal explanations. As we shall see, these two points will prove instrumental in drawing the appropriate lessons from our last and most challenging symmetry problem, to which we now turn.

5 Causation and Sturdy Inference

Thus far, we have seen one explanatory asymmetry—involving the Empire State Building and its shadow—that *need not* be construed as a causal asymmetry, and another—involving regression to mean—that *should not* be construed as a causal asymmetry. However, there is a class of explanatory asymmetries where causation plays a starring role. Barnes' critique of Kitcher, briefly discussed in Section 2, provides an exemplar of this kind of asymmetry. Our treatment of this species of asymmetry involves three main argumentative maneuvers. First (Section 5.1), we motivate the special challenge these examples pose for our view. In particular, appeals to sturdiness without further appeal to causation end up being either too permissive (leaving the symmetry-mongering inference sturdy) or too prohibitive (leaving the correct explanation unsturdy). Then (Section 5.2), we show that appealing to causation delivers the correct verdicts

[8] Admittedly, this gets a bit more complicated when we consider the actual numbers. Perhaps if Davis had not had a blister, his second-half HR/FB would still have been substantially lower than it was in first half, but not as low as 21.3. If this were the case, it would technically defeat the inference. This is remedied by changing the explanandum from having a specific value (21.3 HR/FB) to its having a range of values for his second-half HR/FB (*e.g.* 24 HR/FB or fewer). Introducing ranges would also mirror the common statistical practice of using confidence intervals. We bracket this complexity for ease of exposition.

[9] In addition to the differential effects of regression towards the mean and Davis' blister on his hitting, REGRESSION is shown to be sturdy if and only if, for its remaining competitors, either REGRESSION is undefeated on the supposition that their premises are false, or these competitors are held fixed. Section 5 discusses what it means to be "held fixed" on the defeasibility model.

as to which inferences ought to be sturdy. Finally, we argue that despite this appeal to causation, sturdiness is still the driving force behind the explanatory asymmetry (Section 5.3).

5.1 Permissive and Prohibitive Sturdiness

Let us begin by showing that, for some examples of explanatory asymmetry, avoiding causation raises certain problems. For example, conservation of kinetic energy entails that any moving particle X that collides elastically with a resting particle Y of equal mass will obey the following law:

$$(V_{1X})^2 = (V_{2X})^2 + (V_{2Y})^2 \qquad \text{(Vel. Law)}$$

In this formula, V_{1X} denotes X's velocity up to its collision with Y , while V_{2X} and V_{2Y} denote the velocities of the particles at some time after this collision. For concreteness' sake, let us consider a simple system consisting of two billiard balls, A and B, on a standard billiards table. A moves across the table and collides with B, which was at rest. After the collision, A's velocity has changed, and B takes on a velocity.

As with our previous examples, there are two initial inferences to consider:[10]

$$\text{Vel. Law},\ V_{1A} = 1\,\text{m/s},\ V_{2A} = .6\,\text{m/s} \mid\!\sim_{\Theta} V_{2B} = .8\,\text{m/s} \qquad (\text{Ball A})$$

$$\text{Vel. Law},\ V_{2B} = .8\,\text{m/s},\ V_{2A} = .6\,\text{m/s} \mid\!\sim_{\Theta'} V_{1A} = 1\,\text{m/s} \qquad (\text{Ball B})$$

Much like before, we must show that Ball A's prospects for being sturdy are strong, while Ball B is not sturdy. Following the now-familiar template, consider a would-be competitor to Ball B; say one in which the velocity of ball A is inferred from some prior event. Suppose, for example, that, before colliding with B, A was at rest until struck by a third ball C. In that case, the following defeasible inference would be acceptable (where V_{0C} is the velocity of ball C at time interval $t_0 < t_1$):

$$\text{Vel. Law},\ V_{0C} = 1.17\,\text{m/s},\ V_{1C} = 0.6\,\text{m/s} \mid\!\sim_{\Theta''} V_{1A} = 1\,\text{m/s} \qquad (\text{Ball C})$$

Let us see whether Ball C blocks the sturdiness of Ball B. As before, we negate the premises of Ball C, i.e., we suppose that C's velocity was not 1.17 m/s. If all went according to plan, Ball B would be defeated. But alas, this is not the case: the supposition that C did not collide with A at 1.17 m/s is compatible with Ball B still being a good retrodiction of ball A's velocity.[11] This is because a different set of conditions at time

[10] Recall from Section 3.2 that such inferences are defeasible because of modeling considerations.

0 might have brought it about that $V_{1A} = 1$ m/s. C's velocity could have been higher or lower (making suitable adjustments to the angle of impact and post-collision velocity of C), and it would still follow that $V_{1A} = 1$ m/s. The difference between Ball A and Ball B cannot be made out on a straightforward analogy with our treatment of Shadow in Section 4.1.

While it is a dead-end, this approach to the symmetry problem highlights an interesting feature of the defeasibility model: the importance of the background conditions. In the earlier discussion of the Tower inference, we noted that the existence of other buildings shading out the Empire State Building was merely a *potential* explanation, since its explanantia were known to be false. We could then rule out these potential explanations by holding certain factors fixed, such as the absence of a taller tower behind the Empire State Building. The defeasibility of Tower thus depends in part on what counterfactual suppositions are allowed and what facts are to be held fixed. In the discussion of the billiard ball example so far, we have permitted any counterfactual supposition about ball C's velocity. Ball B is not defeated by negating C's velocity because there are too many ways for C's velocity to be different without changing A's velocity. As a result, the bare fact that C is not at 1.17 m/s does not defeat Ball B. Since nothing is held fixed, all of the inferences we have been considering remain undefeated. For this reason, we call this way of approaching this kind of symmetry problem *permissive sturdiness.*

Since permissive sturdiness fails to achieve the desired asymmetry, perhaps we should hold these other routes to A's velocity fixed. But how should we understand these "routes" in an inferential idiom? In this example, let an "alternative route" be any nontrivial inference other than Ball B or Ball C that has V1A = 1 m/s as its conclusion, *e.g.*:

$$\textsf{Vel. Law},\ V_{0C} = 1.28\,\mathrm{m/s},\ V_{1C} = 0.8\,\mathrm{m/s} \mathrel{|\!\!\!\sim_{\Theta'''}} V_{1A} = 1\,\mathrm{m/s} \quad (\text{Ball C*})$$

"Holding fixed" then amounts to assuming that every alternative route is defeated. By holding the background conditions fixed in this way, all other inferences to A's velocity will be ruled out. This, in turn, ensures that, when we get to Step 2, and negate the premises of Ball C, it follows that $V_{1A} \neq 1\,\mathrm{m/s}$. In other words, inferences such as Ball C* will no longer be able to "take over" for Ball C once its premises are negated in Step 2. In this way, C's velocity acts like a switch for A's velocity, i.e. changes in C's velocity are both necessary and sufficient to change A's velocity. Hence, the negated premises of Ball C defeat Ball B, for we get an inconsistency, $V_{1A} = 1\,\mathrm{m/s} \wedge V_{1A} \neq 1\,\mathrm{m/s}$. The desired result at Step 3 is thereby achieved: Ball B is not sturdy, and hence not an explanation.

Unfortunately, the problem with this way of holding things fixed is that it throws out the baby with the bathwater. Maintaining the assumption that every other way of inferring ball A's velocity except Ball B and

[11] To simplify the presentation, the ambiguous phrase "ball A's velocity' denotes V_{1A}, "ball B's velocity" denotes V_{2B}, and "ball C's velocity" denotes V_{0C}, unless otherwise specified.

Ball C is defeated, let's now turn things around and consider whether the negated premises of Ball B defeat Ball C. Since we have ruled out all of the other possible future trajectories of B, then, if $V_{2B} \neq 0.8$ m/s, V_{1A} will not equal $1\,\mathrm{m/s}$.[12] So, although Ball B will not be an explanation, neither will Ball C (nor, presumably, will Ball A be an explanation for that matter.)[13] If everything outside of the two inferences being compared is held fixed, no inferences are sturdy. Call this *prohibitive sturdiness.*

Since sturdiness depends on comparative defeasibility, and defeasibility depends, in part, on what background conditions are held fixed, sturdiness depends (in part) on the background. We have seen that if we hold *nothing* fixed, then *both* the symmetry-mongering inferences, Ball B and Ball C, are sturdy. That is the problem with being overly permissive. Yet, if we hold *as much as possible* fixed, and then *neither* of these inferences are sturdy. That is the problem with being insufficiently permissive (prohibitive sturdiness). The challenge is to finding principled ways to hold one or another aspect of the situation fixed when comparing inferences.

5.2 Causal Sturdiness

Let us take stock. When we were permissive about the background, we did not limit what might happen *after* ball A achieved its velocity ($V_{1A} = 1$ m/s). This meant that the negated premises of Ball B did not defeat to Ball C. Permissive sturdiness' partly captures the fact that past events cannot be influenced by future events. Prohibitive sturdiness holds a good deal fixed with respect to what happens *before* A gets its velocity of 1 m/s, so that the negated premises of Ball C defeat Ball B. In this way, it captures the idea that future events can be influenced by past events. If we could combine these restrictions, the symmetry of Ball A and Ball B could be broken. But how can this be done nonarbitrarily?

The basic idea is this: systems such as the billiard balls are *causal.* In a causal system, fixing causes makes renders the temporally forward-looking inferences sturdy and it undermines the sturdiness of their temporally backwards-looking cousins. Let us call this *Causal Sturdiness*

> Where C is an actual cause of A, and B is a later event,
>
> $B \mid\!\!\sim_{\Theta} A$ is not sturdy, while $C \mid\!\!\sim_{\Theta} A$ may be.
>
> (Causal Sturdiness)

[12] In effect, this will underwrite a backtracking counterfactual: "Had $V_{2B} \neq 0.8$ m/s, then it would have had to have been the case that $V_{1A} \neq 1\,\mathrm{m/s}$.

[13] Since the arguments showing that Ball C is sturdy (or not) can be extrapolated to Ball A, we focus on showing that the former is a candidate for sturdiness. Analogous considerations apply to the latter.

Applied to the example at hand, this means that if C's velocity is an actual cause of A's velocity, then Ball B is not sturdy, but this does not prohibit Ball C from being sturdy inference.

The crucial concept of the Causal Sturdiness thesis is the idea of an "actual cause." We borrow this notion from Woodward (2003) :[14]

(AC1) The actual value of $X = x$ and the actual value of $Y = y$.
(AC2) There is at least one route R from X to Y for which an intervention on X will change the value of Y, given that other direct causes Z_i of Y that are not on this route have been fixed at their actual values. (It is assumed that all direct causes of Y that are not on any route from X to Y remain at their actual values under the intervention on X.)

Ex hypothesi, the actual value of $V_{0C} = 1.17$ m/s and the actual value of $V_{1A} = 1$ m/s. Hence, (AC1) is satisfied. Furthermore, (AC2) requires some "route" or causal chain from V_{0C} to V_{1A} such that at least one change to V_{0C} leads to a change in V_{1A} when all other causes not on this route are held fixed at their actual value. This entails that had $V_{0C} \neq 1.17$ m/s, then $V_{1A} \neq 1$ m/s, even if all of the other causes of A's velocity remained exactly the same. Note that this was the desirable feature of prohibitive sturdiness: the negated premises of Ball C defeat Ball B, and the latter is prevented from being sturdy.

However, treating C's velocity as an actual cause does not require us to hold the *effects* of A's velocity fixed. Hence, we also inherit the good parts of permissive sturdiness: even if we negate the premises of Ball B, we may still infer A's velocity from C's velocity. The reason for this is that any number of other "retrodictive" inferences leading to A's velocity are compatible with the premises of Ball C. Hence Ball C is not defeated.

The Causal Sturdiness thesis says that when we treat the events under consideration as parts of a system where some events are *actual causes* of others, retrodictive inferences are rendered non-sturdy. This raises a concern: doesn't this solution to the symmetry problem concede too much to causality?

5.3 Causation or Inference: Which Comes First?

Our discussion of the billiards example highlights an underlying tension in the larger debate as to whether causation or inference more fundamental to explanation. On the one hand, the example shows causal explanations to be *sturdy*, which favors an "inference-first" approach to explanation. On the other hand, it also shows *causal* explanations to be sturdy, which favors a "causation-first" approach to explanation. So what role does causation play in grounding the asymmetry in the billiards example?

On our inference-first approach, the asymmetry of explanation ultimately boils down to differences in sturdiness. Causes turn out to be

[14] We have not used Woodward's ultimate formulation of actual causation, (AC*), as it introduces complexities that are unnecessary for the purposes at hand.

especially good ways of achieving sturdiness, but they are not the only means for doing so. For instance, even though Bromberger's classic example appears to involve a causal explanation, we managed to preserve its asymmetry without appealing to any causal relationship between the tower and the shadow. More telling, accounting for the asymmetry of explanations involving regression toward the mean cannot advert to causes. Finally, and perhaps most importantly, even when we deploy causes, our argument is altogether different than the typical argument for establishing explanatory asymmetry via causal asymmetry, i.e. we did not argue as follows:

1. $V_{1A} = 1$ m/s causes $V_{2B} = 0.8$ m/s.
2. Causation is asymmetrical, i.e. if A causes B, then B does not cause A.
3. All explanations cite causes.

$\therefore$ $V_{2B} = 0.8$ m/s does not explain $V_{1A} = 1$ m/s.

As our discussion of regression towards the mean makes clear, we reject the third premise, and, for all that we've argued, we can remain agnostic about the second. Indeed, while we accept the first premise, it also played no role in establishing the asymmetry of causal explanation in the billiard ball example. That premise relies on the causal relationship between the velocities of balls A and B to establish the desired asymmetry. By contrast, the Causal Sturdiness thesis relies on the causal interaction between the velocities of balls C and A to establish that the symmetry-mongering inference Ball B is not sturdy. In short, no part of the argument above is essential to our solution of Barnes' symmetry problem.

Since the argument above is the standard way of arguing for explanatory asymmetry on the basis of causal asymmetry, it is clear that we are up to something else. More precisely, our argument is as follows:

1. All explanations are sturdy inferences.
2. If $V_{0C} = 1.17$ m/s is an actual cause of $V_{1A} = 1$ m/s, then Ball B is not a sturdy inference.
3. $V_{0C} = 1.17$ m/s is an actual cause of $V_{1A} = 1$ m/s.

$\therefore$ Ball B is not an explanation.

Furthermore, prioritizing inferences over causes has several advantages when thinking about explanatory asymmetries. First, some asymmetries are not causal, as was seen in the example of regression toward the mean. Second, causes that fail to be sturdy can be trumped by sturdy non-causal competitors, as was the case with Chris Davis' blister when compared with regression toward the mean. Furthermore, the billiards example now shows that even the asymmetries that, according to the earlier state of the field, appeared to be the exclusive province of causal approaches also admit of an inferential rendering. So, when compared to our inference-first approach, causation-first approaches suffer several disadvantages, and enjoy no distinct advantages.

6 Conclusion

Our aims in this essay have been both critical and constructive. On the critical side, we have shown that not all explanatory asymmetries are causal asymmetries, and their standard causal diagnosis cannot be right. Indeed, the symmetry problem is really a three-headed monster—at least when viewed against the broader dialectic of causal and inferential approaches to explanation. Some asymmetries have an ineluctable causal element; some are decidedly noncausal; and others are fair game for both parties to the debate.

On the constructive side, we have sketched the broad contours of a new version of EAI—what we have called the defeasibility model of explanation. It departs from earlier versions of EAI in its melding of nonclassical and comparative components. This duet achieves its denouement in the concept of sturdiness—roughly the idea that explanations are inferences that succeed where their competitors fail. As we have shown, sturdiness is the common thread that ties together the different kinds of explanatory asymmetries.

This is but an opening salvo in a research program that we hope to develop in greater detail, by extending the defeasibility model to solve other venerable problems in the explanation literature. Earlier versions of EAI face a variety of problems.[15] How should laws be characterized? How to make sense of indeterministic explanations of improbable events, such as the fact that a person's untreated syphilis explains his paresis, despite the fact that only 25% of untreated syphilitics suffer from paresis?

Equally importantly, the symmetry problem has overshadowed two *advantages* that early variants of EAI enjoy over causal approaches. First, inferentialism is a natural way to analyze *non-causal* explanations. However, this paper has only focused on the *asymmetry* of these explanations. In the future, we hope to extend the defeasibility model to the growing stockpile of examples of non-causal explanations (Baker 2005; Batterman 2002; Bokulich 2011; Huneman 2010; Irvine 2015; Lange 2013a,b; Rice 2015; Risjord 2005).

Second, in comparison with causal approaches, earlier proponents of EAI enjoyed what we might call *Humean modesty.* Inference-based approaches argue that explanations are simply inferential relationships between certain empirical statements. Hence, competent language users can explain by wielding inferences that carry no further commitment to a substantive modal or causal ontology. As a result, EAI often avoids the various placement problems associated with modality and causality (e.g. how modality fits within a naturalistic ontology, how modal and causal claims can be known, etc.).[16]

[15] For a review of these challenges, see Salmon (1989) and Woodward (2014).

[16] Indeed, in AUTHOR CITATION, we develop a more precise account of explanation using a broadly inferentialist semantics. This approach to modal and explanatory vocabulary gives those of a Humean bent a compelling story about how one can *use* modal vocabulary without having to *represent* or be *ontologically committed* to metaphysically controversial modal

Since the demise of the covering law model, the symmetry problem hung like an albatross around the neck of proponents of EAI. This paper has sought to loosen that grip, and to allow new approaches to EAI to breathe. By solving the problem of explanatory asymmetry, we have cleared an important barrier to showing how inferential considerations latch onto explanation's deeper structures.

7 Compliance with Ethical Standards

The author declares no potential conflicts of interests with respect to the authorship and/or publication of this article. The author received no financial support for the research and/or authorship of this article.

References

Ariew, A., C. Rice, and Y. Rohwer (2014). Autonomous-Statistical Explanations and Natural Selection. *The British Journal for the Philosophy of Science*.

Baker, A. (2005). Are There Genuine Mathematical Explanations of Physical Phenomena? *Mind 114*(454), 223–238.

Barnes, E. C. (1992). Explanatory Unification and the Problem of Asymmetry. *Philosophy of Science 59*(4), 558–571.

Batterman, R. W. (2002). *The Devil in the Details : Asymptotic Reasoning in Explanation, Reduction and Emergence*. New York: Oxford University Press.

Bokulich, A. (2011). How Scientific Models Can Explain. *Synthese 180*(1), 33–45.

Brandom, R. (2008). *Between Saying and Doing : Towards an Analytic Pragmatism*. Oxford ; New York: Oxford University Press.

Brandom, R. (2015). *From Empiricism to Expressivism*. Cambridge, Mass.: Harvard University Press.

Daniel R. Braunstein and Ken Woolums. (2014). The Home Run Derby Myth. URL `https://fivethirtyeight.com/datalab/the-home-run-derby-myth`. Accessed March 18, 2017.

Bromberger, S. (1965). An approach to explanation. In R. Butler (Ed.), *Studies in analytical philosophy*, Volume 2, pp. 72–105. Oxford: Blackwell.

Bromberger, S. (1966). Why-questions. In R. Colodny (Ed.), *Mind and Cosmos: Essays in Contemporary Science and Philosophy*, pp. 86–111. Pittsburgh: University of Pittsburgh Press.

Cartwright, N. (1983). *How the Laws of Physics Lie*. New York: Oxford University Press.

Friedman, M. (1974). Explanation and Scientific Understanding. *Journal of Philosophy 71*(1), 5–19.

Huneman, P. (2010). Topological Explanations and Robustness in Biological Sciences. *Synthese 177*(2), 213–245.

Irvine, E. (2015). Models, Robustness, and Non-Causal Explanation: A foray into cognitive science and biology. *Synthese 192*(12), 3943–3959.

Kitcher, P. (1989). Explanatory Unification and the Causal Structure of the World. In P. Kitcher and W. C. Salmon (Eds.), *Scientific explanation*, Volume XIII, pp. 410–506. Minneapolis: University of Minnesota Press.

Lange, M. (2009). Why Do the Laws Explain Why? In T. Handfield (Ed.), *Dispositions and Causes*, Mind Association Occasional Series. Oxford University Press.

Lange, M. (2013a). Really Statistical Explanations and Genetic Drift. *Philosophy of Science 80*(2), 169–188.

Lange, M. (2013b). What Makes a Scientific Explanation Distinctively Mathematical? *British Journal for the Philosophy of Science 64*(3), 485–511.

entities (see Brandom 2008, 2015). This idea can be traced back to (Sellars 1957). For similar approaches to the semantics of modal vocabulary see (Thomasson 2007) and (Stovall 2015).

Lange, M. (2016). *Because Without Cause: Non-Causal Explanations in Science and Mathematics*. New York: Oxford University Press.

Lipton, P. (2004). *Inference to the Best Explanation*. International library of philosophy and scientific method. Routledge.

Mitchell, S. D. (2003). *Biological Complexity and Integrative Pluralism*. Cambridge University Press.

Rice, C. C. (2015). Moving beyond causes: Optimality models and scientific explanation. *Noûs 49*(3), 589–615.

Risjord, M. (2005). Reasons, causes, and action explanation. *Philosophy of the Social Sciences 35*(3), 294–306.

Salmon, W. C. (1989). Four decades of scientific explanation. In P. Kitcher and W. Salmon (Eds.), *Scientific Explanation*, pp. 3–219. Minneapolis: University of Minnesota Press.

Schurz, G. (1999). Explanation as unification. *Synthese 120*(1), 95–114.

Schurz, G. and K. Lambert (1994). Outline of a theory of scientific understanding. *Synthese 101*(1), 65–120.

Sellars, W. (1957). Counterfactuals, dispositions, and the causal modalities. In G. Maxwell (Ed.), *Minnesota Studies in The Philosophy of Science, Vol. II*, pp. 225–308. Minneapolis: University of Minnesota Press.

Skyrms, B. (1980). *Causal Necessity: A Pragmatic Investigation of the Necessity of Laws*. New Haven, CT: Yale University Press.

Stovall, P. (2015). *Chemicals, Organisms, and Persons: Modal Expressivism and a Descriptive Metaphysics of Kinds*. Ph. D. thesis, University of Pittsburgh.

Strevens, M. (2008). *Depth: An Account of Scientific Explanation*. Cambridge: Harvard University Press.

Thomasson, A. L. (2007). Modal normativism and the methods of metaphysics. *Philosophical Topics 35*(1/2), 135–160.

Walsh, D. M. (2015). Variance, invariance and statistical explanation. *Erkenntnis 80*(3), 469–489.

Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation*. New York: Oxford University Press.

Woodward, J. (2014). Scientific explanation. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2014 ed.).