The Shared Explanatory Deficit in Humean Accounts of Causation and Laws

Srikar P. Desemsetti

North Carolina State University

PHI 497

Dr. Travis S. McKenna

April 30th, 2026

The Shared Explanatory Deficit in Humean Accounts of Causation and Laws

Introduction

Two central questions in metaphysics concern what causation is and what laws of nature are. Humean philosophers, following the tradition inaugurated by Hume's skepticism about necessary connections, answer both in the same spirit: neither causation nor laws involve anything over and above patterns in the world. To say that one event caused another is to report a certain kind of counterfactual pattern; to say that something is a law is to say that it belongs to the best systematization of those patterns. The approach is parsimonious and avoids positing mysterious necessary connections. Yet it faces a serious challenge from the direction of scientific explanation.

Humean accounts of both causation and laws fail to ground the asymmetric, directed explanatory relations that scientific practice requires. The explanatory demand takes a different form in each case: laws are expected to explain their instances, while causal claims are expected to ground directed dependence between particular events. But in both cases, the Humean framework—by reducing its target to patterns in the mosaic—cannot account for why those patterns have explanatory force rather than merely descriptive accuracy. This is not because the Humean identifies the wrong patterns; Lewis's accounts may well get the extensions of "cause" and "law" right. The problem is that correctly identifying a pattern is not the same as explaining why it holds.

The argument that follows targets explanation as an objective feature of the world, rather than merely a pragmatic or epistemic virtue of our theories. Scientists typically treat explanatory claims as mind-independent: when a physicist says that a law explains a trajectory, or that a short

circuit explains a fire, the claim is not merely that these statements fit our interests or unify our descriptions, but that the world itself exhibits a directed structure of dependence. My argument is conditional on this widely shared assumption. If explanation is instead wholly pragmatic or interest-relative, the Humean may avoid the challenge. But if explanation involves objective, asymmetric dependence relations in the world, the Humean framework faces a serious difficulty.

We will proceed as follows. Section 1 briefly sets out Lewis's Humean accounts of causation and laws. Section 2 articulates the explanatory demand on each, arguing that despite their differences, both require a kind of grounding that goes beyond pattern-description, and defends the key principle underwriting this demand. Section 3 develops the case against Humean laws and directly addresses the strongest Humean replies. Section 4 extends the argument to Humean causation. Section 5 considers the most promising Humean responses and argues that they do not resolve the difficulty. Section 6 concludes.

1. Humean Accounts of Causation and Laws

Lewis's approach to both causation and laws rests on Humean supervenience: the thesis that all facts about a given world W supervene on the spatiotemporal arrangement of local intrinsic properties at W—the "Humean mosaic" (Lewis 1986, ix). On causation, Lewis (1973) argues that C causes E just in case there is a chain of counterfactual dependence linking them: causation is defined in terms of counterfactual dependence, counterfactual dependence in terms of the truth of certain counterfactuals ("had C not occurred, E would not have occurred"), and the truth of those counterfactuals in terms of what holds at the closest possible worlds, where closeness is measured partly by agreement with the actual laws (Lewis 1979, 47). On laws, Lewis (1983) argues that the laws are the theorems and axioms of the Best System—the deductive system that

best balances simplicity and strength as a systematization of the mosaic. As Loewer (1996) clarifies, laws are the most informative and compact summary of the patterns already present.

Both accounts share the same structure: they reduce their target to features of the mosaic and deny that anything over and above patterns is needed. This shared structure is what makes it plausible that they face a common difficulty.

2. The Explanatory Demand

My argument turns on a claim about what causal and nomological statements must do in scientific practice. This claim requires care, because causation and laws do not play identical explanatory roles.

The argumentative structure of the paper can be stated as a dilemma for the Humean. If explanation requires that the explanans stand in some respect independent of the explanandum, then Humean accounts of laws and causation fail to meet this requirement—that is the first horn, developed in Sections 3 and 4. If, on the other hand, explanation does not require such independence, the Humean must adopt a reconceived account of explanation on which systematic counterfactual dependence alone is sufficient for explanatory force. That is the second horn, and it is the Humean's most serious line of reply. Section 5 argues that this reply, despite its initial plausibility, cannot supply the independence that explanation actually requires.

Begin with laws. If L is a law, then L is expected to explain the events to which it applies. When a physicist explains why a projectile follows a parabolic trajectory by citing the law of gravitation, she offers the law as the reason why the trajectory had to go that way, given the initial conditions. The explanation is asymmetric: the law and the conditions explain the trajectory, but the trajectory does not explain the law. And it supports counterfactuals: had the

initial conditions been different, the trajectory would have been different in a specific, law-governed way. As Maudlin (2007) argues, laws are not merely patterns we happen to notice; they are supposed to govern the evolution of the world.

The case of causation is more complicated. Not every true causal dependence claim is a good explanation. I strike a match, and it catches fire. The match's catching fire causally depends on many things: the striking, the presence of oxygen, the dryness of the matchstick. But while "the match caught fire because I struck it" seems explanatory, "the match caught fire because there was oxygen in the room" seems odd—it cites a causally relevant factor that does not seem explanatorily relevant in context. Lewis himself would agree: in his 1973 paper he declines to say what "the" cause of an event is, taking the distinction between causes and background conditions to track pragmatic factors rather than metaphysical ones.

But it does not follow that causal claims carry no explanatory burden. Consider a specific causal claim: a malfunctioning electrical circuit caused a house fire. Call this claim C. C is genuinely explanatory. It supports a counterfactual—had the circuit not malfunctioned, the fire would not have occurred—and it is asymmetric: the short circuit explains the fire, not vice versa. On Woodward's (2003) interventionist framework, C identifies the short circuit as a variable whose manipulation would make a difference to whether the fire occurs, in a way that satisfies certain structural conditions. What C is on the hook for explaining, I will argue, is not merely that the short circuit and the fire co-occur in a certain pattern. C is supposed to explain why the fire occurred rather than not, given the short circuit—in a way that is directed and not symmetric. That directed dependence is what the Humean must account for. In many scientific explanations, this direction is grounded not merely in counterfactual patterns but in the presence of a process connecting cause and effect.

The common demand, then, is not that causation and laws do identical work. It is that both, when they figure in genuine explanations, must ground directed dependence relations that go beyond describing which patterns obtain. For laws, this means explaining why instances must occur given the law and initial conditions. For causal claims, this means grounding why the effect depended on the cause in a directed, non-symmetric way.

The Independence Principle

The argument that follows rests on a principle that requires explicit defense: if X is determined by Y, then X cannot explain Y. Call this the Independence Principle. It says that for an explanation to be genuine, the explanans must be, in some relevant respect, independent of the explanandum. Why accept this?

The reason is structural. Explanation is asymmetric: if A explains B, then B does not explain A. For this asymmetry to hold, there must be some respect in which A is prior to or independent of B—some feature of the explanatory relationship that makes it run in one direction rather than the other. If A is determined by B (that is, if A's content is derived from or fixed by the very facts it purports to explain), then there is no basis for privileging one direction over the other. The relationship is symmetric: B determines A, and A describes B. Neither is prior.

Consider an analogy. A class's grade-point average is determined by the individual grades. The average accurately describes a pattern across the grades, and it is counterfactually sensitive to them (had the grades been different, the average would differ). But the average does not explain why any individual student received their grade. It cannot, because the average is derived from those very grades—it has no independent standing from which to do explanatory work. To explain a student's grade, one must cite something independent: their performance on

assignments, their understanding of the material, their study habits. The same structure, I argue, applies to Best System laws and the mosaic.

One might object that this principle begs the question against the Humean, who denies that explanation requires any such independence. I address this objection in Section 5. For now, the point is that the Independence Principle is not an arbitrary philosophical stipulation but a consequence of the asymmetry of explanation—a feature that even the Humean acknowledges.

Put differently, the Independence Principle is not an additional constraint on explanation beyond asymmetry; it is an attempt to articulate what asymmetry itself requires. If the content of an explanans is fixed by the very facts it explains, then there is no principled basis for privileging the direction from explanans to explanandum rather than the reverse. The relationship becomes one of mutual determination or descriptive equivalence, rather than explanatory priority. The Independence Principle therefore makes explicit a condition that the asymmetry of explanation already presupposes.

3. The Explanatory Deficit in Humean Laws

On the Best System Account, the laws are those regularities that figure in the simplest and strongest systematization of the mosaic. Return to the projectile: a physicist explains the cannonball's trajectory by citing Newton's law. On the Humean picture, this law is the regularity that best balances simplicity and strength within the overall systematization. It describes a pattern that obtains across the totality of events. But the mosaic contains the trajectory as one of its constituent facts, and the law is determined by the mosaic. By the Independence Principle, the law cannot explain the trajectory, because the law's content is derived from facts including the trajectory itself.

A medical analogy makes the structure vivid. A hospital has a 30% postoperative infection rate. This statistic accurately describes a pattern across individual cases, and it is counterfactually sensitive to them. But citing the rate does not explain why any particular patient got an infection—it is drawn from those very cases. To explain why this patient got infected, a physician must cite something independent of the aggregate pattern: the specific pathogen, the compromised immune response, the lapse in sterile technique. The statistic describes; the independent causal story explains.

The Humean has a powerful reply: Best System laws are not mere summaries in the way that an infection rate is. As Beebee (2000) argues, the Best System Account does not simply catalog what happens; it identifies regularities that earn their place by optimally balancing informativeness and simplicity across the entire mosaic. They are counterfactually robust, projectible to unobserved cases, and form the basis of scientific inference. This makes them much more theoretically distinguished than a brute statistical summary.

This reply narrows the gap but does not close it. Grant that Best System laws are special regularities. The question is whether being the optimal systematization of patterns is sufficient to explain why those patterns hold. The law of gravitation, on the Best System Account, earns its status because gravitational regularities pervade the mosaic. But then the regularities are what make the law a law, and the law is what is supposed to explain the regularities. The explanatory relation runs in both directions at once, which is precisely what the asymmetry of explanation prohibits. The law's theoretical distinction—its simplicity, its strength, its projectibility—does not change the fact that its content is determined by the patterns it purports to explain.

Armstrong (1983) and Dretske (1977) argue that this is why laws must involve a necessitation relation between universals—something that stands behind the regularities and

makes them hold. Maudlin (2007) takes a different route, treating laws as metaphysically primitive: they do not reduce to anything in the mosaic but are fundamental features of the world that govern its evolution. These proposals differ in their positive accounts, but they share a common diagnosis: if laws are to explain their instances, they must have a status that is not derived from those instances.

4. The Parallel Deficit in Humean Causation

A parallel deficit arises for Lewis's account of causation, though as argued in Section 2, the explanatory demand takes a different form. Lewis reduces causation to counterfactual dependence, evaluated using the laws as part of the similarity metric for possible worlds. Since the laws are Best System regularities, the causal facts inherit whatever explanatory limitations the laws have. But the causation case also raises a distinct problem concerning the directed character of causal explanation.

Return to the claim that the short circuit caused the house fire. Lewis's account says we are reporting that in the closest possible worlds where the short circuit does not occur, the house fire does not occur either. This counterfactual is true and captures a genuine dependence. But the explanatory question is whether Lewis's framework can account for why this counterfactual tells us that the short circuit is responsible for the fire—responsible in a directed way that mere counterfactual correlation does not guarantee.

The Humean might reply that the counterfactual dependence itself is the explanation—nothing more is needed. But consider what directed explanation requires. Hall (2004) distinguishes two concepts of causation: dependence (counterfactual covariation) and production (a spatiotemporally continuous process connecting cause to effect). The short circuit's explanation of the fire involves both: the fire counterfactually depends on the short circuit, and

the short circuit produces the fire through a specific physical process—the electrical fault generates heat, which ignites nearby material. Lewis's account captures the dependence but has no resources for capturing the production—it treats causation as exhausted by counterfactual patterns.

Why does this matter for the explanatory deficit? Hall's distinction suggests that directed causal explanation—why the fire depended on the short circuit in a non-symmetric way—is grounded in the productive process connecting them, not merely in a pattern of counterfactual covariation. Lewis's account captures only the dependence side. It correctly handles cases like the barometer: the closest worlds where the barometer reads differently are worlds where atmospheric conditions differ, so Lewis rightly concludes the barometer does not cause the storm. That verdict is correct, and the counterfactual analysis delivers it. But delivering the right verdicts about which events are causes is not the same as capturing what makes a causal claim explanatory. The barometer case tests extensional adequacy—whether the account picks out the right causal relata. The present question is different: it is whether Lewis's account can ground the directed, asymmetric explanatory force that a causal claim carries once we have identified the cause. And here, having no resources for production, the account must locate that directionality elsewhere.

This matters for the Independence Principle in the following way. Lewis's account of causal asymmetry appeals to the asymmetry of overdetermination: effects tend to have more independent traces than causes, which is a contingent feature of our world's thermodynamic structure (Lewis 1979, 49). Lewis uses this to explain why counterfactual dependence runs in the direction it does—from cause to effect, not the other way around. But notice what this means for the explanatory claim. When we say that the short circuit explains the fire, the directional force

of that explanation is, on Lewis's picture, ultimately grounded in global statistical patterns in the mosaic: the fact that fires tend to leave more traces than short circuits, that the thermodynamic arrow points in a particular direction across the totality of events. The short circuit's explanatory priority over the fire is not grounded in anything local to their particular relationship—it is grounded in a global feature of the mosaic from which the asymmetry is read off.

Before pressing this point, it is important to distinguish two questions that can easily be run together. One is the global question of why causal relations in our world exhibit a temporal direction at all—why causation runs from past to future rather than the reverse. The other is the local question of why a particular causal claim—such as the claim that the short circuit explains the fire—is asymmetric in the way required for genuine explanation. Lewis's appeal to thermodynamic asymmetry is addressed to the global question: it tells us why the counterfactual dependence in our world runs in the direction it does across the board. But the present worry concerns the local one. When a scientist cites the short circuit to explain the fire, the directionality of that explanation is taken to arise from features of the particular causal relationship, not from large-scale statistical features of the universe as a whole. The thermodynamic story, even if correct as a global account, does not by itself deliver the local explanatory asymmetry that the particular case requires.

But this is precisely where the Independence Principle bites. The mosaic includes the fire. The global patterns that determine the thermodynamic asymmetry—and thereby determine that the short circuit is the cause and the fire the effect, rather than vice versa—are themselves partly constituted by the fire's occurrence. So the directional explanatory relationship between short circuit and fire is grounded in patterns that the fire itself helps to constitute. The explanatory asymmetry is, in the relevant sense, derived from facts that include the explanandum. And by the

Independence Principle, an explanatory relation cannot be grounded in that way: the explanans must be independent of the very facts it purports to explain. Lewis's account locates the source of causal direction at the wrong level—not in anything intrinsic to the particular short circuit–fire relationship, but in the aggregate structure of a mosaic that the fire itself partly shapes.

5. The Humean Response

The strongest Humean response challenges the Independence Principle directly. On this view, to explain an event is simply to show how it fits into the Best System—how it instantiates the laws and how it is counterfactually connected to other events via those laws. Explanation does not require any metaphysical priority of explanans over explanandum. The demand for such priority, the Humean argues, reflects a pre-scientific picture of necessitation that physics has shown to be unnecessary. Many philosophers have found this compelling: accounts that characterize explanation as the identification of systematic counterfactual dependence—precisely the resources Lewis has—have gained significant traction (Woodward 2003).

This response has real force, and it deserves careful attention. The core Humean claim is that fitting into the Best System is itself sufficient for explanation: the short circuit explains the fire because the Best System, which is the most powerful and simple systematization of the mosaic, licenses the counterfactual connecting them. There is no further fact about production or necessitation that needs to be cited—the systematic counterfactual structure just is what explanation consists in.

But this response, I argue, quietly concedes the Independence Principle rather than refuting it. Consider carefully what "fitting into the Best System" amounts to in the case of the short circuit and the fire. The fire occurs; it instantiates a Best System regularity; and that

regularity was determined by the totality of events in the mosaic, including the fire itself. The "explanation" the Humean offers reduces to this: the fire fits a pattern that was derived from events including the fire. The Humean insists that systematic counterfactual dependence is sufficient for explanation, but to say this is to redefine what explanation requires—it is to say that asymmetry need not be grounded in anything independent of the explanandum, but only in the global structure of the systematization.

The problem is that this redefinition cannot recover the directionality that genuine explanation requires. When a physicist says that the law of gravitation explains the projectile's trajectory, she means that the law, together with the initial conditions, determined the trajectory—that alternatives were ruled out by something that stands independently of the trajectory itself. On the Humean picture, the law of gravitation is itself determined by the mosaic, which includes the trajectory. The law does not constrain the mosaic; it is the mosaic's best self-description. And a self-description cannot rule out alternatives in the sense required: it can only report what is the case, not what had to be the case given something prior and independent.

The Humean has a natural rejoinder here. The redefinition of explanation is not arbitrary, the Humean will say—it is principled. The Best System is not just any pattern derived from the mosaic; it is the uniquely optimal one, the systematization that best balances simplicity and strength. When the fire instantiates a Best System law, it does not merely fit some pattern that happened to be derived from events including itself. It fits the most powerful, most compressed, most projectible account of how the world works. And that, the Humean will insist, is genuine explanation: not mere pattern-fitting, but subsumption under the best possible systematization. This is why Lewis and his followers think they can handle explanatory asymmetry: the Best

System identifies which direction the counterfactuals run, and that is all that direction in explanation requires.

This rejoinder, however, does not escape the problem—it restates it. The question is not whether the Best System is optimal, but whether optimality confers the kind of independence that explanation requires. Return to the grade average. Suppose we do not just take a raw average, but instead compute a sophisticated weighted average—one that balances recency, difficulty, and consistency in just the right way to produce the most informative and predictively powerful summary of the class's performance. That weighted average is more theoretically distinguished than a raw average. It is projectible, it supports counterfactuals, it is the "best" summarization of the grade distribution. But it still does not explain why any particular student received their grade, for exactly the same reason as before: however sophisticated the method of derivation, the summary is still derived from the individual grades themselves, including the grade in question. Optimizing the derivation procedure does not create independence between the summary and the facts it summarizes. The same point applies to the Best System. The fire's fitting the Best System is not an explanation of the fire because the Best System, however optimal, is still determined by a mosaic that includes the fire. Theoretical distinction does not substitute for independence.

The Humean program correctly identifies which patterns obtain and correctly tracks which counterfactuals are true. What it cannot do is ground the directed explanatory force that scientific practice attributes to laws and causal claims alike. The independence that explanation requires is not delivered by optimizing the way patterns are summarized—it requires that the explanans stand apart from the facts it explains in a way that the Best System, by its very nature, cannot. That is the shared explanatory deficit, and it is one the Humean cannot close.

6. Conclusion

Humean accounts of causation and laws share a common structure: both reduce their targets to patterns in the mosaic. I have argued that both fail to ground the directed, asymmetric explanatory relations that scientific practice requires.

The argument rests on the Independence Principle: if the explanans is determined by the explanandum, the explanatory relation lacks the asymmetry that genuine explanation demands. For laws, the Best System Account identifies laws as optimal systematizations of the mosaic, but a systematization determined by the facts it purports to explain cannot ground the asymmetry between law and instance. For causation, Lewis's counterfactual analysis correctly tracks which events are causes, but the directional force of its causal explanations is grounded in global thermodynamic patterns in the mosaic—patterns that the effect itself helps to constitute. In both cases, the explanatory relationship is derived from facts that include the explanandum, and no amount of theoretical optimization of the derivation procedure can substitute for the independence that genuine explanation requires.

The strongest Humean response—that explanation just is systematic counterfactual dependence, and that the Best System's optimality is sufficient to ground it—does not resolve the difficulty. The Humean's rejoinder that the Best System is not merely any pattern but the uniquely optimal one restates rather than escapes the problem: however sophisticated the derivation, a summary determined by facts that include the explanandum cannot explain those facts. Theoretical distinction does not substitute for independence. The Humean program correctly identifies which patterns obtain. What it cannot do is explain why those patterns have the directed explanatory force that science attributes to them.

References

Armstrong, D. M. *What Is a Law of Nature?* Cambridge: Cambridge University Press, 1983.

Beebee, Helen. “The Non-Governing Conception of Laws of Nature.” *Philosophy and Phenomenological Research* 61, no. 3 (2000): 571–594.

Dretske, Fred. "Laws of Nature." *Philosophy of Science* 44, no. 2 (1977): 248–268.

Hall, Ned. “Two Concepts of Causation.” In *Causation and Counterfactuals*, edited by John Collins, Ned Hall, and L. A. Paul, 225–276. Cambridge, MA: MIT Press, 2004.

Lewis, David. "Causation." *Journal of Philosophy* 70, no. 17 (1973): 556–567.

Lewis, David. "Counterfactual Dependence and Time’s Arrow." *Noûs* 13, no. 4 (1979): 455–476.

Lewis, David. "New Work for a Theory of Universals." *Australasian Journal of Philosophy* 61, no. 4 (1983): 343–377.

Lewis, David. *Philosophical Papers, Volume II.* Oxford: Oxford University Press, 1986.

Loewer, Barry. "Humean Supervenience." *Philosophical Topics* 24, no. 1 (1996): 101–127.

Maudlin, Tim. *The Metaphysics Within Physics*. Oxford: Oxford University Press, 2007.

Woodward, James. *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press, 2003.