

# Ceteris Paribus Hedges: Causal Voodoo That Works

Michael Strevens

Draft of April 2010

## ABSTRACT

What do the words *ceteris paribus* add to a causal hypothesis, that is, to a generalization that is intended to articulate the consequences of a causal mechanism? One answer, that looks almost too good to be true, is that a *ceteris paribus* hedge restricts the scope of the hypothesis to those cases where nothing undermines, interferes with, or undoes the effect of the mechanism in question, even if the hypothesis's own formulator is otherwise unable to specify fully what might constitute such undermining or interference. This paper proposes a semantics for causal generalizations on which *ceteris paribus* hedges are able to deliver on this promise: it argues that the truth conditions for a causal generalization depend in part on the—perhaps unknown—nature of the mechanism whose consequences it is intended to describe. It would follow that the truth conditions for causal hypotheses are typically opaque to their own formulators. The paper concludes with a discussion why opacity might benefit, rather than hinder, causal investigation.

## 1. Regularity and Hedge

High-level regularities—patterns of biological, psychological, social, or economic events—are, in our kind of universe, convoluted things. For example, a government’s printing money to pay its debts tends more often than not to cause an increase in the rate of inflation. But under a number of different circumstances, printing money may not cause an increase in inflation (if, for example, the extra currency is hoarded in mattresses rather than spent, so taken back out of circulation). Because these circumstances are rather diverse, an attempt to specify the economic regularity with any degree of precision will be a daunting undertaking, requiring presumably many clauses, sub-clauses, parentheses, and footnotes. Some writers have thought it plausible that an exact specification of the regularity would have to be infinitely long (Pietroski and Rey 1995, 102).

The reasons for the convoluted nature of high-level regularities are many, but one is more important than the others: the regularities that interest science tend to be the consequences of causal mechanisms, and causal mechanisms are not even close to being sure-fire things, for three reasons. First, they require certain enabling conditions to operate, but these conditions are not always present. Second, even when the conditions are right, the operation of a mechanism is always vulnerable to interference from the outside, to some orthogonal force that derails the causal train of events. Third, even when a causal process runs to completion, other causal processes may reverse or otherwise undercut its effects. (Not all dogs are four-legged in part because some are amputees.) To capture precisely the consequences of a causal mechanism, then, you must capture all such possibilities for causal failure or reversal. An exact specification of such conditions, and so of the regularity itself, must be

if not infinite then at least very long and complex.

This complexity—this intricacy in the twists and turns of the high-level regularities—poses a problem, it is generally agreed, for our understanding of the nature and workings of the high-level sciences. To introduce the problem, suppose that a statement of a high-level law should entail its corresponding Humean generalization, where by a law statement's "corresponding Humean generalization" I mean a precise specification of the pattern of events to which the law in question would give rise, if the law statement were correct. The law statement *All Fs are Gs*, for example, should entail that all actual *Fs* are *Gs* (which is why an *F* that is not *G* is regarded as a counterexample to the putative law).

From this assumption it follows that law statements must be at least as complex as their corresponding Humean generalizations. But if all (interesting and correct) high-level Humean generalizations are extremely long, complex, and convoluted, then all high-level law statements must be at least as convoluted. Two methodological problems then arise.

First, if the degree of complexity is very high, it is hard to see how researchers in the high-level sciences can formulate law statements that have any chance of being correct. Since coming to know the nature of the high-level laws requires formulating, then testing and confirming, correct high-level law statements, it follows that it is practically impossible to for science to discover the high-level laws.

Second, even if it were possible to formulate law statements of sufficient length and complexity to specify the high-level laws, it would be a formidable task to discover which of many similar complex law statements was correct. For any detail that might plausibly appear in the specification of the law under investigation, there will be two statements that differ only on that detail. Discovering the correct law statement would seem to involve as many tests as there are details.

There are several possible responses to these worries. First, you might

give up on the idea that the sciences, or the high-level sciences at least, take as one of their principal goals the discovery of laws (Giere 1999). It is not enough, note, merely to collapse the distinction between laws and other generalizations, since the problems stated above for laws look to apply equally to any class of scientifically significant generalizations. This first response, then, is quite radical: it is to abandon the scientific search for high-level regularities altogether.

A second response is to understand high-level law statements (or generalizations, or whatever you want to call them) in such a way that they do not entail their corresponding Humean generalizations. You might, for example, think that the purpose of a high-level law is to assert the existence of a tendency or a capacity (Cartwright 1989; Lipton 1999). Though a capacity presumably gives rise to various patterns of events, the assertion of its existence need not logically entail this pattern.

A third response is to understand high-level laws in such a way that their corresponding Humean generalizations need not capture the convolutions described above. High-level laws, you might hold, are “statistical” or “inexact” (Hausman 1992); their statements entail the existence of certain patterns of events, but these patterns are themselves rather rough—rough enough to accommodate, rather than to be refuted by, the results of causal breakdown. Perhaps the corresponding Humean generalizations are vague, or perhaps they allow for many exceptions (“Most *F*s are *G*s, most of the time”).

It is a fourth response that is the subject of this paper: there is (so the response goes) a certain generic hedge—in this paper, I will suppose that it is the expression *ceteris paribus*—that, when added to a law statement, imbues the statement with additional content in virtue of which it entails its corresponding generalization without its having to specify explicitly the generalization’s convolutions. For example, although *Printing money causes inflation* entails only the simple and therefore false Humean generalization *Printing money is always accompanied by a rise in inflation*, the addition of

a *ceteris paribus* hedge changes everything: *Ceteris paribus, printing money causes inflation* entails a far more nuanced Humean generalization that takes into account all the opportunities for the causal mechanism linking money-printing and inflation to be stymied by the absence of enabling conditions, interference, or reversal. I emphasize that this is a view about the content added by the linguistic item *ceteris paribus* to a *law statement*, not a view about the laws themselves. It is unremarkable to hold that the actual laws of nature necessitate real rather than fictional patterns of events; what is intriguing is the possibility that, with the help of a short Latin expression, we can frame short and simple sentences that entail actual event patterns in all their complexity.

Such a view implies that our hypotheses about the laws—our attempts to formulate true law statements—typically have what I will call opaque conditions of application. By a law statement’s *conditions of application* I mean simply a rider on a law that restricts the range of systems to which it applies. For example, in a statement Mendel’s second law of the form *If two genes lie on different chromosomes, they assort independently* the condition of application is the “if” clause, restricting the scope of the law to genes on different chromosomes. More generally, in a statement of the form *In conditions Z, Fs are G*, the conditions of application are *Z*.

A law statement’s conditions of application are partly *opaque* if they are not all known to the statement’s formulator. If the addition of a *ceteris paribus* hedge to a law statement amounts to a requirement that the causal mechanism operate unimpeded, it will in almost every case add opaque content to the statement, because the formulator of the statement—a scientist who is perhaps just beginning to investigate the nature of the mechanism in question—will normally be unable to specify explicitly a complete set of conditions sufficient for the mechanism to function, or roughly equivalently, will be unable to recognize in all cases whether or not such conditions hold. (There is no need to agonize about the characterization of opacity; any one

of a number of related notions would do to make my point.) The opacity of a hypothesis about a law, then, implies that the scientific community is not in a position to spell out the full significance of a theory that it has itself formulated.

Opacity raises two questions. First, it may seem too good to be true that, by adding the words *ceteris paribus* to your hypotheses about the laws, you can endow them with a sophisticated causal know-how that is otherwise beyond your grasp. Second, this power of *ceteris paribus* hedges may seem to be not only miraculous but useless. What is the practical significance of content in a hypothesis unless the investigators know that it is there?

The principal aim of this paper is to address the first worry, by showing that there is a simple and natural way of understanding the semantics of law statements (or at least, statements of *causal* laws) on which they can have the kind of opaque content advertised above. But I will also briefly discuss the second question, explaining how opaque content can make a significant and salutary difference to the course of empirical inquiry.

\* \* \*

The account I give of the power of *ceteris paribus* hedges is limited to their role in *causal generalizations*, by which I mean generalizations that are intended to describe certain patterns of events that are consequences of the operation of some single kind of causal mechanism, which I call the hypothesis's *target mechanism*. (It is the actual facts about this target mechanism, I will argue, that determine the significance of the hedge.)

More specifically, I focus on causal generalizations of the form *Fs are Gs* or *F-ness causes G-ness* that are, if true, presumed to hold in virtue of a causal mechanism connecting *F-ness* to *G-ness*. My generalizations are directional, then: they attribute to the antecedent an active role in a causal process that brings about the consequent. (Contrast a generalization that associates two effects of a common cause such as *Barometer drops are followed by storms*.)

A few remarks. First, as far as I can see there is no special syntactic

form that marks out the causal generalizations with which I am concerned. A natural formulation uses the generic: *Printing money leads to inflation*, *Ravens are black*, and so on; causation need not be mentioned explicitly.

Second, I sometimes use the term “causal hypothesis” rather than “causal generalization”, both for variation and to stress that my subject matter is scientists’ attempts to represent laws rather than the laws themselves. (I will, incidentally, no longer use the term “law”, which seems too grandiose in the context of many of the high-level sciences.)

Third, I will assume that the causal mechanisms in question are deterministic.

Fourth, the term *ceteris paribus* performs more than one function in the philosophical literature. Sometimes it has its literal meaning “all other things being equal”. Sometimes it is supposed to refer to a special kind of law. The phrase “*ceteris paribus* conditions” may be used to mean something similar to my “conditions of application”. And finally, the term may be understood as a kind of hedge to causal (or other) generalizations: when the words “*ceteris paribus*” are added, the meaning of the generalization changes. It is exclusively this last use that interests me here.

## 2. Truth Conditions for Causal Generalizations

The canonical form of a directional causal generalization is, I am supposing, *In conditions Z, Fs are G*. (The reader will encounter no problems, I hope, in generalizing what I have to say to other similar forms, such as *Fs are followed by Gs*, *Fs behave in manner G*, and so on.) What makes such a generalization causal is the supposition that there is a causal mechanism by which *F*-ness helps to cause the *G*-ness in question—the “target mechanism”.

Let me give a few examples, including some already mentioned above:

1. If a gas’s density and pressure are not too high, then when its temperature is held constant, its pressure is inversely proportional to its volume

( $PV = k$ , i.e., Boyle's law).

2. When genes are located on different chromosomes, they assort independently (Mendel's second law).
3. Printing money causes inflation.
4. Ravens are black (Hempel's first law).<sup>1</sup>
5. Paranoid schizophrenics hear voices.
6. Adult humans think about biological species in essentialist terms.
7. Hunter-gatherers share large food items with all members of their band.
8. The luminosity of Cepheid variable stars is directly proportional to their period.

What kind of truth conditions does a causal generalization without a *ceteris paribus* hedge have? Let me make a suggestion to get things started: the causal generalization *In conditions Z, Fs are G* means

There exists a causal mechanism that has as its only enabling conditions or components *Z* and *F*, and that brings about *G*.

But that is too strong; as inspection of the examples above shows, most causal generalizations do not specify every enabling condition and component of a mechanism. Perhaps the truth conditions should be weakened like so:

There exists a causal mechanism that has among its enabling conditions and components *Z* and *F*, and that brings about *G*.

---

1. In this and some of the other examples, there is reason to doubt that the antecedent property is supposed by the generalization's formulators to play a direct role in the putative causal mechanism in question. Such cases require additional elements to be added to the account of *ceteris paribus* hedges given in this paper; these complications will be left to another time.

This semantics for causal generalizations is, however, too weak. It makes such generalizations almost trivially true: for just about any choice of  $Z$ ,  $F$ , and  $G$  there will be some Rube Goldberg mechanism that links the three as specified.

What next, then? One possibility is to retreat from making any explicit reference to a causal mechanism in the truth conditions. The assumption that the statistical connection between  $F$  and  $G$  is due to a causal connection would be exiled to the conversational context, and the generalization itself would be given truth conditions such as “When  $Z$  holds, all  $F$ s are  $G$ ”.

A more interesting strategy is to move elements from the context to the truth conditions rather than vice versa. When scientists formulate a causal generalization, they typically have a certain mechanism in mind concerning which they wish to make their claim. They do not conceive of themselves as making an existential claim (“somewhere out there, there is a mechanism that . . .”), but a claim about the nature and consequences of this particular mechanism. Mendel’s law is about the mechanism of meiosis; Boyle’s law is about the intrinsic behavior of gases (thus not about Rube Goldberg causal pathways that travel outside the gas—e.g., the weasel sees that the pressure has increased and hits the switch to decrease the volume); “All ravens are black” is about ravens’ natural coloration mechanism (not about the efforts of un-hinged raven bleachers); and so on. When a causal hypothesis is framed, then, it is supposed to make a claim about a particular, contextually determined mechanism—the target mechanism.

I propose that the truth conditions for causal generalizations make explicit reference to this target mechanism, as follows:

The contextually determined target mechanism  $M$  has among its enabling conditions and components  $Z$  and  $F$ , and brings about  $G$ .

or more elegantly and roughly equivalently:

By way of the target mechanism  $M$ , the conditions  $Z$  and the property  $F$  bring about  $G$ ,

(These truth conditions require some precisification. Does the hypothesis assert that  $F$  and  $Z$  always bring about  $G$ , or only that they do so under further, unspecified circumstances? I will take care of this issue later.)

This proposal immediately raises a range of questions about the so-called target mechanism. How are such mechanisms picked out, especially at the early stages of scientific investigation when not much is known about the subject matter? What if the scientists in question are sufficiently confused that their investigative intentions fail to pick out a mechanism, as would happen, for example, they intended their hypotheses to describe the phlogiston consumption mechanism, or the astrological influence mechanism? What if scientists have no particular mechanism in mind? These concerns will be addressed later in the paper; for now, please give me the benefit of the doubt, as I move on to the question of *ceteris paribus* hedges.

### 3. The Semantic Contribution of *Ceteris Paribus* Hedges

What does *ceteris paribus* mean? How do the truth conditions for a causal generalization change with the addition of a *ceteris paribus* hedge? Most important, can such a hedge contribute opaque content?

#### 3.1 *Approaches to Ceteris Paribus Hedges*

There are, broadly speaking, three approaches to understanding the effect of *ceteris paribus* hedges (though the literature is large and complicated, and not every contribution fits neatly into the following schema). In each case, the hedge may be understood as responding to the problem of “causal breakdown”—the fact that causal mechanisms may not get started, or may not run to completion, or may have their effects reversed.

On the *softening* approach, adding *ceteris paribus* to the generalization *Fs are G* loosens the connection asserted to hold between  $F$  and  $G$ . The

simplest softening approach equates *Ceteris paribus, Fs are G* and *Most Fs are G*. You might add that the exceptions should not be systematic (Fodor 1991). Another possibility is that the hedged generalization means *Fs have a tendency to be G* (Kincaid 1990; Lipton 1999). Either way, adding the hedge to a causal generalization implies that, because of causal breakdown, there are some *Fs* that are not *G*, but does not imply which *Fs* these are.

On the *narrowing* approach, adding *ceteris paribus* strengthens the generalization's conditions of application, so reducing the range of systems in which the connection between *F* and *G* is asserted to hold. Thus *Ceteris paribus, Fs are G* means *In conditions Z, Fs are G*, for some appropriate choice of *Z* (Hausman 1992, §8.2). In the case of a causal generalization, *Z* will specify some or all of the conditions under which the target mechanism breaks down.

On the *annotating* approach, the addition of a *ceteris paribus* hedge does not alter what is said by a causal generalization; rather, it has a pragmatic function connected to causal breakdowns, usefully commenting on them in some respect—for example, warning the user that the corresponding Humean generalization is not exceptionlessly true, or that the generalization is in some sense incomplete (Earman and Roberts 1999 and perhaps Pietroski and Rey 1995).

Only on the narrowing approach, you will observe, can a *ceteris paribus* hedge add to a causal generalization the content needed to entail a convoluted Humean regularity: on the pragmatic approach, the hedge has no effect on the generalization's content, while on the softening approach, the hedge modifies the generalization's content, but not in such a way as to specify which *Fs* are not *G*, and so not in such a way as to capture the contours of causal breakdown. It is a narrowing approach, then, that I will offer in what follows. (I do not exclude the possibility that there is something to the softening and pragmatic approaches—a hedge might have more than one function.)

### 3.2 Truth Conditions for Hedged Generalizations

The truth conditions for *Ceteris paribus*, in conditions  $Z$ ,  $F$ s are  $G$  are, I propose, as follows:

When conditions  $O$  hold, then by way of the target mechanism  $M$ , conditions  $Z$  and the property  $F$  bring about  $G$ ,

where  $O$  is the set of conditions required for the successful operation of  $M$  (apart from  $Z$  and  $F$ ). If you compare these truth conditions to the truth conditions for an unhedged generalization, which were as follows:

By way of the target mechanism  $M$ , the conditions  $Z$  and the property  $F$  bring about  $G$ ,

you will see that the impact of the hedge is to add an additional rider, restricting the scope of the generalization to cases in which the conditions  $O$  hold. This is therefore, as promised, a “narrowing” account of *ceteris paribus* hedges.

What are these conditions  $O$ , these “conditions required for the successful operation of the target mechanism”? They are, I stipulate, the minimal set of conditions necessary to guarantee that (a) the enabling conditions for the mechanism hold, (b) nothing interferes with the mechanism’s operation, and (c) nothing reverses or undoes the effect of the mechanism’s operation. When the operation conditions hold, then, a deterministic mechanism is guaranteed to operate successfully.<sup>2</sup>

Observe that the operation conditions in question are the actual operation conditions for the actual mechanism that the generalization’s formulators intend it to describe. They are not the conditions that the investigators think

---

2. I do not think that the operation conditions would have to be infinitely long if spelled out in English, for the reasons given by Earman et al. (2002). But even if they were, that would be no objection to the truth conditions for hedged hypotheses given here; on the contrary, it would be an advantage of the *ceteris paribus* language that it allows us to state finitely hypotheses that would otherwise consume infinite resources to specify.

are required for the mechanism's operation; they are the conditions that really are required for operation.

Typically, scientists will have only very partial knowledge of the target mechanism's operation conditions. If, thanks to the introduction of a *ceteris paribus* hedge, a hypothesis about a mechanism contains in its truth conditions a specification of all the conditions of operation, then the content of the hypothesis is in part opaque to its users. In other words, scientists literally do not comprehend the content of their own hypotheses. Or rather, they do not comprehend *all* the content; they certainly comprehend some, and this is good enough, as I will eventually show, for their purposes.

How is opacity possible? If my semantics of hedged hypotheses is correct, then the *ceteris paribus* hedge reaches out, as it were, into the world, grabs the facts about the operation conditions that are in part unknown to science, and puts them into the hypothesis itself. The result is that the hypothesis restricts its claims to situations in which the operation conditions hold. Provided that there is a fact of the matter about the target mechanism and about its operation conditions (much more on this below), there is nothing inherently mysterious about such a semantics. But it has two peculiar consequences. First, the formulators of a causal hypothesis may not know what predictions the hypothesis makes about any given  $F$ , because they do not know exactly what conditions must hold in order for the hypothesis to predict that a given  $F$  is a  $G$ . Second, if everything goes according to plan, a causal hypothesis will entail a Humean generalization that successfully traces the convoluted contours of a high-level regularity, even though the formulators of the hypothesis are themselves unable to specify those contours. Exact high-level truths can be formulated, then, from a position of relative ignorance.

I expect you have some questions. How does this opacity mesh with the traditional scientific method of hypothesis testing? Why think that there is a fact of the matter about the "target mechanism's conditions of operation"? Is there any evidence from science that *ceteris paribus* hedges work this way?

Why would you want to introduce opaque content to your hypotheses, in any case? I will try to answer them all, in reverse order: the fourth and third in section 3.3; the second in section 4; the first in section 5.

### 3.3 *Opacity in Science*

Why opacity? Or more exactly, why restrict your causal hypotheses to cases in which your target mechanism's operation conditions hold even when you cannot recognize whether or not the conditions hold? The answer lies in the functional characterization of a causal hypothesis given above: a causal hypothesis is supposed to specify the consequences of the operation of a certain causal mechanism. It follows that the hypothesis ought to make predictions only about cases in which the causal mechanism operates successfully. It is simply not in the business of specifying what happens when the mechanism does not run to completion or has its effects undone. So, if it is to perform its function efficiently, it ought not to pronounce on such cases. In short, it ought to say *When the mechanism operates, you get Fs that are G*. But if the mechanism is incompletely understood, such a claim must be opaque: the investigator does not know exactly what the mechanism consists in, and so they cannot recognize in every case whether the mechanism has operated or not.

A causal hypothesis is opaque, then, because the subject matter of the investigation is opaque. Investigators want to understand the behavior of mechanism *M*, so they quite appropriately formulate hypotheses that restrict themselves to describing the behavior of *M*. But because they do not understand exactly what *M* is, their hypotheses will contain a restriction that is also incompletely understood.

What reason is there to think that causal hypotheses in science actually contain opaque conditions of application? Let me consider two examples.

*Ceteris paribus, all ravens are black*. These words are normally intended to describe the effects of the natural raven coloration mechanism, whatever

it may turn out to be. Coloration that is clearly not due to the mechanism is therefore not considered relevant to the truth of the hypothesis: a raven bleached white is no refutation of the hypothesis, because the bleached raven's color is not due to the natural mechanism.

Now imagine a group of scientists testing the raven hypothesis who are unaware that one of the enabling conditions for the coloration mechanism is the presence of adequate copper in the ravens' diet. They come across a group of ravens living in a copper-poor environment; the ravens are gray.<sup>3</sup> Are these ravens proof of the hypothesis's falsehood? To answer this question, suppose that the scientists discover ten years later that copper is necessary for the proper operation of the raven coloration mechanism. Will they retrospectively regard the gray ravens as having refuted the hypothesis that they entertained a decade before?

The answer, I suggest, is that they will regard the gray ravens as irrelevant, because their hypothesis was intended to describe the consequences of the natural coloration mechanism, and the grayness of the ravens was no more a product of that mechanism than the whiteness of bleached ravens. Thus, they will regard their hypothesis as having had, at the time of the discovery of the gray ravens, an implicit restriction that excused it from predicting that these ravens would be black. In other words, they will regard their hypothesis as having had an implicit rider saying roughly *Provided that there is enough copper in the diet . . .* This is a rider that they were incapable of spelling out at the time; it therefore endowed their hypothesis with opaque content.

But they ought not to be surprised at this: they intended at the time that their hypothesis apply only to the products of a particular causal mechanism, and they knew at the time that their knowledge of the mechanism was incomplete. Thus they intended all along that their hypothesis have a rider

---

3. This is a fictional case. Copper is the key ingredient of tyrosinase, the enzyme that catalyzes the production of the pigment melanin, which is responsible for ravens' blackness. But as far as I know, if ravens get so little copper that they lack tyrosinase, they will be dead rather than gray.

with opaque content—that it have a rider with the content *Provided that the conditions of operation for the natural coloration mechanism hold . . .*, where their knowledge of such conditions was clearly incomplete.

As another example, consider the hypothesis that the absolute magnitude  $M$  of Cepheid variable stars is related to their period  $P$  by the formula  $M = -2.81 \log(P) - 1.43$ . For many years a numerical relationship of this sort was known, but the actual mechanism responsible for Cepheids' variation was unknown. Suppose that during this time, a variable star is found that fits the profile for Cepheids (in its spectral type, the qualitative aspects of its variation, and so on) but that does not quite fit the formula. Is this star a counterexample to the hypothesis? At the time of its discovery, it may certainly seem so. But now suppose that it is decades later, and it has become clear that the variability of the star in question is caused by a mechanism different in some ways from the mechanism underlying the variability of the classical Cepheids. Something like this has happened several times in the history of the study of variable stars; in each case, once the causal facts became known, the quantitatively anomalous star was treated not as a counterinstance to the period/magnitude hypothesis but as a new kind of variable star lying outside the scope of the hypothesis.<sup>4</sup>

There are various ways to explain this methodological trend. You need not appeal to *ceteris paribus* hedges; you might rather propose that the term “Cepheid” was all along intended to be individuated by the causal mechanism underlying the variability of a certain “baptismal group” of stars with respect to which the term was introduced. Which explanation of the trend is chosen does not affect my principal contention, which is that the content of the original hypothesis was partially opaque, and that the opacity was contributed by the unknown nature of an actual causal mechanism. More

---

4. The facts are not quite as simple as in my hypothetical example, but some close approximations would include among others: the separation of the type II Cepheids from the classical Cepheids, the separation of the  $\delta$  Scuti and SX Phoenicis stars from the RR Lyrae variables, and perhaps to some extent the separation of the RR Lyrae stars from the Cepheids.

specifically, the hypothesis declared itself to be describing the consequences of only the mechanism in question; it therefore made no predictions about a star with a different variability mechanism, even though the formulators of the hypothesis were unable to distinguish stars with the one mechanism from stars with the other.

### 3.4 *Are Hedged Generalizations Trivial?*

Could it be that, by contributing to a causal generalization conditions that guarantee the successful operation of the target mechanism, a *ceteris paribus* hedge trivializes the generalization? (The grandfather of all such worries is the concern that *Ceteris paribus, Fs are G* means *Fs are G, except when they are not*, in which case a hedge transforms its generalization into an empirically vapid analyticity.)

On my view, a hedged causal generalization is not at all trivial. Hedging *All ravens are black* may protect it from some potential counterexamples—from section 3.3’s gray ravens, for example—but it is nevertheless subject to empirical refutation. What the hypothesis says is that the natural coloration mechanism for ravens makes them black. Such a claim could well have turned out to be false. We might have discovered that all ravens are naturally gray (the black ones we saw first were suffering from a rare ailment), or more plausibly that ravens come in a variety of colors. The effect of a *ceteris paribus* hedge is to focus a causal generalization on the consequences of a particular causal mechanism. It makes an empirically substantive claim about those consequences—perhaps, as in the case of the Cepheids, a quantitative claim. If that claim is false, the hypothesis is false. In short, while a hedge “protects” a causal hypothesis from refutation by states of affairs not caused by the mechanism in question, states of affairs that are caused by the mechanism may, and often will, disconfirm the hypothesis.

Here is a somewhat more subtle worry. In the post-war period, many economists came to believe that there was a robust relationship between

inflation and unemployment: the higher the inflation rate, the lower the unemployment rate. Call this *Phillips' hypothesis*. Confidence in Phillips' hypothesis may explain national policies in the late 1960s and early 1970s of tolerating high inflation in order to boost employment (though there were other forces at work as well). The consequence was stagflation: contrary to Phillips' hypothesis, both inflation and unemployment increased. Many economists would say that the hypothesis is refuted.

There is, however, a causal mechanism that has precisely the consequences stated by Phillips' hypothesis. That mechanism has as one of its enabling conditions that inflationary expectations should remain constant (a condition that is unlikely to hold in the real world if the government is manipulating the inflation rate). But then if Phillips' hypothesis allowed the benefit of *ceteris paribus* hedge, you might think it will count as true, in virtue of this mechanism. This seems to be a case where a *ceteris paribus* hedge makes it too easy for a hypothesis to qualify as correct.

My reply is that if Phillips' hypothesis were intended to capture a causal mechanism in which inflationary expectations were constant (whether the mechanism were known to have this enabling condition or not), then it would be true. In reality, it was not so intended: it was intended to apply to the real causal mechanisms driving the post-war economy in the West, mechanisms that allowed all too easily for a change in inflationary expectations. What it says of these mechanisms is false. Thus, the hypothesis is false.

The moral of the story is that it is not enough, for a hedged hypothesis to be correct, that there exist some mechanism whose consequences it captures. The mechanism in question must be the mechanism to which the hypothesis is intended to apply. Let me next turn to the question of how such intentions typically function in science.

## 4. Picking Out Mechanisms

The power of *ceteris paribus* hedges to capture well-defined, finely determined sets of conditions for causal breakdown hinges on scientists' ability to pick out well-defined, finely determined mechanisms as the objects of their inquiry, and thus as the subjects of their causal generalizations. I do not claim that scientists *always* succeed in picking out a determinate target mechanism, and thus a determinate set of operation conditions, for their causal hypotheses—more on this below—but I do claim that they *often* succeed, even under conditions of considerable ignorance. How is this possible?

### 4.1 *Examples of Mechanism Determination*

Let me give some examples, starting with the Cepheid variable stars. When the generalization relating the luminosity and period of the Cepheids was first formulated by Henrietta Leavitt, nothing was known of the mechanism responsible for their variability (as noted in section 3.3). Leavitt noted that a number of variable stars in the Small Magellanic Cloud showed a qualitatively very similar pattern of variation, “diminishing slowly in brightness, remaining near minimum for the greater part of the time, and increasing very rapidly to a brief maximum” (Pickering 1912, 1), while also fitting the striking luminosity/period relation. She conjectured that variable stars in our own galaxy that showed the same variation pattern owed their variability to the same mechanism, and so would also fit the luminosity/period relation. (She was at first unable to test this hypothesis, because she did not have a sufficiently reliable way of determining the distance and therefore the absolute magnitude of the closer stars. She did not know the distance to the Magellanic Cloud stars, either, but she did know that they were all roughly the same distance.)

Leavitt's hypothesis (now sometimes called “Leavitt's law”) is, I think you will agree, supposed to characterize the consequences of a certain unknown

mechanism for variability. The identity of that mechanism is determined by something like the following intention in Leavitt's and her readers' minds: the hypothesis should describe the consequences of *whatever mechanism causes the variability of the Magellanic stars in Leavitt's study*. (You might add: and also causes the variability of those closer stars that display the same pattern of variation.) The scope of Leavitt's hypothesis is therefore all variable stars having the same mechanism for variation as the Magellanic stars.

Three things are needed for this method of determining a mechanism to succeed. First, there must be a "baptismal group" of exemplars. Second, there must be an observer-independent "same mechanism as" relation, that is, a criterion for individuating mechanisms that is capable of determining facts of the matter about which stars do and do not share a certain mechanism for variation without additional input from the formulator of the hypothesis, who might not have a clue about the causes of stellar variation. Third, there must be a single mechanism causing the behavior of all or almost all of the members of the baptismal group—it must not be the case that there are several different mechanisms, none statistically dominant. I will discuss the question of mechanism individuation, along with the question of what happens when the third criterion is not satisfied, shortly.

As a second example, consider *Ceteris paribus, ravens are black*. The hypothesis is, as I have observed, supposed to describe the consequences of the natural mechanism for raven coloration—thus, ravens painted white, bleached, or otherwise artificially colored do not qualify as counterinstances. It appears that in this sort of case, we have a readymade phrase—"natural mechanism"—that is capable of picking out a certain mechanism as the subject matter of generalizations like the raven hypothesis, without our having much idea how that mechanism works.

It is a very interesting question how the term "natural" functions referentially. Although I have my own ideas, I will not lay them out here.<sup>5</sup> Let

---

5. But I will say that talk of the "natural mechanism" can succeed in picking out a certain

me simply note that in some cases, we have a vocabulary that allows us to pick out determinate mechanisms “by name” (compare our ability to pick out natural kinds by name).

We can also pick out and formulate hypotheses about the consequences of unnatural or pathological mechanisms. Consider the following causal generalization: *Ceteris paribus, paranoid schizophrenics hallucinate voices*. It is intended to describe the result of a causal process that is in some sense typical of paranoid schizophrenia, but that is as far as we know not natural. The mechanism is also not yet at all well understood. How do we pick it out? We identify a fairly well-defined class of individuals sharing certain symptoms, the paranoid schizophrenics, without having any understanding of the causes of schizophrenia, and we specify that our generalization is supposed to capture the consequences of the mechanism that is responsible for the symptoms in most individuals in the group. The case is similar to the Cepheids, then, with two exceptions. First, rather than a small “baptismal class” we begin with a large class that we hope contains all or almost all individuals in which the mechanism in question is at work. Second, we are consequently perhaps rather less confident about the possibility that all individuals in the class experience their symptoms for exactly the same reason; we therefore intend our generalization to describe the mechanism that is responsible for the symptoms in the *majority* of cases, but we do not require that it be a mechanism that causes the symptoms in *almost all* cases.

The schizophrenia generalization raises the possibility that elements of a causal generalization itself may be used to pick out its mechanism: the mechanism that psychiatrists have in mind as the target of *Paranoid schizophrenics hallucinate voices* might be determined by their pointing to a certain class of patients saying “whatever mechanism causes those people to hallucinate voices”. Does this not revive the worry that opaque content makes causal

---

mechanism even if various biological or metaphysical assumptions about naturalness are false.

generalizations trivial? In the case at hand, does not the existence of the target mechanism—given the way that it is picked out—guarantee the truth of the generalization? It does. But it does not follow that the generalization is trivially true. Indeed, it fails to be true under just the circumstances you would want, namely, if the mechanism or mechanisms responsible for paranoid schizophrenia do not cause its sufferers to hear voices (or if there are no such mechanisms). How can this be? The generalization’s failure to be true is not due to its saying something false about its target mechanism, but due to its having no target mechanism. (This sort of defect, which is distinct from falsehood, is discussed in section 4.3.) A causal generalization’s picking out its target mechanism in this self-regarding way does not, then, make it any easier for the generalization to be true; consequently, such pickings out may safely be allowed.

So far, I have discussed examples in which the ignorance of a hypothesis’s original formulators is profound: Leavitt, the original observers of raven color, and twentieth century psychiatrists had very little idea as to the structure of the mechanisms that served as their objects of inquiry. But in many cases, investigators do have some particular workings in mind. This is often true in economics: when economists propose a hypothesis such as *Printing money leads to inflation*, they are able to describe to some extent, if not completely, how the mechanism works.

#### 4.2 *Individuating Mechanisms*

If you have an observer-independent criterion to individuate mechanisms—if you have a “same mechanism as” relation that can sort mechanisms based on the actual causal facts, even when they are unknown to you—then it is relatively easy to point to a particular mechanism as the subject of your investigation, even when you know very little about the workings of that mechanism. As the cases of Cepheids and schizophrenia show, you need only to find a group of exemplars, and specify that your hypothesis concerns the

mechanism at work in most or all of those exemplars. Where, then, does the mechanism individuation criterion come from?

Let me begin by describing what such a criterion must do. On the one hand, it must distinguish different mechanisms, independently of our knowledge of causal differences. It must be capable of determining that the mechanisms causing luminosity variation in classical Cepheids, RR Lyrae variables, and Wolf-Rayet stars are different, so that when we say “Same mechanism as that”, pointing to  $\delta$  Cephei, our words pick out other Cepheids but not RR Lyrae or Wolf-Rayet types. On the other hand, certain differences must be ignored. Every Cepheid is different—in size, in the precise details of its composition, and so on—but we do not want these differences to count for the purposes of mechanism individuation, or else when we say “Same mechanism as that” pointing to  $\delta$  Cephei, we will pick out only that single star.

To specify an observer-independent criterion for mechanism individuation, then, is to specify a criterion for determining which causal facts matter and which do not in deciding the question of “sameness of mechanism”. It is, in other words, to find an observer-independent standard for causal coarse-graining.

I propose that two phenomena are brought about by the same causal mechanism just in case they have the same causal explanation. The classical Cepheids have the same mechanism for variability, then, because their variability is explained in the same way. The Cepheids and the Wolf-Rayet stars have a different mechanism for variability because their variability is explained in a different way. The causal facts that matter for the purposes of mechanism individuation are, in other words, the explanatorily relevant facts. Differences among the Cepheids do not make for differences in mechanism because they are not explanatorily relevant differences—they are not differences with respect to the kinds of factors that would be cited in an explanation of Cepheids’ pattern of variation. The differences between the

Cepheids and the Wolf-Rayet stars, by contrast, include factors explanatorily relevant to their variability.

In order to find a mechanism individuation criterion, then, simply look to the literature on causal explanation, with the following desiderata in mind.

First, you will need an account of explanation on which an explanation takes the form of a causal model, by which I mean a generic description of a causal process, thus of a mechanism type. Almost every proponent of the causal approach to explanation would, I think, claim that their account can be configured to satisfy this demand.

Second, you will need an account that will result in an appropriate coarse-graining, thus an account on which not every causal detail is explanatorily relevant to a system's conforming to a causal generalization. The right account will, for example, give exactly the same explanation for any Cepheid's conforming to Leavitt's hypothesis, despite the small causal differences between different Cepheids. This desideratum rules out Salmon's (1984) account of explanation, on which all the causal details are explanatory.<sup>6</sup>

Third, you will need an account on which the facts about explanatory relevance are sufficiently observer-independent to carve out the mechanisms without the close supervision of the scientific community. This likely rules out the sort of account offered by van Fraassen (1980), on which the facts about explanatory relevance are determined in part by a relation specified individually for each explanatory inquiry (though in principle, you could implicitly associate such a relation with each of our causal generalizations).

What is left? Many accounts of scientific explanation in the literature satisfy the desiderata—to name a few, those of Lewis (1986), Woodward (2003), Strevens (2009) and even, though it is not usually considered a causal account, Kitcher's (1989) unification account. On any of these views, many details of the causal history of any given Cepheid, such as the precise trajectories of molecules, are irrelevant to the explanation of its variation. Thus you have a

---

6. Salmon later recants on this very question (Salmon 1997).

basis for a coarse-grained scheme of mechanism individuation.

More specifically, apply any of these accounts of explanation to the task of explaining some particular Cepheid's varying in accordance with Leavitt's hypothesis. You will get a causal model that specifies just those causal factors that were explanatorily relevant to the star's fitting the hypothesis. I suggest that you will get precisely the same causal model for every Cepheid. This model may be regarded, then, as the schema for the operation of the Cepheid variation mechanism.

A mechanism's conditions of operation are whatever conditions are required for the corresponding causal model to be instantiated—which is to say, a mechanism's causal model and its list of operation conditions are one and the same. I make no distinction, then, between aspects of a causal process that are intrinsic to the mechanism and aspects that are extrinsic but required for the mechanism's operation. Everything is intrinsic, which is to say, I use the term *mechanism* so that everything that is causally relevant to a mechanism's operation is a part of the mechanism. This sense of *mechanism* does not quite correspond to the everyday sense of the word—it is a term of art—but it is the artful, not the everyday, notion that best suits the semantics for hedged causal generalizations advocated in this paper.

Another way in which the present notion of mechanism is artful is its indifference to mechanisms' spatiotemporal or organic integrity. The explanation of, say, a stock market crash may span many disparate and otherwise unrelated events that come together to cause the crash; these converging causal chains therefore constitute the mechanism responsible for the crash. Mechanisms do not always come in boxes.

To conclude: let the observer-independent criteria for explanatory relevance, whatever they are, individuate mechanisms. For every distinct causal explanation of the instantiation of a property *G*, then, there is a distinct mechanism by which *G* can be caused; these are the mechanisms whose consequences causal hypotheses are intended to describe.

### 4.3 Failures of Mechanism Determination

When everything goes well, a scientist who formulates a causal hypothesis succeeds in picking out a target causal mechanism, the consequences of which the hypothesis is supposed to describe. But what if things go wrong?

One possible, and perhaps not uncommon, problem is for the scientist to pick out a group of related mechanisms rather than a single mechanism. Suppose, for example, that researchers frame the hypothesis *Diabetes causes hypoglycemia*, intending it to pick out a consequence of a pathological causal mechanism operating in a group of patients that include both type 1 and type 2 diabetes sufferers. Unknown to the researchers, there are two different mechanisms operating in their sample: in type 1 diabetes, the body has lost its ability to produce insulin, while in type 2 diabetes, insulin is (usually) produced normally, but cells have acquired “insulin resistance”, that is, an inability to use insulin effectively.

The hypoglycemia hypothesis, then, has no determinate target mechanism. Suppose that the hypothesis is hedged; what then? A *ceteris paribus* hedge is supposed to restrict the scope of a hypothesis to cases in which the target mechanism’s conditions of operation hold, but the two diabetes mechanisms have somewhat different conditions of operation. There are various ways to deal with the problem. You might, for example:

1. Understand a *ceteris paribus* hedge as limiting the scope of the hypothesis to systems in which the intersection of the two sets of conditions of operation hold.
2. Understand the significance of the *ceteris paribus* hedge as varying with its application. When the hypothesis is tested against the medical history of a type 1 diabetes patient, the hedge picks out the operation conditions for the type 1 mechanism; likewise for a test against a type 2 patient.
3. Understand the *ceteris paribus* hedge as adding indeterminate condi-

tions of application to a hypothesis, or perhaps as adding the intersection of the two sets of operation conditions along with some further indeterminate conditions.

I will not advocate any one of these proposals; rather, I suggest that they all do an adequate job of capturing the methodological implications of the situation: while the foundations of the investigation are not entirely secure, and a hedged hypothesis will consequently not have quite the kind of content that its formulators suppose it to have (i.e., a restriction to the operation conditions for a single kind of mechanism), the hypothesis can nevertheless play a useful role in the ongoing investigation. You might compare the case to one in which a hypothesis contains a theoretical term with indeterminate reference, such as *jade* in the early days of geochemistry (so the story goes) or *mass* in pre-relativistic physics (Field 1973). Provided that the scope of their indeterminacy is not too wide, such terms are, far from being scientific dead weight, extremely valuable organizational tools in the period before their limitations are revealed by the advance of scientific knowledge.

A more serious deficiency in a causal hypothesis is its having no target causal mechanism at all. Consider, for example, the case of “hysteria” in women, a common diagnosis in Victorian Europe that was thought to account for a range of symptoms, such as faintness, insomnia, and irritability, but which is no longer thought to have any medical basis. A scientist might frame the hypothesis *Hysteria causes insomnia* intending to capture a causal mechanism common to most hysteria patients, but because the symptoms of hysteria patients (where they were real at all) were due to a wide array of causal mechanisms having little in common, such an intention seems likely to fail to pick out any target mechanism. The generalization is what I will call *baseless*.

On the semantics for causal generalizations given above, a hedged generalization has truth conditions of the form: *When conditions O hold . . . the property F brings about G*, where *O* are the target mechanism’s conditions for

operation. If the generalization has no target mechanism, it therefore has a gaping hole in the middle of its truth conditions. It is semantically defective. (Generalizations with indeterminate mechanisms have a similar sort of hole, but it is not so gaping and can be patched.)

What to do? Nothing, I suggest. The sort of causal inquiry that gives rise to baseless hypotheses is seriously flawed by anyone's lights; according to my semantics, the hypotheses themselves have a flaw that simply reflects the fundamental problem, that scientists are investigating the consequences and structure of a putative mechanism that does not exist.

It is an interesting methodological question, of course, how scientists deals with such cases—how they eventually back out of such investigative dead-ends. On my view, they must do so by recognizing that many of their hypotheses are not false but semantically ill-formed. But this is not entirely new; the same must be said for any branch of scientific theory in which some theoretical terms fail to refer, such as phlogiston theory, biorhythmics, and indeed the study of “hysteria”.<sup>7</sup>

## 5. Hedge and Method

A *ceteris paribus* hedge contributes to a causal generalization's truth conditions a restriction in the generalization's scope that goes beyond anything represented explicitly: the hedge confines the generalization's range of application to cases in which the operation conditions for the target mechanism hold, that is, to cases in which the target mechanism is guaranteed to operate without interference or reversal. Such a hedge is very convenient if it is your aim to spell out, with minimal verbiage, high-level causal laws that entail their corresponding Humean generalizations, and so that trace the intricate boundaries of high-level regularities.

---

7. Unless, of course, the terms in question have a “descriptivist” rather than a “Millian” semantics—a complex issue that I will glide over here.

The methodological advantage of hedges is less clear, I think, when you are very far from being able to specify the laws and regularities under investigation, that is, when you are first formulating hypotheses about the consequences of the causal workings of the world, typically in great ignorance of those workings. In such circumstances, your hedged hypotheses will contain content that is to a considerable extent opaque to you. The hypotheses will restrict their scope to conditions in which their target mechanisms operate, but you will have very little idea what this restriction amounts to. How can such a restriction possibly be of any use to you? Indeed, how can opaque content not put you at a disadvantage, by obscuring the question whether any particular piece of data confirms or disconfirms the hypothesis?

Before I address these questions, let me remind you that the opaque content of a *ceteris paribus* hedge looks in principle to be methodologically apt: if you have framed a hypothesis with the sole aim of describing the consequences of a certain causal mechanism, it is natural to understand the hypothesis as pronouncing only on cases in which the mechanism in question actually operates—just as a hypothesis about raven color ought only to make predictions about the color of ravens, and not about the color of other birds. The question is whether, in virtue of the restriction's opacity, it is in practice useless or worse.

To the first question, then: is an opaque hedge useless? Only that which is internal to science can affect the course of science—so you might suppose. Opaque conditions of application, even if present, are therefore simply irrelevant to our understanding of confirmation and other elements of scientific method.

This chain of reasoning goes wrong at the *therefore* in the last sentence. Although the content of unknown conditions of application cannot make a difference to scientific procedure, knowledge of their possible existence can. Consider my canonical hypothesis *Ceteris paribus, in conditions Z, all Fs are G*. Suppose that the formulators of the hypothesis are confronted with an

apparent negative instance of the hypothesis, that is, an  $F$  in conditions  $Z$  that is not  $G$ .

If the causal hypothesis's truth conditions are not at all opaque, then its conditions of application are  $Z$  and nothing more. It follows that the scientists have no choice but to regard the datum as a negative instance of their hypothesis (assuming, for the sake of the argument, that they are certain that the datum is indeed an  $F$ , that conditions  $Z$  do indeed hold, and that the datum is indeed not  $G$ ). The hypothesis must be abandoned, and one of three options endorsed:

1. Abandon the research program as ill-advised,
2. Continue testing with a new hypothesis created by relaxing or softening the connection between  $F$  and  $G$  asserted by the old hypothesis, for example: *In conditions  $Z$ , most  $F$ s are  $G$* , or
3. Continue testing with a new hypothesis created by narrowing the scope of the old hypothesis, that is, by revising the old hypothesis's explicit conditions of application so that they rule out the negative instance.

The possibility of opaque conditions of application opens up a further option: treat the datum, at least provisionally, as not a genuine instance, thus not evidentially relevant, on the grounds that the conditions required for the operation of the mechanism under investigation were not satisfied. The possibility of opacity legitimizes, then, a methodological response to apparent negative instances that is simply not possible in a science conducted using strictly transparent hypotheses.

Further, the truth conditions I have offered not only open up a new methodological option but provide guidance on whether or not it should be exercised: you may dismiss an apparent negative instance only if you have good reason to think that the conditions required for the mechanism's operation did not hold. Or more generally, your subjective probability that the

apparent exception is not a genuine counterinstance should be proportioned to your subjective probability that the operation conditions did not hold.

Of course, you will not know enough about the mechanism to be sure that the conditions of operation did not hold, but in many cases you can speculate shrewdly. Upon finding a small population of gray ravens living in a copper-poor environment, for example (see section 3.3), you might reason that ravens elsewhere are black with such regularity that the grayness of your sample is more likely to be a consequence of developmental privation than a sign of indeterminacy in the coloration mechanism. Other evident abnormalities due to the same deficiency might lend weight to this diagnosis. The possibilities for educated guesswork are legion, even if you know next to nothing about the internal workings of the mechanism in question.

There is no question, then, that opaque conditions of application have an impact on the course of scientific investigation. But is it a salutary impact?

Let me sketch an argument that opacity is unhealthy. To ascertain whether a piece of evidence is relevant to a hedged causal hypothesis, you must determine whether the hypothesis's conditions of application are satisfied. If the conditions of application are opaque, then you may not have all the resources you need to make this judgment successfully, which is to say that the relevance of the datum to the hypothesis may be itself opaque. This amounts to a kind of evidential externalism, on which the relevance of a piece of putative evidence to a hypothesis—the question whether the evidence confirms the hypothesis, disconfirms it, or neither—depends on external matters of fact that an investigator may be not be in a position to grasp.

Evidential externalism is certainly not a standard position in confirmation theory. I would guess that the majority of confirmation theorists would sooner concede to Williamson (2000) that the facts as to what your evidence is are externalist (because observation statements must be true to count as evidence, but few or no evidence statements are known to be true with certainty) than concede that the facts about whether or not the confirmation

relation holds between two sentences are subject to empirical investigation.

It does not follow that evidential externalism is mistaken. Indeed, the prospect is with us already. The question whether, say, a specimen of blue metal is evidentially relevant to the hypothesis *All gold is yellow* turns on the question whether the specimen is itself a sample of gold or something else. On the commonly accepted views about the reference of natural kind terms, chemical naifs may be unable to determine whether a substance lies within the extension of one of their own natural kind terms. If they are unable to decide whether the substance is gold, they are unable to decide whether it is a counter-instance to, or entirely irrelevant to, their hypothesis. Evidential relevance is for them partially opaque, hence external.

Still you might ask: what good can evidential externalism be? That is a question that has a long and interesting answer; I will not, however, try to give the answer here. Let me rather make the following observation. At the very beginning of this paper, I defined a causal generalization as an attempt to describe the consequences of a particular causal mechanism. Behind this stipulation lies an assumption about the methodology of science, that scientific practice individuates causal inquiry by mechanisms—that a scientist who asks about causal consequences is in the first instance asking about the consequences of a particular mechanism.

Grant me this assumption: the natural unit of scientific inquiry is the causal mechanism. What kind of evidence is relevant to such an inquiry? Clearly, the results of the mechanism's operation. What kind of evidence is irrelevant? Among other things, effects that look like they might be due to the mechanism's operation, but that turn out to be produced by some other mechanism.

You have an externalist standard for evidential relevance right there. Relevance and irrelevance depend on what is and is not produced by the mechanism under investigation, but these facts are potentially opaque. The only reasonable standard of evidential relevance for causal inquiry as I conceive it

is an externalist standard.

It does not follow that the hypotheses around which the inquiry is structured are themselves partially opaque. By insisting on a standard, transparent semantics for causal generalizations, evidential internalism can be preserved. But there is little reason to preserve it. The question of what evidence is relevant to the inquiry cannot but be potentially opaque and thus external, because the nature of the object of the inquiry—the target mechanism—is itself is partially unknown (and who could want it otherwise?). To opt for transparent hypotheses does not remove this opacity, but merely moves it: the opaque, externally answered question becomes that of which hypotheses are relevant to the inquiry, that is, which hypotheses make predictions about the consequences of the target mechanism only and which make predictions about the operation of other mechanisms that are outside the scope of the inquiry.

One last question. Does it matter whether a causal hypothesis is explicitly hedged or not? If opaque conditions of application are apt for scientific inquiry into causal mechanisms, why not hold that all causal hypotheses have such conditions, or to put it another way, that all causal hypotheses are implicitly hedged with a *ceteris paribus* clause? Let me conclude by endorsing this view: the sole semantics for causal generalizations is the semantics I have given for hedged hypotheses. Semantic opacity and evidential externalism permeate causal inquiry through and through.

## References

- Cartwright, N. (1989). *Nature's Capacities and Their Measurement*. Oxford University Press, Oxford.
- Earman, J. and J. T. Roberts. (1999). Ceteris paribus, there is no problem of provisos. *Synthese* 118:439–478.
- Earman, J., J. T. Roberts, and S. Smith. (2002). Ceteris paribus lost. *Erkenntnis* 57:281–301.
- Field, H. (1973). Theory change and indeterminacy of reference. *Journal of Philosophy* 70:462–481.
- Fodor, J. A. (1991). You can fool some of the people all of the time, everything else being equal: Hedged laws and psychological explanations. *Mind* 100:19–34.
- van Fraassen, B. C. (1980). *The Scientific Image*. Oxford University Press, Oxford.
- Giere, R. N. (1999). *Science without Laws*. University of Chicago Press, Chicago.
- Hausman, D. M. (1992). *The Inexact and Separate Science of Economics*. Cambridge University Press, Cambridge.
- Kincaid, H. (1990). Defending laws in the social sciences. *Philosophy of the Social Sciences* 20:56–83.
- Kitcher, P. (1989). Explanatory unification and the causal structure of the world. In P. Kitcher and W. C. Salmon (eds.), *Scientific Explanation*, volume 13 of *Minnesota Studies in the Philosophy of Science*, pp. 410–505. University of Minnesota Press, Minneapolis.

- Lewis, D. (1986). Causal explanation. In *Philosophical Papers*, volume 2, pp. 214–240. Oxford University Press, Oxford.
- Lipton, P. (1999). All else being equal. *Philosophy* 74:155–168.
- Pickering, E. C. (1912). Periods of 25 variable stars in the Small Magellanic Cloud. *Harvard College Observatory Circular* 173:1–3. Article written by Henrietta Leavitt.
- Pietroski, P. and G. Rey. (1995). When other things aren't equal: Saving ceteris paribus laws from vacuity. *British Journal for the Philosophy of Science* 46:81–110.
- Salmon, W. C. (1984). *Explanation and the Causal Structure of the World*. Princeton University Press, Princeton, NJ.
- . (1997). Causality and explanation: A reply to two critiques. *Philosophy of Science* 64:461–477.
- Strevens, M. (2009). *Depth: An Account of Scientific Explanation*. Harvard University Press, Cambridge, MA.
- Williamson, T. (2000). *Knowledge and Its Limits*. Oxford University Press, Oxford.
- Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press, Oxford.