# High-Level Exceptions Explained

Michael Strevens

*Erkenntnis* 79, 1819–1832, 2014

Abstract

Why are causal generalizations in the higher-level sciences "inexact"? That is, why do they have apparent exceptions? This paper offers one explanation: many causal generalizations cite as their antecedent—the *F* in *Fs are G*—a property that is not causally relevant to the consequent, but which is rather "entangled" with a causally relevant property. Entanglement is a relation that may exist for many reasons, and that allows of exceptions. Causal generalizations that specify entangled but causally irrelevant antecedents therefore tolerate exceptions.

## 1. Inexact Generalizations in the High-Level Sciences

It is a fact of life in the high-level sciences that even true generalizations have exceptions, or at least, apparent exceptions. *Ravens are black*—but leucistic specimens are white. *Shield volcanos erupt effusively*—but not pyroclastic shields, which erupt explosively. *Oldest children are conservative rather than creative*—but not Newton and Einstein. For this reason, the generalizations are called inexact, or non-strict, or "ceteris paribus".

What is the source of this inexactness? Is it something about the world, or something about the way we represent the world? Are the exceptions real or only apparent? Are there multiple sources of inexactness? These are some of the most important philosophical questions about the high-level or special sciences, that is, about biology, psychology and linguistics, sociology and anthropology, and economics—and for that matter, about chemistry and non-fundamental physics.

There are, I think, two broad approaches to understanding special-science inexactitude. (They do not exhaust the possibilities, I should add: between two great rivers many lesser watercourses may flow. Further, they are as you will see compatible: a generalization may be inexact for two separate reasons, each reason captured by one of the approaches.)

On the first approach, high-level generalizations contain tacit conditions of application. Were the propositional content of, for example, *Ravens are black* to be made explicit, it might read "In conditions *C*, ravens are black", where *C* is some list of requirements that is not entirely satisfied by white ravens—some list of requirements that puts leucistic, painted, and bleached ravens outside the scope of the generalization. (The hedge *ceteris paribus* is often thought to indicate the presence of these tacit conditions.)

The exceptions to a high-level generalization are not, on this view, real exceptions; they are exclusions. The leucistic raven, for example, does not represent a counterexample to, or a negative instance of, the claim that ravens are black; rather, it is excluded from the class of birds asserted to be black in virtue of the unspoken qualification "In conditions $C \ldots$".

Further, the existence of these apparent exceptions is not due to some feature of the world, some looseness in the connection between ravenhood and blackness. It is explained rather by a feature of our system for representing the world, namely, the feature that allows a generalization to make exclusions tacitly as well as explicitly.

Call this the *narrowness* explanation for inexactitude, turning as it does on an invisible condition that, by narrowing the scope of a generalization, makes it appear to be inexact to those unaware of the condition's presence in its propositional content. Different versions of the narrowness explanation—which applies roughly to the category of generalizations that Reutlinger et al. (2011) call "exclusive ceteris paribus laws"—will give different accounts of the nature of the invisible condition and the reason for its invisibility.[1] Hausman (1992, 136) and Lange (2002) derive the content of the condition from scientific practitioners' theoretical aims and working assumptions; Spohn (2002), Schurz (2002), and Nickel (2010) from various notions of normality.[2] Another strategy (exemplified by Hüttemann (this volume)) is to understand the rider as confining generalizations' claims to systems in isolation or under other "ideal" conditions.

The second broad approach to understanding inexactitude is what I call the explanation from *softness*—the softness, that is, of the connection between a generalization's antecedent and consequent. On this sort of view, *Ravens are black*, when properly understood, asserts the existence of a connection

1. On some of these accounts, the generalizations will be what Earman et al. (2002) call "non-lazy" and what Reutlinger et al. (2011) call "indefinite"; on others, they will be lazy/definite. The issues are, however, beyond the scope of the present paper.

2. For Spohn and Nickel, what counts as normal depends in part on practitioners' aims and assumptions. Schurz proposes that science should "reconstruct" generalizations framed in terms of normal conditions so as to eliminate their normality riders in favor of a stochastic connection between antecedent and consequent, or in other words, that narrowing generalizations should be converted into what I will call soft generalizations; see also Schurz (this volume).

between ravenhood and blackness that ensures that many or most ravens are black while allowing for genuine exceptions, that is, ravens that fall within the scope of the generalization but that have some other color.

A very simple version of the softness approach understands the connection to be statistical in the shallowest sense: to say *Ravens are black* is (in the relevant scientific contexts) to assert the proposition *Most ravens are black*. A Humean "best system" approach to high-level exceptions amounts, I think, to a metaphysically more sophisticated version of this implementation of the softness paradigm (Schrenk this volume; Unterhuber this volume).

A related kind of softness is, you might say, stochastic rather than merely statistical: *Ravens are black* means that there is a high physical probability that any particular raven is black. The putative exceptions to the generalization, in this latter case, are those for which the probabilistic mechanism fails to fire, the cases where Nature's dice come up "snake eyes".[3] Arguments for a stochastic approach to the majority of inexact high-level generalizations are given in this volume by Roberts and Schurz. Reutlinger is more skeptical.

Yet another variant on the softness approach reads high-level generalizations as asserting the existence of "causal tendencies" (Lipton 1999). Depending on the interpretation of the term "tendency", this might amount to any of a wide variety of softness explanations of inexactitude, each with its own distinctive conception of a causal connection between *F* and *G* that is compatible with the occasional non-*G* *F*.

On the softness approach, the reason for high-level generalizations' inexactitude is something about the world, or at least about that aspect of the world picked out by the generalization—the connection asserted to exist between the antecedent and consequent properties. More exactly, it is because the special sciences concern themselves with connections—statistical, or tendential, or stochastic—that tolerate genuine exceptions, that they are run through with

3. That is, two dice yield two '1's. The Romans called this the dog throw, though non-black dogs are hardly rarities.

inexactness.

If *Ravens are black* said what it appears to say on the surface—that all ravens are black—it would be, in virtue of leucistic and other non-black ravens, false. Both the softness and the narrowness approaches save it from falsification by supposing that it expresses a proposition that differs from its surface form: on the narrowing approach, *In conditions C, ravens are black*, and on the softening approach *Relation R holds between ravenhood and blackness* (with the precise logical form to be determined by the nature of *R*).

On another view, perhaps important enough to constitute a third great river, inexact generalizations say something—possibly something as simple as *All ravens are black*—that is literally false. A story needs then to be told about the practical value in science of falsehoods of this sort; however, I will say no more about the third river here.

In my view, both the narrowness and the softness approaches are correct. There are, that is to say, at least two wholly distinct reasons why high-level generalizations are inexact: they contain implicit hedges, creating exclusions that appear to be exceptions, and they concern among other things "soft" connections between properties, allowing real exceptions. Further, many generalizations are inexact for both reasons, which is to say, they are tacitly hedged claims about soft connections. Further still, there are at least two separate sources of softness, which may themselves cohabit in the same causal connection.

We need several complementary theories, then, to explain inexactitude in the high-level sciences: a theory of narrowing and two distinct theories of softening. My theory of narrowing—of tacit hedges—is presented in Strevens (2012a). One of the two sources of softness, causal stochasticity or indeterminism, is dealt with to some extent in Strevens (2008a) and Strevens (2011), and in work by many other writers, including the authors in this volume cited above. It is the other source of softness, which allows for inexactitude even given causal determinism, that is the topic of this paper. These various theories of inexactitude are independent; consequently, I will not assume in my presentation any particular theory of narrowness or probabilistic causation.

The account of softness offered in what follows is restricted in two ways. First, it applies only to causal generalizations, by which I mean generalizations that are intended to capture the consequences of a single causal mechanism or process—as *Ravens are black* is intended to capture the consequences of the process of natural raven development, or *Cepheid stars are more luminous, the longer their periods* is intended to capture the consequences of a certain mechanism for stellar variability. (Numerous writers on high-level laws, exceptions, and ceteris paribus hedges make similar assumptions; see for example in this volume alone Nickel, Pemberton and Cartwright, and for generalizations of the indefinite variety, Schurz.)

Second, I will for expository simplicity restrict my discussion to generalizations with the elementary form *Fs are G*. As I explain near the end of section 4, this is not the sacrifice of scientific realism that it might seem; the simple form covers many complex causal claims found in the scientific journals. In any case, the extension to generalizations with more elaborate logical forms is not terribly problematic.

## 2. Non-Causal Antecedents

Writers such as Lipton (1999), articulating a "soft" approach to inexact causal generalizations in the special sciences, have located the softness in high-level causality itself. High-level causal connections, the thought goes, are frequently "oaken" rather than "iron" (Armstrong 1983, 147), meaning that they do not always lead inexorably from antecedent to consequent.[4]

The purest kind of soft causality is irreducibly indeterministic: when *F* is causally connected to *G* in such a way, then though the conditions are

4. Armstrong's connections are not in the first instance causal; they are rather relations of necessitation that (according to his later work) when instantiated manifest themselves as relations of singular causation.

auspicious and the proper sacrifices have been made, an instantiation of *F* may not be accompanied by an instantiation of *G*. This is a familiar characteristic of processes governed by fundamental physical probabilities, of course, but you might think that there can also be a sort of non-probabilistic fundamental-level causal indeterminism, intimated by such terms as "tendency" or "propensity", due to Nature's imperfect control over the course of events (Earman 1986).

More typical is the view that the failure of a soft causal connection between high-level properties usually or always has a low-level explanation (Armstrong 1983; Lipton 1999). To see how this works, let me focus on domains in which a failure always has an explanation, thus on domains in which the relevant low-level causal processes are fully deterministic. In such worlds or realms, if there is a soft or high-level causal connection between *F* and *G*, then there is some set of initial or boundary conditions *C* such that, in all and only those cases where *C* holds, there is a hard causal connection between *F* and *G*, that is, a connection that does not tolerate exceptions (or perhaps it is better to say: there is a hard causal connection between *CF* and *G*).

It is important to distinguish the high-level and the low-level causal connections. The high-level connection (I am supposing) exists for every *F* and permits exceptions; the low-level connection exists only for *F*s in conditions *C* and does not. The high-level causal generalization *Fs are G* is true because it concerns the high-level connection: without this high-level connection to serve as subject matter you would have not a softness explanation of inexactitude but the sort of view mentioned in the previous section, according to which inexact generalizations are literally false.

Views that mix high-level indeterminism with low-level determinism face the question: why do scientists frame soft generalizations about high-level causal connections rather than focusing on the underlying hard connection, framing generalizations that attempt to capture the content of the conditions *C*? Why do scientists ever seek, in formulating their hypotheses, to "soften" when they could "narrow"?

Because the condition $C$ is hopelessly complex, says one standard answer. Strevens (2012a) undercuts the response, showing that in many circumstances it is surprisingly easy to specify the conditions. But this is not an argument I will pursue here.

My goal in what follows is rather to investigate a kind of softness that can permeate even those generalizations that aim to capture hard, or deterministic, causal processes. As remarked above, this project is quite compatible with the presumption that there is also softness due to causal indeterminism. Inexactness may have many sources, but only one is to be explored in what follows.

Assume, then, for the sake of the argument, that the high-level causal relation in question—the causal connection that is the subject matter of a high-level causal generalization—is deterministic. It follows that if $F$ has a high-level causal connection to $G$, and if all the attendant conditions required by that connection are satisfied (no interfering factors, no reversal of the outcome once the causal process is done, and so on), then every $F$ will be, and must be, $G$.

Suppose also, again for the sake of the argument, that causal generalizations specify their hypothesized causal connections' attendant conditions: they take the form *In conditions C, Fs are G*, where $C$ are the conditions for the operation (and non-reversal) of the putative underlying $G$-producing mechanism. Given determinism, how could such a causal generalization permit exceptions? How could the observation of an $F$ in conditions $C$ that is not $G$ be anything other than a falsifier?

The answer is that causal generalizations' truth conditions do not require the antecedent property $F$ to be causally connected to the consequent property $G$—indeed, that they allow $F$ to play no role whatsoever in the $G$-producing process. A generalization that takes advantage of this flexibility describes a consequence of a certain causal mechanism, then (as my definition of a causal generalization requires), but it is not a mechanism in which $F$, or any part of $F$, causally participates. In short, $F$ is allowed to be wholly causally irrelevant, in every scientifically or philosophically important sense, to $G$. Consequently, even if $F$ is present, some crucial part of the $G$-producing mechanism may be missing; in that case, you have a genuine exception to the generalization, an exception that is compatible with—for reasons I will later explain—the generalization's holding true.

In suggesting that the antecedents of some causal generalizations are irrelevant to their consequents, I am gainsaying a received view, according to which an irrelevant antecedent is always a defect (Davidson 1967; Fodor 1989). Let me present a few examples suggesting that the received view is wrong—that the antecedents of many scientifically respectable causal generalizations stand apart, causally, from the mechanisms that those generalizations describe.

Consider the philosophically august science of ornithology: there is nothing wrong with *Ravens are black*, yet on at least some theories of taxonomy, ravenhood is a causally impotent property, either because being a raven is entirely a matter of lineage, and so ravenhood is a purely historical property or—on the view that species are individuals (Hull 1978)—because ravenhood is no more than a reification of a proper name, a tag picking out an individual that points to nothing whatsoever belonging to the causal economy. The same consideration applies to any generalization that ascribes physiological or psychological properties to a species or other taxon, including many scientifically more interesting specimens: *Humans tend to think about the natural world in essentialist terms*, *The outer coat of the HIV virus is highly mutable*, and so on.

Outside the world of living things, it is true that pre-1977 Ford Pintos have a disturbingly high chance of exploding if struck from behind with sufficient force. But being a pre-1977 Ford Pinto is not what causes the explosion; rather, it is four bolts that protrude from the differential and, in the wrong sort of collision, puncture the fuel tank. Being a Pinto of a certain era goes along with this unfortunate design, but it does not cause it.

Finally, obese people have a stronger tendency to become diabetic than people of normal weight. The mechanisms are not yet fully understood, but some recent research suggests that obesity does not play a direct role in promoting diabetes; rather, a high-fat diet both brings on obesity and triggers excessive glucose production in the liver, leading eventually to the development of insulin resistance and thus type 2 diabetes (Wang et al. 2009). The connection between obesity and diabetes is typically stated in causal language—the Salk Institute press release about the discovery of the mechanism just described is titled "How obesity increases the risk for diabetes"—yet the underlying mechanism is driven not by obesity but by a diet that typically goes along with obesity, much like the Pintos' rupturing bolts.[5]

These examples do not constitute irrefutable arguments for irrelevance. In defense of a universal requirement of causal relevance you might argue, for example, that some of the properties I have mentioned are on closer examination causally relevant to their consequents at least in part. (The individualist metaphysics of species is controversial, and those raven ancestors are an important part of the evolutionary process leading to blackness.) Or you might argue that in some cases they are causally relevant in a special, high-level way: they are "relevant$_2$" (Lepore and Loewer 1987) or "quausally relevant" (Horgan 1989). Or you might declare the generalizations in which the irrelevant antecedents figure to be unfinished, arguing that the antecedents function only as placeholders for the causally relevant properties that belong in the generalizations as of right.

I will not try to resolve these debates here. That would require, at the least, a thorough examination of the nature of causal relevance, which would take me far out of my way. Let me ask skeptical readers, instead, to play along for the sake of the game. I will try to show you why we sometimes use causally

5. It is an open question whether obesity itself reinforces the mechanism by which glucose is over-produced. If it does then obesity is after all directly causal, but the important point here is that the causal generalization is assertible even if obesity is merely an effect of the real cause(s) of type 2 diabetes.

irrelevant antecedents in causal generalizations, and what connection they ought to have, if not causal, to their consequents.

Two questions about causally irrelevant antecedents, then: Why allow them, alongside their causally relevant counterparts? What rules govern their deployment?

## 3. The Role of Causally Irrelevant Antecedents

What are causal generalizations—I mean the linguistic entities or mental representations—for? They are for stating causal regularities, laws, tendencies, causal connections, and so on—for stating the facts about causal mechanisms. But that is not all. They are also for formulating, developing, and testing hypotheses about causal mechanisms, and thus for representing or attempting to represent what amounts, at best, to partial knowledge of causal mechanisms.

The first of these functions requires causal generalizations to represent causal facts and to cite causally relevant properties. The second, however, requires further representational resources. This is because there is often a certain asymmetry in our causal knowledge, when that knowledge is only partial: we know the identity of the caused property, but not the identity of the properties doing the causing.

Even scientifically naive individuals such as young children or subsistence farmers who have never heard of DNA know, for example—or certainly, they strongly suspect—that the blackness of ravens is a consequence of a causal mechanism. But they do not know, and they know that they do not know, what is doing the causing. They think that there is something about ravens that causes blackness, but they have little idea what that something is.[6]

6. On the evidence for the existence of this causal belief in naive biology, see Strevens (2000) and Gelman (2003). Some psychologists have thought that biological naïfs are essentialists who understand the property of ravenhood itself as causing the characteristic properties of ravens: their color, their behavior, and so on. Strevens provides evidence against this claim and in favor of the view stated in the main text: normal, biologically uninformed individuals believe that there is something about ravens that causes blackness, but they are

A similar state of intermediate knowledge is common in many domains: until recently, we could say that there is something about influenza viruses that can be fatal to humans; that there is something about Cepheid stars that makes their brightness vary regularly; that there is something about massive objects that causes them to attract one another; that there is something about alkali metals that causes them to react violently with water; that there is something about humans that makes them more facile at reasoning with modus ponens than with modus tollens—but that we have little idea what is doing, in each instance, the causing.

In some cases, we suspect that the antecedent property fingered by our generalization is causally relevant, as in the case of mass, and in many cases we suspect that the antecedent is not itself what is doing the causing, as in the case of being human. In the first sort of case, then, we surmise, but do not know, that the "something about mass" that does the causing is mass itself; in the second sort of case we are pretty sure that the "something about humans" is not humanity but certain psychological tendencies that humans may once not have had (if they were bad even at modus ponens until recently) and may in the future lose (once genetic therapy has cured our contrapositive clumsiness).

Either way, we need some apparatus to represent what we are relatively sure of: that for two properties $F$ and $G$, there is something about $F$s that, when conditions are right, causes $G$. Many of our causal generalizations convey precisely this information, I propose; they therefore allow (without requiring) that $F$ itself is causally irrelevant, in every significant sense, to $G$. When we say that ravens are black, then, we are saying that something about ravens causes blackness without saying also that the something is ravenhood itself.

In that case, what are we saying? First, we are postulating the existence of an underlying property $P$ that causes blackness in ravens (again, when

---

not committed to the cause's being the essence of ravenhood.

conditions are right—but I will from now on omit this aspect of the claim when convenient).

Second, we are postulating some connection between ravenhood and the property *P*. What sort of connection? It must be a relation broad enough to be realized, at one extreme, by the identity mapping (to allow the possibility that it is the antecedent property itself that does the causing), and at the other extreme, by something consistent with the total causal irrelevance of ravenhood to blackness, while at the same time in every case building a bridge between ravenhood and *P* that is strong enough to bear the inferential and explanatory traffic that we expect our causal generalizations to sustain.

## 4. Entanglement and Non-Causal Relevance

To qualify as an antecedent in a causal generalization, a property must stand in what I will call a relation of *entanglement* to a cause of the consequent (Strevens 2008a, 2012b). Causally irrelevant properties qualify as antecedents, then, in virtue of their being entangled with causally relevant properties. To claim that ravens are black is therefore to claim for some, typically unknown property *P*, that ravenhood is entangled with *P* and that *P* causes blackness. More generally, to claim that *F*s are *G* (in a causal tone of voice—as opposed to, say, a statistical tone of voice) is to claim that *F*-ness is entangled with some perhaps unknown *P* and that *P* causes *G*. (On the way that the unknown *P* enters into these claims, see Strevens (2012a). For the purposes of this paper, you can think of the claims as existentially quantified: to make an *F*/*G* causal claim is to assert that there exists some *P*, such that *F* is entangled with *P* and *P* causes *G*—though I do not in fact think that this analysis is quite correct.)

What is entanglement? A property *F* is entangled with a property *P* just in case there is a robust tendency for *F* to be accompanied by *P*, there is a single explanation for this tendency, and *P* is an intrinsic property of *F*s.

I will break this formulation down into four conditions: two for the robust connection between *F* and *P*, one for the "single explanation" criterion, and a fourth for the intrinsicness condition. Only the first of the four conditions plays a role in explaining causal generalizations' softness. The other three conditions are described here to motivate the claim that entanglement qualifies a causally irrelevant property to play the antecedent role in a causal generalization, on the grounds that entanglement captures what we have in mind when we say "*Something about F*s causes *G*".

The first condition: for a property *F* to be entangled with a property *P*, most *F*s must have *P*. As a consequence, the causal generalization *Fs are G* implies that most *F*s for which *C* holds are *G*, where *C* is the set of conditions required for *P* to cause *G* and assuming that there is no correlation between the entanglement's falling through and *C*'s failing to hold. (That is, of course, compatible with most *F*s, or even the overwhelming majority of *F*s, failing to have *G*: to cite the famous example, *Sperm fertilize eggs*.)

The second condition exists to ensure that causal generalizations with entangled antecedents offer counterfactual support. A merely statistical connection between *F* and *P* is not sufficient, the second condition says, for entanglement: there should be a subjunctive tendency for *F*s to have *P*, or in other words, *F*-ness should be robustly connected to *P*-hood. What this amounts to is the truth of a wide range of counterfactual conditionals. In the case of the entanglement of ravenhood and the coloration mechanism *P*, for example, many conditionals such as the following should hold true:

> If this particular raven had had a slightly different diet, it would still have had *P*, and
>
> If these two particular ravens had mated (in fact they did not), their offspring would have had *P*.

More schematically: for *F* to be entangled with *P*, it should be the case not only that most actual *F*s have *P*, but that (a) most of the *F*s would still have had *P* in many circumstances where things went slightly differently, and (b) if things, going slightly differently, had brought new *F*s into existence, most of these *F*s would also have had *P* (at least for most such slight differences).[7]

There is of course a fair amount of latitude in this prescription; the robust connection it characterizes is a matter of degree (and so you might suppose—though I will not develop the thought here—that entanglement itself is a matter of degree). A complete theory of causal generalizations would have more to say: it would perhaps specify a minimum level of robustness, a way of quantifying "many" and "most", and it would further characterize the relevant counterfactual antecedents. (What is meant by things going "slightly differently"?) But none of this helps to explain how causal generalizations with entangled but non-causal antecedents tolerate exceptions, so I will stop here.

Concerning the third condition for entanglement, I will be equally concise. In virtue of the first two conditions, *F* is entangled with *P* only if most *F*s have *P* in actuality and (roughly) many *F*s have *P* subjunctively. The third condition requires that there is a single reason, or explanation, that accounts for the instantiation of *P* by (almost) every such actual and subjunctive *F*. The third condition also requires that any actual or subjunctively nearby thing that has *P* for this reason is an *F*, so that *F* gathers together in its extension all the things that have *P*, or if conditions were slightly different would have *P*, for the given reason. As some examples of entanglement at the end of this section will suggest, the shared explanation need not be causal.

Ravenhood satisfies the third condition, then, because (almost?) all ravens have the raven coloration mechanism *P* for the same reason—something to do with raven inheritance and development— and because the things that have the raven coloration mechanism for this reason are (almost?) all ravens.

What motivates the third condition? It is the principle, rooted in the notion that causal generalizations are devices for formulating and testing

7. On the grounding of these and other such counterfactuals in actual, scientifically significant properties of ravens and their environment, and their connection to the counterfactual support offered by the causal generalization as a whole, see Strevens (2008b).

hypotheses about explanatory structure, that causal generalizations should go hand in hand, and therefore one by one, with explanations: a generalization connecting properties *F* and *G* should have as its instances all and only those *F*/*G* pairs in which *F*'s being accompanied by *G* has a single shared explanation. The principle motivates also the defining characteristic of causal generalizations, that they describe the consequences of a single causal mechanism. (There is no prohibition, of course, on generalizations that individuate explanatory reality differently, or more coarsely. But they are not *causal* generalizations; they are not the usual and default generalizations deployed in the investigation of causal-explanatory structure.)

Finally, the fourth condition: the entangled property *P* must be an intrinsic property of the *F*s, as the raven coloration mechanism is an intrinsic property of ravens. Note that the requirement concerns not what's intrinsic to *F*-ness but what's intrinsic to the things that *F*-ness is predicated of—ravens, not ravenhood. Without the satisfaction of the intrinsicness condition, it is not correct, I think, to describe *P* as "something about" *F*s.

* * *

As examples of entanglement, consider two causal generalizations from the beginning of this paper: *Shield volcanos erupt effusively*, and Sulloway's (1996) hypothesis *Oldest children are conservative rather than creative*.

A volcano's being a shield volcano is not a cause of its erupting effusively rather than explosively: certain structural properties of the volcano cause both effusive eruptions and (because the lava is very liquid) its characteristic broad, shield-like morphology. The property of being a shield volcano earns its place as an antecedent because it is entangled with such structure: there is a particular reason why shield volcanoes and only shield volcanoes tend robustly to have this structure. The entanglement in this case exists because of a causal mechanism, but the antecedent property is not itself doing the causing—it and the consequent are, rather, effects of a common cause. (That *F* is caused by *P* is not sufficient, note, for *F*'s being entangled with *P*; it must in addition be the case that *P* is the usual cause of *F*. In common cause generalizations, then, entanglement does not exist for causal reasons alone.)

That a child has younger but no older siblings is likewise not a cause (according to Sulloway's thesis) of conservative tendencies. The tendencies are rather caused by something that goes along with being an oldest child, namely, the psychological concomitants of the child's having a preeminent position in the sibling power structure. Why this "going along"? Because older siblings tend to be stronger and more experienced than younger siblings. Arguably, this tendency has a causal explanation (certainly, the explanation is partly causal), but whether the property of "being older" itself plays a causal role in that explanation is doubtful. Nevertheless, there is clearly an entanglement: the tendency, whatever its roots, is counterfactually tenacious and exists for a single reason in (almost) every case.

Entanglement may exist for many other reasons as well. Because water is constituted by $H_2O$ molecules, the property of being water is entangled with the property of being largely $H_2O$. (I leave open the question whether water is identical to $H_2O$, though I suspect a negative answer.) A desire for chocolate may be entangled in humans with a certain kind of neural activity, though the externalist element of the desire's content means that it cannot be (wholly) constituted by such an activity. Finally, the moral goodness of an act may be entangled with properties of the act that motivate humans to perform it—though the motivation is necessarily purely causal and goodness may be a non-natural property with no causal power. Consequently, non-natural moral properties can appear in the antecedents of causal generalizations. I describe these possible cases of entanglement to entice you rather than to persuade you; they of course require far more discussion than I can give them here.

* * *

I am proposing that entanglement finds its way into the semantics of simple causal generalizations as follows: *In conditions C, Fs are G* means something like

> By way of such and such a mechanism, certain properties entangled with *F* bring about, when conditions *C* hold, *G*.

Remembering that both *F* itself and its constituent parts are entangled with *F*, such a generalization may attain truth in a number of different ways. The *G*-producing mechanism may comprise *F*, some part of *F*, something distinct from but entangled with *F*, or a mixture of these—for example, a constituent of *F* along with an entangled property. This semantic liberality allows causal generalizations to perform their function—to provide investigators with a vehicle for stating the putative consequences of a mechanism—even when the investigators know little or nothing about the workings of the mechanism.

For simplicity's sake I have restricted this discussion to causal generalizations having the form *Fs are G*. It is clear, I hope, that the usefulness of the notion of entanglement is not confined to these cases. Consider, for example, a functional generalization linking a person's body mass index, hence their degree of obesity, to their probability of developing type 2 diabetes. As explained in section 2, it is quite possible that obesity is not a direct cause of diabetes but is rather entangled with such a cause, namely, a high-fat diet. The story is easily made quantitative: different levels of obesity are entangled with different levels of fat intake, which then probabilify diabetes to different degrees.

It is worth noting also that even the elementary *Fs are G* template covers a wider range of causal claims in science than you might suppose. Many logically or mathematically complex causal generalizations have the form *Systems of type S exhibit behavior B*. Here *S* might describe, for example, a certain kind of economic system—a feudal economy or a type of market—and *B* might pick out a complex pattern of behavior specified by a mathematical model or a computer simulation. (The model might in turn consist of structural equations that, like the obesity/diabetes generalization, have causal interpretations of their own.) If my interpretation of causal generalizations is on the right track, then for the *S*/*B* generalization to hold true, it need not be the case that *S* picks out the causal structure responsible for the behavior *B*; it is sufficient that *S* be entangled with such a structure, in precisely the sense described above.

Some concluding remarks on the logic, in a broad sense, of entanglement. First, it is as promised a wide-ranging relation: it encompasses identity—a property is always entangled with itself—and many far weaker connections. One property can be entangled with another for causal reasons (e.g., they are effects of a common cause), for metaphysical reasons (e.g., the instantiation of the other is constitutive of the instantiation of the one), or for any other reason that counterfactual dependence might exist.

Second, entanglement is asymmetric: that *F* is entangled with *P* does not imply that *P* is entangled with *F*.

Third, entanglement does not imply explanatory relevance to the entangled property. If *F* is entangled with *P* then there exists an explanation of *F*'s being robustly connected to *P*, but *F* may not play an active role in this explanation—as when *F* is an effect of *P*, or *F* and *P* are effects of a common cause.

Fourth, entanglement *does* imply explanatory relevance, though not causal relevance, to anything caused by the entangled property: if *F* is entangled with *P* and *P* causes *G*, then *F* plays a role in explaining *G* (or so I argue in Strevens (2008a)). The entangling property *F* in effect hitches an explanatory ride on the entangled property *P*.

Fifth, entanglement tolerates exceptions. Which brings us back to inexactitude.

## 5. Entanglement and Inexactitude

Though one property is entangled with another, the one is not always followed by the other. The entanglement of *F* with *P* requires a tendency for *F*s to have *P*, but that tendency need not be, and often is not, exceptionless. Thus though there exists an entanglement between ravenhood and the normal raven coloration mechanism, not all ravens possess the normal mechanism; the albinos and leucistic ravens, in particular, do not. Likewise, some oldest children do not hold a position of power and authority over their younger siblings and some shield volcanoes lack the structure that brings about effusive eruptions—though in both cases, the one usually goes along, for robust reasons, with the other.

Consider a causal generalization *Fs are G*, where *F* is not itself causally relevant to *G*, but is merely entangled with the *G*-producing mechanism *P*. If the tendency underlying the entanglement is not exceptionless, then some *F*s will lack *P* and so will lack *G*.[8] These are exceptions to the generalization, exceptions that can exist even if the causal mechanism in question, the mechanism by way of which *P* produces *G*, cannot fail—because they are cases in which some crucial part of the infallible mechanism is simply not present.

Such *exceptions* must be distinguished from what I called earlier in this paper *exclusions*. Let me conclude by putting the two side by side, adding for good measure exceptions due to the other source of softness, causal indeterminism.

Every well-formed and well-founded causal generalization *Fs are G* is connected to a putative mechanism *P* having conditions of operation *C*. (The determination of the mechanism and conditions of operation is discussed in Strevens (2012a), but not here.) When *F* is not itself the mechanism *P*, such a generalization asserts the existence of two relations: first, a relation of entanglement between *F* and *P*, and second, a causal relation running from *P* and *C* to *G*. Together these imply that *F*s in conditions *C* have a tendency to *G*-ness.[9]

Now take some *F* that is not *G*. What happened? There are three possible

8. It does not logically follow that they lack *G*: they may have *G* for some other reason, like a leucistic raven dyed black. Such individuals conform to the generalization but are not instances of the generalization, because their *G*-ness is not explained in the right way.

9. Except in a perverse scenario in which *C* holds only in situations where the entanglement falls through.

answers:

1. Despite the entanglement, the *F* in question lacks some part of *P*. Such *F*s—leucistic ravens, for example—are *exceptions* to the generalization.

2. For this particular *F*, conditions *C* did not hold. Since the generalization makes claims only about cases where *C* holds, such *F*s—bleached ravens, for example—are (tacit) *exclusions*: they are cases that, on closer examination, fall outside the generalization's scope.

3. All prerequisites for the operation of the causal mechanism—both *P* and *C*—were present, but the mechanism "misfired". This is possible only if the causal relation between *P* and *G* is in some way indeterministic, perhaps stochastic. Such exceptions merit a name of their own; I call them *negative instances*.[10]

In the first case, the absence of *G* is allowed by the kind of softness described in this paper. In the second case, it is allowed by narrowing. In the third case, it is allowed by softness due to indeterminism, at either the high or the fundamental level. These exhaust, I think, the varieties of inexactitude to be found in the causal generalizations of the high-level sciences.

**Acknowledgments**

Thank you to Laura Franklin-Hall, the editors, and two anonymous referees.

10. I further suggest that the term "exception" should be applied, in a stochastic context, not to cases where an *F* lacks *G* but to cases where, because of local circumstances, an *F*'s probability of being *G* departs from the probability asserted by the generalization—for example, where the probability is small though the generalization asserts that it is large. Using this terminology, negative instances and exceptions both arise from softness, but softness of different sorts.

## References


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

Davidson, D. (1967). Causal relations. *Journal of Philosophy* 64:691–703.

Earman, J. (1986). *A Primer on Determinism*. D. Reidel, Dordrecht.

Earman, J., J. T. Roberts, and S. Smith. (2002). Ceteris paribus lost. *Erkenntnis* 57:281–301.

Fodor, J. A. (1989). Making mind matter more. *Philosophical Topics* 17:59–79.

Gelman, S. A. (2003). *The Essential Child: Origins of Essentialism in Everyday Thought*. Oxford University Press, Oxford.

Hausman, D. M. (1992). *The Inexact and Separate Science of Economics*. Cambridge University Press, Cambridge.

Horgan, T. (1989). Mental quausation. *Philosophical Perspectives* 3:47–76.

Hull, D. (1978). A matter of individuality. *Philosophy of Science* 45:335–360.

Hüttemann, A. (this volume). Ceteris paribus laws in physics. *Erkenntnis*.

Lange, M. (2002). Who's afraid of ceteris paribus laws? or: How I learned to stop worrying and love them. *Erkenntnis* 57:407–423.

Lepore, E. and B. Loewer. (1987). Mind matters. *Journal of Philosophy* 84:630–642.

Lipton, P. (1999). All else being equal. *Philosophy* 74:155–168.

Nickel, B. (2010). Ceteris paribus laws: Generics and natural kinds. *Philosophers' Imprint* 10(6):1–25. URL = <http://quod.lib.umich.edu/p/phimp/3521354.0010.006>.

———. (this volume). The role of kinds in the semantics of ceteris paribus laws. *Erkenntnis*.

Pemberton, J. and N. Cartwright. (this volume). Ceteris paribus laws need machines to generate them. *Erkenntnis*.

Reutlinger, A. (this volume). Do statistical laws solve the problem of provisos? *Erkenntnis*.

Reutlinger, A., G. Schurz, and A. Hüttemann. (2011). Ceteris paribus laws. In E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. Spring 2011 edition. Metaphysics Research Lab, CSLI, Stanford, CA. URL = <http://plato.stanford.edu/archives/spr2011/entries/ceteris-paribus/>.

Roberts, J. T. (this volume). Ceteris paribus law statements as vague, self-referential, self-locating, statistical, and perfectly in order. *Erkenntnis*.

Schrenk, M. (this volume). Better best systems and the issue of ceteris paribus laws. *Erkenntnis*.

Schurz, G. (2002). Ceteris paribus laws: Classification and deconstruction. *Erkenntnis* 57:351–372.

———. (this volume). Ceteris paribus and ceteris rectis laws: Content and causal role. *Erkenntnis*.

Spohn, W. (2002). Laws, ceteris paribus conditions, and the dynamics of belief. *Erkenntnis* 57:373–394.

Strevens, M. (2000). The essentialist aspect of naive theories. *Cognition* 74:149–175.

———. (2008a). *Depth: An Account of Scientific Explanation*. Harvard University Press, Cambridge, MA.

———. (2008b). Physically contingent laws and counterfactual support. *Philosopher's Imprint* 8(8):1–20. URL = <http://quod.lib.umich.edu/p/phimp/3521354.0008.008>.

———. (2011). Probability out of determinism. In C. Beisbart and S. Hartmann (eds.), *Probabilities In Physics*. Oxford University Press, Oxford.

———. (2012a). Ceteris paribus hedges: Causal voodoo that works. *Journal of Philosophy* 109:652–675.

———. (2012b). The explanatory role of irreducible properties. *Noûs* 46:754–780.

Sulloway, F. J. (1996). *Born to Rebel: Birth Order, Family Dynamics, and Creative Lives*. Pantheon, New York.

Unterhuber, M. (this volume). Do ceteris paribus laws exist? A regularity-based best system analysis. *Erkenntnis*.

Wang, Y., L. Vera, W. H. Fischer, and M. Montminy. (2009). The CREB coactivator CRTC2 links hepatic ER stress and fasting gluconeogenesis. *Nature* 460:534–537.