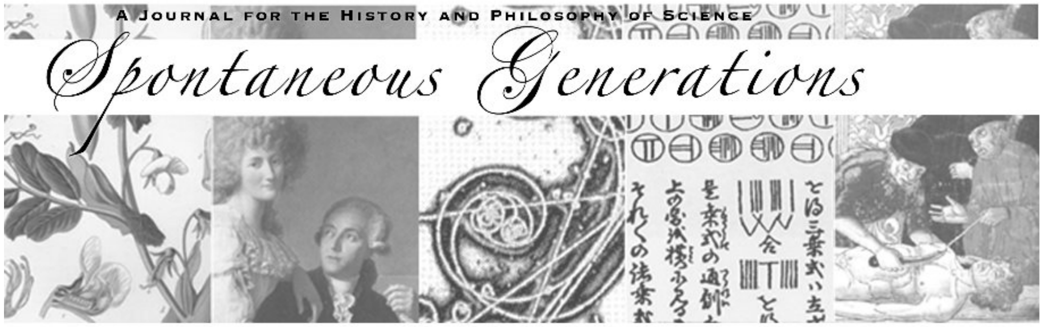

REVIEW: Michael Strevens, Depth: An Account of Scientific Explanation

Author(s): Anthony J. Kulic

Source: *Spontaneous Generations: A Journal for the History and Philosophy of Science*, Vol. 4, No. 1 (2010) 292-299.

Published by: The University of Toronto

DOI: 10.4245/sponge.v4i1.14195

EDITORIAL OFFICES

Institute for the History and Philosophy of Science and Technology
Room 316 Victoria College, 91 Charles Street West
Toronto, Ontario, Canada M5S 1K7
hapsat.society@utoronto.ca

Published online at jps.library.utoronto.ca/index.php/SpontaneousGenerations
ISSN 1913 0465

REVIEWS

Michael Strevens. *Depth: An Account of Scientific Explanation*. xvii + 516 pp. Cambridge: Harvard University Press, 2009.*

Anthony J. Kulic†

Michael Strevens' *Depth: An Account of Scientific Explanation* is an impressive recent contribution to the philosophical literature on explanation. While clearly influenced by several of the leading theories of the later twentieth century, Strevens' account is firmly rooted in the causal tradition, with his most notable intellectual debts in this regard owing to David Lewis and Wesley Salmon. Still, Strevens sees the work of these theorists as flawed in important respects and his "kairetic account" of explanation is meant to provide answers to problems his predecessors left unresolved (or poorly resolved, as the case may be). Before examining Strevens' account we should identify the more significant of these problems and briefly survey the contexts in which they arose.

Rigorous work on explanation began in the mid-twentieth century with the Deductive-Nomological (DN) model of explanation. Under the DN model, an *explanandum*–a sentence describing the intended object of explanation–is a consequence of an *explanans*–a set of sentences adduced to account for the object of explanation. Furthermore, an explanans must include at least one statement of scientific law. Hence, explanation is both deductive *and* nomological. The DN model remained the leading account of explanation until critics noted that certain prima facie acceptable explanations exhibit asymmetric features to which the DN model is insensitive. For example, consider that using *i*) a particular flagpole's height, *ii*) its angle relative to the sun, and *iii*) laws that describe the rectilinear propagation of light, one can derive *iv*) the length of the flagpole's shadow. According to the DN model, then, *iv* is explained in virtue of a derivation from *i*, *ii*, and *iii*. However, a derivation conforming to the DN criterion for explanation obtains in the reverse direction, where *iv*, *iii*, and *ii*, may be used to derive *i* Yet it is highly counterintuitive to claim that the reverse derivation *explains* the flagpole's height. Counterexamples like this one expose at least two flaws of the DN model: 1) it is insensitive

* Received 22 July 2010.

† Anthony is a doctoral candidate at the Institute for the History and Philosophy of Science and Technology at the University of Toronto, where he studies philosophy of science in connection with metaphysics.

to the asymmetrical nature of certain explanations; and, 2) given 1, the DN model fails to suitably specify *sufficient conditions* for explanation. It should be noted, however, that the former might be resolved via some satisfactory answer to the latter.

Early attempts at addressing the problem of giving sufficient conditions for explanation prompted theorists to advance analyses based on statistical relevance (SR), where sufficient conditions for explanation are identified with the "difference-making" features that give rise to the event of some target explanandum. For example, under Salmon's (1971) SR model of explanation, statistical relevance is stated as follows: given some class or population *C*, a property *P2* is statistically relevant to another property *P1* if and only if the probability of *P1* conditional on *P2* and *C* is different from the probability of *P1* conditional on *C* alone.

While undeniably capturing a vital difference-making component of explanation, the SR model is problematic on various fronts. One of these concerns certain counterintuitive results with respect to low probability events. Consider a quantum mechanical explanation for the following event: a wine glass is knocked from a high countertop to a tile floor and shatters. Given the nature of quantum mechanical explanation, the probability distribution for the 'dropped-glass event' (i.e. the difference-maker) functions as the explanans for the shattering event (i.e. the outcome). This seems correct because the theory predicts the glass will shatter with high probability. But, recall, the SR model doesn't impose a high probability requirement, only a difference-making requirement. Consequently, the probability distribution for the dropped-glass event counts as an explanation for each and every outcome in the partition of possible outcomes. The problem becomes acute when considering the unlikely outcomes in the partition. For example, in the extremely unlikely event the glass passed through the floor and emerged in the room below unperturbed–a phenomenon in quantum theory known as quantum tunneling–the same probability distribution that "explained" the shattering event would likewise explain the tunneling event. Hence, under the SR model, an explanans is capable of supporting both an explanandum *and its negation*.

Of the many problems that arise for the SR model of explanation, perhaps the most serious is its wholesale omission of information pertaining to causal relationships that seem obviously relevant to scientific explanation. Recognizing this as a serious weakness, Salmon offered a reworked account he called the Causal Mechanical (CM) model of explanation. Under the CM model, to explain some event *E* is to trace the causal processes and interactions that give rise to *E*. But while Salmon's CM model seems to ably handle the problem of explanatory asymmetries (since CM explanations are strictly unidirectional, from causes to effects), it falls short of answering the problem of explanatory relevance: i.e., which causal processes and events ought to count as relevant to a target explanandum? Consider, e.g., a collision of curling stones. What licenses appeal to the moving stone's linear momentum to the exclusion of other causal variables as crucial for explaining the stationary stone's subsequent movement? Regrettably, Salmon's CM model doesn't include a method for distinguishing between relevant and irrelevant causes.

In *Depth*, Strevens purports to offer an account of explanation that, among other solutions, resolves the problem of explanatory relevance. At 516 pages the book is a goliath, but its bulk reflects the rigour Strevens applies to its subject matter. *Depth* consists of five parts, each of which focuses on some aspect of Strevens' 'kairetic account' of explanation. His views related to causality, laws and probabilities are found in parts I, III and IV, respectively; part II contains his statement of the kairetic account of explanation for deterministic events; and part V is a comparably short summation of the main points raised in parts I–IV. Additionally, Strevens demonstrates how the elements discussed in earlier chapters can work to solve explanatory problems that arise in the special sciences.

Strevens organizes the book by examining the various aspects of scientific understanding with an eye to their explanatory structure, as opposed to their domain of investigation. Hence, explanations from quantum mechanics, statistical mechanics, evolutionary genetics, medicine and sociology are treated together given their shared probabilistic structure; likewise, deterministic explanations are treated together. Since constraints on space preclude a thorough discussion of each of the book's five parts, I will instead give a concise overview of Strevens' kairetic account. By way of conclusion I offer some commentary and criticism.

So, why the *kairetic* account? Oddly enough, Strevens doesn't provide an etymology for what is evidently a neologism. However, "kairos" is one of two terms the ancient Greeks used in reference to the phenomenon of time. The other, "chronos", translates as the standard, quantitative sense of time. "Kairos," on the other hand, is a *qualitative* notion that translates as "an opportune time" or "situational context." Since temporal considerations aren't central to the kairetic account, the name seems apt if Strevens dubbed it with the latter sense in mind (more on this later).

Strevens' account of explanation is "ontological" insofar as he takes explanations to involve the scientific discovery of sets of *explanatory facts* about the world, where facts are understood as true propositions, and where laws and regularities are invoked to derive explanations from facts (except where laws/regularities are themselves the objects of explanation).

Strevens places himself squarely in the causal tradition by supposing that the only relation germane to explanation is causality: "[w]hat explains a given phenomenon is a set of causal facts... The communicative acts that we call explanations are attempts to convey some part of this explanatory causal information (but not just any part...)" (p. 6). Here, the parenthetical claim alludes to the problem of explanatory relevance–only the causes relevant to some explanandum matter. So, e.g., seismologists should ignore facts about the gravitational influence of distant galaxies upon Earth when they explain why an earthquake occurred. But what principle determines which causal information we should include and which to exclude?

Strevens outlines the procedure for constructing a kairetic explanation as follows:

> An attempt at a deterministic kairetic explanation of an event *e* begins with a veridical, deterministic, atomic causal model for *e*. That the model is deterministic and causal means that (a) the setup of the model entails the target *e*, and (b) this entailment mirrors a real-world relation of causal production. Such a model is subjected to the optimizing procedure; the result is an explanatory kernel, a model containing only factors relevant to the explananadum. A standalone explanation is built from explanatory kernels. (p. 358)

Although the influence of the DN model is evident here, the principal difference between the two concerns their respective conceptions of entailment. While DN explanation is taken as a purely logical entailment from premises (explanans) to a conclusion (explanandum), Strevens conceives the entailment relation as wholly causal. His notion of causal entailment assumes a primitive notion of *causal influence*, which he accepts as unproblematic. Causal influence is simply the relation that obtains between any cause and its effects, with the former bearing a unidirectional causal influence upon the latter. The entailment thus derives from the causes in question supporting the right counterfactual conditionals.

According to Strevens, there are broadly two categories of thing that will feature in an explanandum: events (i.e., happenings, occurrences) and laws (i.e. generalizations.) Strevens' metaphysics of events is coarse-grained, allowing for *high-level events* as objects of explanation. High-level events are abstracted from the specific properties of their constituent low-level physical entities and events, giving rise to *states of affairs*. An earthquake, recall, is an example of the sort of high-level event/state of affairs that might feature as a target of explanation.

When investigating earthquakes seismologists needn't worry about the behaviour of the particles composing the rock involved in the seismic activity resulting in the earthquake. Rather, their concerns extend only to details relevant to an explanation for why the quake occurred; together these details constitute an *explanatory kernel*. To get at these details the scientist chooses a "veridical, deterministic, atomic causal model" (i.e. a theory, or component model thereof) that posits entities like, e.g., tectonic plates that exert a causal influence upon one another. One should note, however, that the entities appealed to in such explanations are not themselves objects of explanation, but rather objects of scientific *description*. Using the tectonic model, then, the seismologist formulates an explanation by showing how movements between specific plates causally entail the quake.

Strevens also applies his kairetic procedure to give an explanatory account of laws and regularities. When an event *E* is taken to represent an instance of some law or regularity *L*, *L* is explained by citing the explanatory kernel of the particular event *E* (as well as other events like *E* taken to exemplify the same law or regularity). However, in order to achieve the generality required of a law or regularity one needs to eliminate the initial conditions particular to *E*. Doing so allows one to derive the *causal mechanism* that underlies both the regularity and any instances thereof.

Recall that in order to derive an explanatory kernel the chosen model is subject to an *optimizing procedure*. This is notable because the optimizing procedure dictates sufficient conditions for explaining some target. Moreover, the optimizing procedure makes sense of Strevens' decision to name his account "kairetic," since optimization is applied to a specific "situational context." The optimizing procedure is a kind of filtration process; it eliminates genuine causal influences that, while influential to some degree, ultimately make no difference to our grasping the explanation for some event. To borrow a recurring example from *Depth*, Mars' gravitational influence on Earth, while substantial, is irrelevant to an explanation of Rasputin's death. On the other hand, facts about whether he was poisoned, shot, bound and gagged, and thrown into an icy river certainly bear mentioning. For Strevens, prior to optimization (and abstraction) an hypothetically comprehensive explanans would make reference to each and every causal influence on the event referred to in the explanandum. Optimization thus involves scientists determining the minimal set of facts sufficient to *causally entail* the target event of the explanandum. Consequently, any relevant fact missing from an explanatory kernel will fail to produce a causal entailment, and the kernel thereby fails as a *standalone explanation*.

The most challenging but theoretically rich section of *Depth* is Part

IV, where Strevens considers cases in which it seems possible to give useful probabilistic explanations. According to Strevens, any successful standalone explanation requires that the model featuring in the explanans causally entail the event of the explanandum. Hence, "the aim of a probabilistic explanation is to isolate the causal difference-makers, where the causal difference making relation is the same high-level causal relation as in deterministic explanation..." (p. 336). This, of course, requires considerable unpacking, which Strevens does by providing extended answers to a series of questions about the formal criteria for probabilistic explanation. They are as follows: "Can a low probability explain an event? Do higher probabilities explain events better than lower probabilities? Can a factor that decreased the probability of an event explain that event?" (p. 336).

There is insufficient space in this review to give even a cursory pass over the sophisticated answers Strevens provides to these questions. But it is worth raising a more general conclusion he supports.

Given certain phenomena, scientists will explain/predict outcomes probabilistically, even when the phenomena in question are assumed to be deterministic. Classical statistical mechanics and population genetics are paradigm examples of theories that impose probabilities onto systems assumed to be governed by deterministic processes (as opposed to, say, quantum mechanics, whose domain of application is often held to be *irreducibly* probabilistic). Since we seem precluded from tracking all the deterministic causal relations relevant to such systems, probabilities are introduced for purposes of theoretical and epistemological economy. On this picture, we formulate predictively useful probabilistic explanations that capture the difference-making causal relations we're concerned about, but we do so at the expense of a maximally precise deterministic account. Strevens thinks this is false: our foisting probabilities onto deterministic systems comes at no cost whatsoever. Even if exhaustive deterministic accounts were computationally tractable, he argues, the probabilistic accounts are on balance superior. Quite convincingly, Strevens shows that probabilities *must* be invoked to explain certain important characteristics of high-level deterministic systems–characteristics that even a fully deterministic account would neglect to capture.

This is just one interesting conclusion from a section of *Depth* that can't be given adequate treatment in a review of this length. Suffice it to say, Strevens' treatment of the problems arising from statistical approaches to explanation are superlative. Part IV of *Depth* is essential reading for anyone interested in probabilistic explanation.

Although there are potential problems with some aspects of Strevens' positive account, I want to draw attention to a substantial problem of omission that, on my view, limits the theory's otherwise broad applicability. In particular, Strevens ignores a kind of scientific explanation that is neither causal nor irreducibly probabilistic. In fact, Strevens fails to recognize this form of explanation in the book's preface, where he acknowledges other omissions (p. xv). Since this form of explanation isn't causal, it threatens the view (one that I'm not certain he fully endorses) that all scientific explanation must be causally derived.

The kind of explanation to which I refer derives from theoretical physics; it is a kind of "geometrical explanation" (Nerlich, 1979). On this account, certain phenomena are best explained by appealing to the geometric structure of spacetime as specified in General Relativity. Consider, e.g., the trajectory of a free particle moving in spacetime. How does one give a satisfactory "why-explanation" of the particle's specific trajectory? Since the free particle travels independent of any force (i.e. it is "field-free"), and since it isn't subject to the exchange of some conserved quantity, it isn't subject to a causal influence. In such cases physicists advert to the geometric structure of spacetime to explain its behaviour. Specifically, they show that the particle's movement is explained by virtue of its following a path along a geodesic.

I have offered a summary of Strevens' excellent book and advanced a criticism of omission rather than an objection to other potentially problematic aspects of his positive program. The counterexample I present constitutes a problem for Strevens to the extent that he assumes his theory capable of handling all forms of scientific explanation. Clearly, in its present form it does not. Nevertheless, this lapse should not deter potential readers from engaging this elegantly written, philosophically rich and innovative account of scientific explanation. To claim that Strevens' work is first-rate is, quite simply, to understate the case.

ANTHONY J. KULIC
IHPST, University of Toronto
91 Charles St. W.
Toronto, ON
Canada M5S 1K7
anthony.kulic@utoronto.ca

REFERENCES

Nerlich, Graham. 1979. What Can Geometry Explain? *British Journal for Philosophy of Science* 30: 69-83.

Salmon, Wesley. 1971. Statistical Explanation. In *Statistical Explanation and Statistical Relevance*, ed. Wesley Salmon, 29-87. Pittsburgh: Pittsburgh University Press.