Harvard Review of Philosophy 14 (2006): 4-19.



METAPHYSICS

# Theories and Models in Metaphysics

*By Peter Godfrey-Smith*

## §1. Introduction

METAPHYSICS IS ONCE AGAIN A THRIVING SUBDISCIPLINE WITHIN PHILOSOPHY, DESPITE A long tradition of challenges to the very viability of the metaphysical enterprise. The criticisms have not so much been satisfactorily answered as shouldered aside by the vigorous development of the field. Some focused meta-theoretic discussion has recently arisen within mainstream metaphysics.[1] The present paper is written more from an outsider's vantage point. I attempt to give a new meta-theory for some parts of metaphysics. The central claim is that much metaphysical work, especially of the contemporary systematic kind, might best be understood as *model-building*, in a specific sense of this term that draws on recent philosophy of science.

Such a claim faces meta-theoretic problems of its own. If metaphysicians are engaged in model-building, surely they know this already, or at least can be easily induced to recognize it once the framework and terminology are introduced. But the account I offer is quite far from the usual self-conception seen in contemporary metaphysics.[2] Indeed, most metaphysicians to whom this view has been presented so far have actively resisted the analysis. So what status is the analysis supposed to have? Is it a "rational reconstruction" of metaphysical work, or something more like a psychological hypothesis?

I see the analysis developed here as having several possible roles. First, metaphysicians might be trying to do something impossible, but succeeding in doing useful model-building despite themselves. In that sense, I offer something like a rational reconstruction. Those who find the usual arguments against the viability of metaphysical inquiry convincing might see the analysis as a way to salvage some useful content from metaphysical work. But there is also a psychological interpretation of the analysis that is more adventurous. Here we see metaphysicians as actually engaged in a form of modeling, despite their ideology, in virtue of the constructs they produce

*Peter Godfrey-Smith is professor of philosophy at Harvard University. He works primarily in the philosophy of biology and the philosophy of mind. He also has a more general interest in the philosophy of science, and makes excursions into metaphysics and epistemology. He has published two books so far:* Complexity and the Function of Mind in Nature *(1996) and* Theory and Reality *(2003). He is currently working on one book about natural selection and another about John Dewey.*

and the psychological mechanisms employed. It would be foolish to suppose that such a hypothesis could be applied to all metaphysical discussion, but it might be true of an important part of the field. Though metaphysics is a heterogeneous business, one especially ambitious form of metaphysical work has achieved a kind of pure expression in recent decades. This has made the "modelish" aspects of metaphysics more conspicuous than before.

This hypothesis does not assert a mismatch merely between ideology and actual practice, however. There is also a mismatch between some aspects of the practice and others. I will argue that metaphysicians sometimes develop good models, but try to treat them in further work in a way that is inappropriate for theoretical constructs of this kind. So this part of my analysis has a more prescriptive side. I suggest that some kinds of metaphysical debate might be more productive and avoid some needless wrangling if practitioners were to consciously take on the meta-theoretic perspective offered here. A case study is developed in the last section of the paper.

## §2. Systematic and Analytic Metaphysics

HERE ARE SOME EXAMPLES OF CLAIMS IN METAPHYSICS:

(i) The best theory of properties is a "universals plus particulars" view. Trope resemblance nominalism is the second best. (Armstrong)
(ii) Modal facts are facts about other possible worlds, which are as real as our actual world. (Lewis)
(iii) Laws of nature are relations between universals. (Armstrong, Dretske, Tooley)
(iv) All there is to the world is a vast mosaic of local matters of particular fact. ("Humean Supervenience," Lewis)
(v) Causal relations are chains of counterfactual dependence relations between events. (Lewis)[3]

I will distinguish two aspects of this sort of work, though they are often combined. These aspects will be labeled "systematic" and "analytic" metaphysics. What I call "systematic" metaphysics is work intended to be about the world itself, and not about the relation between particular ordinary concepts and the world. I use the term "analytic metaphysics" for work that includes a significant role for investigation of what our concepts commit us to (and questions of that kind), as well as investigation of what there is.

For example, assessments of the above sample claim *(v)*, concerning causation, tend to include a great deal of this conceptual work, at least in practice. Assessments of claim *(i)*, which concerns the reality of universals, usually do not. This paper will focus primarily on systematic metaphysics, but will discuss analytic metaphysics in §5.

What is the self-conception that we see in systematic and analytic metaphysics over recent years? An appropriate description might be "cautious attempted theorizing." Armstrong, for example, often says that the best we can do in metaphysics, most of the time, is to consider various candidate "package deals," that is, combinations of views that can be shown

to work well together. But if we ask what these package deals are "candidates" *for*, the answer appears to be that they are candidate descriptions of how the world is really constituted. It is hard to tell whether or not there are universals, for example, but their existence is a genuine theoretical option. Work of this kind resists deflationary analyses of what it means to assert metaphysical doctrines, what it means to favor one option over another. Some positions are more parsimonious, convenient, and elegant, but there is held to be a gap between those virtues and truth. In a particular case, it might be found that some range of prima facie rivals are completely equivalent to each other, but that is something to argue for, and it is not supposed to be the usual case.

Lewis's comments about "Humean supervenience" in his introduction to *Philosophical Papers Vol. II* (1986a) provide another useful example. Lewis says that Humean supervenience is a contingent claim about our world (and worlds like it). Lewis is willing to consider the possibility that physics might show Humean supervenience to be false, but he thinks it can be defended against many other objections.

Considerable pressure can be put on metaphysical theorizing of this kind. Examples include:

(i) *Epistemic pressure:* Any phenomenon in the relevant domain can be accommodated by all the well-developed rival views. We have no rational way of choosing between the options.
(ii) *Verificationist and other semantic pressure:* The apparent contrasts literally make no sense. We think we are posing and assessing rival options but the words are empty.
(iii) *Disconnect from science:* Insofar as traditional metaphysical questions are real, the only way to adjudicate them is to treat them as foundational problems in physics and nearby sciences.

How do the practitioners of systematic metaphysics respond to this pressure? Some willingly take on extreme caution, at least when speaking outside the immediate give-and-take. Some hope for the right kind of continuity, at the end of the day, with science. I think that many just find it personally inescapable that at least some of the contrasts they debate are real. The metaphysician might have no available philosophy of language or epistemology that makes good sense of this fact, but the contrasts do seem non-empty.

I will outline a different response. It includes a role for continuity with science, but it is also related to a family of more deflationary views that use concepts of fiction and metaphor.

## §3. Model-Based Theorizing

My suggestion is that the theoretical constructs developed in systematic metaphysics are best seen as models. Metaphysical system-building is model-building.

I say this using a very specific sense of "model," a sense in which models yield a particular kind of understanding. The term "model" is used in a multitude of ways in science and philosophy. In this paper I am *not* drawing on the technical sense seen in model theory, and I distance the present discussion from the (often misleading) literature which applies that concept to the analysis of theorizing in the philosophy of science.[4] It is best to start afresh from a particular scientific usage of the term, and a particular scientific activity in which the term often appears. This activity within science is more self-conscious now than it was fifty to one hundred years ago.

A rough definition of the relevant sense of "model" can be given as follows: A model is an imagined or hypothetical structure that we describe and investigate in the hope of using it to understand some more complex, real-world "target" system or domain. Understanding is achieved via a resemblance relation, that is, some relevant similarity, between the model and the real-world target system. This account draws especially on the work of Ronald Giere (1988). Giere originally developed this view as part of an analysis of how *all* theorizing works in science. But such a broad claim is both implausible and inessential to the uses to which I will put the key ideas here. It is far more plausible to argue that some theoretical science, though not all, operates as *model-based science*.[5] Model-based science takes an *indirect* approach to representing complex or unknown processes in the real world. The modeler's first move is to specify and try to understand a hypothetical structure, often using mathematical methods. It is a separate question—and often a very subtle one—to work out what sort of similarity there is between the model and events and processes in the real world. Much of the time, in fact, the focus of the scientific work is the first stage, where the aim is to understand the properties and behaviors of various important "model systems." Will an idealized market of a particular kind clear? Will a particular imagined ecological system reach a stable equilibrium? Can a feed-forward neural net with no "hidden units" learn a simple grammar?

I said this "indirect" approach to theoretical investigation is one option. The other option is to try, as well as we can, to *directly describe the target system*—what it contains, how it works, what it will do. We might offer this "direct" description with great caution, but employing caution is different from taking a deliberate detour through an idealized hypothetical system.

In this account of model-based science, a key role is played by some philosophically notorious features of resemblance relations. Especially since Goodman (1955), philosophers have treated resemblance as a slippery and context-sensitive matter. Here, this is not so much a problem as an important "moving part" within the analysis. Two scientists can use the same model to think about the same target system, but with different sorts of resemblance in mind. Sometimes the relevant resemblance relation is extremely subtle; part of the informal schooling of a modeler is giving them a sense of what makes for useful, illuminating resemblance relations between models and real systems.

Defending this view in philosophy of science requires putting considerable weight on the contrast between indirect and direct strategies of

representing a target system. Sometimes that can look difficult. The "direct" approach is surely compatible with some idealization and approximation. Is this really distinct from modeling? In some discussions, this contrast can look difficult to sustain. In the present context, however, this is not a problem. The part of model-based science that I will focus on is the part that is most clearly different from traditional conceptions of theorizing.

Models of the kind I have in mind are often found in frontier areas where our knowledge is poor and the system is complex. The models are highly schematic. The aim is to explore the behavior of structures that might be of the right *kind* to postulate in a direct description of the target at some later stage. Often, the aim is to look at systems that exhibit dependence structures that have some of the right features. That is, the aim is to find structures by which *something like* the special properties of the target system can be generated from plausible kinds of ingredients.

A certain kind of instinct is important in this work. A model that has been built might have many features that have no analogues in the real world. (An example is the unrealistic error signal in "back-propagation" learning models in cognitive science. Another, which might be especially suggestive to metaphysicians, is a recent biological model which treats the *absence* of organisms as the occupation of that space by a species of non-organisms, the species "Void," which competes and reproduces along with the other organisms in the ecology [Mitteldorf and Wilson 2000].) These unrealistic features exist in the model to make it possible to explore the behavior of other features of the model which *may* be important. This can be the beginning of a genuine understanding of the target system, and it works via subtle, partial resemblance relations between model and target. A good modeler has an instinct for distinguishing useful idealizations from simplifications that merely "fudge" the problem.

The familiar category of "how-possibly explanation" is relevant here. That concept captures some of what is going on. But what I am describing is then a special kind of how-possibly explanation, because a model will in many cases have features that are *not* relevantly possible. The model is not, as a whole, a literal candidate for how the target behavior might really be produced. Or at least, this is not necessary for a model to be useful.

A useful scientific illustration of this work can be found in biological work on the origins of life, and other early stages in the evolution of life. This is very much a "frontier" modeling field. The origin of life is a complex problem, and much of the crucial information was lost billions of years ago. Consequently, the research involves a lot of modeling of possible scenarios; modeling of how various kinds of early conditions might in principle give rise to structures with basic capacities we associate with life. A modeler might start with some assumptions about a chemical soup—often not a soup of real chemicals, but of idealized chemicals. The aim is to show how ingredients of a certain kind, with lots of energy around, could give rise to processes of reproduction or replication, to the appearance of cell-like units that have insides and outsides, to heredity, complex metabolism, and cooperation of the kind that makes multi-cellular life possible.

One line of work follows a "replicator first" approach. The aim is to look for some crucial simple molecule that can replicate itself and hence initiate an evolutionary process. In contrast, "autocatalytic network" approaches explore larger self-perpetuating chemical networks (Kauffman 1993, Maynard Smith and Szathmary 1995). Each view temporarily ignores some aspects of the problem. The aim is to specify some reasonable schematic ingredients and get them to produce relevant analogs of the important behaviors. When we do, we have the beginnings of an understanding of how life-like activities are related to the simpler state from which they arose. This style of modeling is not restricted to work on life's origins. Later steps in early biological evolution are often treated in the same way. A term used sometimes in areas like this is "toy model." This is not an insult, but a descriptive term, often self-applied, which indicates the kind of extreme idealization discussed above.[6] Work of this kind can be controversial in some quarters, however. It is highly speculative, and often somewhat unempirical. Within the part of science I know best, evolutionary theory, highly idealized modeling of this kind dates back to the 1920s (Fisher 1930). The prevalence of this style of work has increased in recent decades, however, and has also become more self-conscious (Levins 1966).

So far I have emphasized the role of ignorance and complexity in pushing us towards this sort of modeling. But models of this kind can also be used when the target systems are better known. In these cases, the aim of the model is usually to achieve high *generality*—to abstract away from a huge amount of empirical detail. Again, the aim is often to show how certain kinds of structures suffice for the generation of some important class of behaviors. The work guided by generality is similar to the work guided by complexity and ignorance in not being strongly guided by the desire for an exact fit to empirical data.

This is the kind of scientific work that I want to compare to systematic metaphysics. We now reach the core of my proposal. Many of the constructions of systematic metaphysics should not be seen as hypotheses in the normal sense. They should instead be seen as models. Work in systematic metaphysics explores model structures that can have genuine *application* to the world, but of a special and indirect kind. The relation I am here asserting between a scientific activity and a philosophical one is not merely one of analogy. It is stronger than that, or it can be in some cases. I suggest that some models in systematic metaphysics have the *same kind of relation* to the world that we see in some model-based science.

In science, this kind of modeling is now fairly self-conscious and has its own subculture. In systematic metaphysics, the usual self-image that people have is different (the term "model" might be used in systematic metaphysics, but not usually to invoke the strategy I describe here). So it is important for my account that a person can be engaged in modeling while having a different self-understanding. Modeling in science used to be more like that. My suggestion, more strongly, is that something like the modeling we see explicitly in science is a human activity that occurs very naturally in certain intellectual contexts. It might involve a distinctive kind of cognitive

processing. There is quite a long tradition in cognitive psychology of positing a general faculty of "model-based reasoning" (Johnson-Laird 1983, Gentner 2002). I do not want to tie my analysis here to the details of that tradition in psychology, but some parts of it are congenial and suggestive. What is most useful, perhaps, is the idea that modeling is something that our minds naturally do when confronted with certain kinds of theoretical problems. Model-based *science* would then be seen as a self-conscious, formalized manifestation of a more basic psychological capacity. Against that background, my suggestion is that metaphysical systems and debates show the marks of that distinctive kind of cognitive capacity. This is the more adventurous, psychologistic version of my analysis that was mentioned in my introduction.

The products of systematic metaphysics can function as models, be useful as models, despite being surrounded by the ideology of direct theorizing. But it is probably better for practitioners to have an explicit understanding of what is going on. This kind of understanding has probably been helpful in the case of scientific modeling. That thought provides the basis for the more prescriptive, revisionary remarks about metaphysical theorizing that will come in the final section below.

Not all metaphysical discussion, and probably not all of systematic metaphysics, is concerned with models. For example, physicalism (in at least some forms) is not a model in my sense. Physicalism is a theory—a scientific theory, primarily, but one with a philosophical role. It also provides a framework within which modeling can be done.[7] So some ideas in systematic metaphysics do not operate as models. But some do.

The next question to ask is what difference it makes to a discipline when models become the overt currency of theoretical work. A key difference has to do with the ways in which particular ideas should be assessed. Let us return for a moment to the scientific case—here I draw again on Giere. In a science textbook, we might see a description of a simple pendulum in Newtonian terms, or a description of how a simple biological population grows. A person might ask of that description: Is it really true? The description will certainly be true of some imaginary pendulums or imaginary populations. But what is the relation between the description in the book and entities in the real world? There, the short answer is that the description is usually *not* literally true of its intended real-world targets. But the model description can specify a model system that is usefully similar to the target. By understanding one, we learn something about the other.

Truth has not fallen out of the picture entirely. We can look at a model and explicitly assert various hypotheses about which target systems it resembles and in which respects the resemblance holds. Giere calls these claims, in the scientific case, "theoretical hypotheses." These hypotheses can be true or false. And sometimes the match between model and target is so close that there is little need to worry much about the detour through resemblance relations; we can just say "these equations are approximately true of this real system." But in other cases, and especially when close attention to the relations between rival options is needed, the detour *is*

important. Then if we want to say something that is really true, we should say: the following is a good model of this system in such-and-such respects.[8]

Much of the time, we need not obsess about saying things that are literally true. We just work within the model, develop it, show its resources, and assail those who would use other models instead. This is how our general capacity for model-based understanding seems to operate. But if a synoptic, reflective, understanding is our goal, then it is important to be able to step back and say exactly what we are up to.

### §4. Models, Fictions, and Metaphors

With my main proposal on the table, I will look briefly at the relation between my view and some other unorthodox construals of metaphysical discourse. I have in mind a family of theories, including these examples:

(i) Carnap's later view, in "Empiricism, Semantics, Ontology" (1956), with its notion of "linguistic frameworks" and a distinction between internal and external questions. External questions—which often look like metaphysics—are assessed pragmatically and are properly directed at terms themselves.
(ii) Yablo's adaptation of Carnap's view, in "Does Ontology Rest on a Mistake?" (1998). Yablo treats metaphysical discourse as metaphorical, as occurring within make-believe. He adds that lots of highly functional discourse in science and elsewhere is permeated by this sort of metaphorical usage.
(iii) Fictionalism, especially of the kind discussed by Rosen (1990) for the case of possible worlds.

My intention is to partially ally the model-based view with this tradition, while emphasizing its distinct handling of some issues. Other views in this tradition somewhat undersell the substantive role of systematic metaphysics.

Here is a key distinction. Suppose we have a body of metaphysical theory—a theory of possible worlds, for example. Various views can be counted as *instrumentalist* if they hold that the value of this body of theory consists solely in the ways it helps us in *other* kinds of theoretical or practical activity. The utility of metaphysics is entirely output-oriented. The ideas play an organizing role, a highlighting role, a paraphrasing role of some kind, with respect to ordinary (perhaps "empirical") claims. I see fictionalism, Yablo's view, and probably Carnap's original view, as being in this instrumentalist category. For Carnap, for example, the pragmatic assessment of the utility of linguistic frameworks is real and important. Metaphysical discussion is—at best—a somewhat awkward and misleading way of framing such assessments. What is denied by instrumentalism in this sense is that there is some *other* kind of application or fidelity to the world that the metaphysical apparatus might have, besides generating useful output.

The model-based analysis of metaphysics is not supposed to be instrumentalist in this sense. Consider the scientific case again. One thing a person might do with a model is make predictions. And a person might say:

"this model and its target are similar with respect to input-output profile, and that is all I care about." But there is not usually a dichotomy between "realist" and "instrumentalist" applications of a scientific model. These are extreme points on a continuum, or, really, in a space of possible applications. There are lots of ways we can take a model to have subtle and partial resemblance relations to the target system, lots of ways we can regard the model as capturing *some* coarse-grained hidden dependence structures in a target system.

The same applies, potentially, in the philosophical case. "Prediction" is not the output of a metaphysical system that is usually envisaged, but rather some kind of organizing or systematizing of empirical information. In other respects, though, the situation is similar. Instrumentalist views are arrived at by setting a parameter that determines how a model is to be treated at a particular extreme value.

For example, we can think of the Lewis/Kripke/Stalnaker array of possible worlds as a model. Here the "modelish" talk is not so far from existing practice. In part, that is because of the role of the model-theoretic sense of "model," but in part, I suspect, it is because of the surfacing of motivations similar to mine. We are presented with an imaginary structure; we think that something useful can be done with it, but we are not sure about its exact status and relation to the real world. One well-known view of possible worlds is fictionalism (Rosen 1990). In my terms, this is one way to construe the model, and it involves an instrumentalist stance in the sense just discussed. Further, I think this is probably an entirely appropriate view of possible worlds—superior to a view that looks for some extra resemblance or correspondence between the modal array and some aspect of the real world. The framework is judged not just *by* its fruits but solely *for* them, where one important fruit is the systematic treatment of modality in terms of quantification. So fictionalist positions can then be seen as special cases of a model-based view.

### §5. Analytic Metaphysics

THE FOCUS SO FAR HAS BEEN ON "SYSTEMATIC" METAPHYSICS. BUT MOST CONTEMPORARY metaphysics is not purely of this kind. Much familiar work includes a mixture of two projects. One is the project of systematic metaphysics, understanding the furniture of the world itself. The other is looking at the relationships between important *concepts* that we have, and the world as described by systematic metaphysics. Philosophers take concepts like *cause*, *law*, *freedom*, or *belief*, and see whether these concepts usefully apply to anything in the world as our systematic metaphysics sees it. Or, they may ask whether these concepts apply to the world as described by systematic metaphysics in conjunction with empirical information. This project can also include consideration of other kinds of "coordination" with the world that our conceptual apparatus might enable, beside simple descriptive relations; here I include "projectivist" analyses and the like.

So this work involves a dual investigation. Part of it is systematic metaphysics, and part is an investigation of what causes, freedom, or beliefs

are supposed to be like. This dual project is the explicit emphasis of Jackson's work (1998), and the "Canberra plan."[9] It is less overt, but omnipresent, in Lewis, and less prominent in Armstrong. This dual project is somewhat transformed by my view of systematic metaphysics. In this section, however, I will discuss a feature of the other side of the work, the more "conceptual" side.

Philosophers investigate the facts on the conceptual side mostly via intuitions about imagined cases. This is becoming more controversial. Do we really think that philosophical intuitions are a good guide to the actual patterns in ordinary thought and talk? Some think so (Jackson 1998, ch. 2). But there is another option, which makes better sense of actual practice. Analytic philosophers tend to adopt, tacitly or explicitly, idealized and cleaned-up variants of the actual patterns of everyday talk. These cleaned-up patterns can be compared more readily to specific systematic metaphysical views. I suggest that people greatly underestimate how much of this there is.

How does work of this kind actually proceed? I take as my example Lewis's and some post-Lewisian work on counterfactuals and causation (Lewis 1973, Collins, Hall, and Paul 2004). On one side we have a piece of systematic metaphysics: a framework of possible worlds, Humean supervenience, and counterfactual relations between events, based on similarity. On the other side, we have patterns in how we talk and think about causation in ordinary life and science. The aim is to see how the two sides match up.

Lewis is interesting for his ability to hear nuances in how people actually talk, but also for his willingness to subtly regiment the linguistic/conceptual patterns that he aims to capture. In the case of causation, Lewis's strategy was to take singular causation as primary, in comparison with general causal facts, and also to base the analysis on a weak, non-exclusive notion of one thing being "a cause" of another, as opposed to being "the cause." This meant that the form of causal claim to which Lewis accorded the central place was "X was a cause of Y." Once we are thinking inside Lewis's metaphysical picture, this seems quite natural. However, it is also true that outside of philosophy, hardly anyone seems to say that one event "was a cause of" another. In descriptions of singular causal relations, I suggest that it is fairly common to hear "the cause," and more common to hear the verb "caused," along with the "special causal verbs" whose role Anscombe (1971) highlighted. "X was a cause of Y," in the sense analyzed by Lewis, is not quite ordinary language. I don't deny that it appears outside of philosophy, but it does so in a special studied and hedged form of usage. Lewis focused on "was a cause of," because it had better contact with the systematic metaphysical picture he was working with. Since then, much of the literature has followed suit.

What has happened here? I suggest that analytic metaphysics is constantly adopting, and sometimes questioning, idealized pictures of ordinary usage in these areas. Some philosophers claim there is something called "our concept" of cause whose "structure" is imperfectly reflected in

ordinary talk. But many other philosophers are wary of moves like this, and rightly so. They are tendentious and also unnecessary. It is perfectly reasonable for a philosopher to say, instead: "I am not a linguist, and I am not going to usurp the linguist's role. I am going to operate with a cleaned-up and compact variant of the actual patterns of talk and thought in this area. I am going to see how this way of thinking and talking *would* relate to the world as described in my systematic metaphysics." The reader will see where this is heading. Analytic metaphysicians engage in something akin to modeling at *both* ends of the process. That is how they can avoid being hostage to both the linguists and the physicists.

Both ends of the process look like modeling, but I do not want to overstate the case. The kind of idealization on the conceptual side that is essential to contemporary work in analytic metaphysics *can* be described in the language of modeling, but it is not the kind of process in which the special features of modeling emphasized in earlier sections of this paper are important. So I certainly would not insist on describing it that way.

### §6. Humeanism and the Mosaic Model

In this final section I will discuss a particular metaphysical debate in detail. The aim is both to illustrate the claims outlined in earlier sections, and to argue that an explicit awareness of the nature of modeling might facilitate progress in metaphysical discussion. My case study will be the view that David Lewis calls "Humean supervenience," along with some of its denials and other ideas in the same area.

Lewis describes the doctrine of Humean supervenience as follows: "all there is to the world is a vast mosaic of local matters of particular fact, just one little thing and then another" (1986a, p. ix; see also Lewis 1994). We have a mosaic of local instantiations of natural qualities, and no more. Lewis's is a richer mosaic than some others, however, because it is not strongly nominalist. Lewis is friendly to the idea that some classes are more "natural" than others (Lewis 1983).

We should distinguish between the idea of the "world as mosaic" itself, and Humean denials of necessary connections between distinct parts of the mosaic. Seeing the world as a mosaic need not itself imply a "loose and separate" view of the relations between parts. In some mosaics, once we have laid tiles of one shape, we are constrained in the shapes of later tiles laid. Lewis's mosaic, though, is of the Humean loose and separate kind.

For both Lewis and Hume, a central goal is the denial of dubious forms of metaphysical connection. Hume's denial is based in epistemological doctrines. Lewis refers to the kinds of connections he rejects as "rubbish" (1986a, p. x). But of course, not all connections between distinct things are rubbish; we spend much of our lives tracking and utilizing them. So the Lewis project becomes one of showing that a mosaic view can accommodate genuine, legitimate kinds of connection. The result is a series of analyses: of laws, counterfactuals, causes, persistence through time, and probabilities.[10] The analyses proceed by finding good candidates in the mosaic for well-

functioning everyday and scientific claims about causation (and so forth) to be about.

These can be described as analyses of various kinds of connection in terms of facts about *patterns*. So the problem, at least for some philosophers, can be put by saying that what often need explanation are the natural facts that *generate* patterns in the world. Laws are a clear example, and chances an especially hard one. Significant weight then falls on a project in analytic metaphysics: showing that the kinds of connectedness we are *committed* to and *need* are compatible with Humean supervenience. Richer connections can be dismissed as philosophical fantasies.

For dissenters, the connections that are admitted by Lewis are not enough. A clear example is Armstrong's treatment of laws (1983, 1987). Armstrong posits a higher-order universal, N, a relation of nomic necessitation that holds between universals. This is a non-Humean extra connector that constitutes relations of natural law. The introduction of N, for Armstrong, also makes sense of an important set of epistemic facts. For Armstrong, "induction" is analyzed in terms of an explanatory inference from local facts to a natural law, and then inference from the law (and additional information) to the presence of other local facts.

Armstrong's account is driven, here and elsewhere, by a desire to give a genuine explanation for facts about connectedness that other philosophers either deflate or take as unanalyzed. As a project of this kind, Armstrong's view has what I see as a model-like shortcoming, about which he is entirely explicit. Our aim is to explain connectedness, but who will connect the connectors? The problem arises in both general and specific forms. In general, for Armstrong, the universals must be attached somehow to the particulars. This is apparently a relation. In response, Armstrong has postulated a "fundamental nexus," a *sui generis* link that is exempt from his usual explanatory demands. In Armstrong's more recent work, the ontological primacy of states of affairs is supposed to help with the problem. The same kind of problem arises in a more specific form in the need to say how it is that N connects to the universals it ties together in law.

These problems—who will connect the connectors?—have an interesting history in metaphysics, as Armstrong notes. Plato's "third man" argument and F. H. Bradley's "relations regress" are in the same family. For philosophers like Bradley, the upshot of the problem is very much an anti-mosaic one.

The dialectic here is familiar. We knew already about the features of Humean supervenience and Armstrong's N discussed in this section. But I hope that now we can see these facts in a new way. What we are seeing here are the limitations of certain kinds of models. We build a structure with ingredients of a certain kind. We show how some relevant features and behaviors can be generated. But it is easy to push the model too hard. To treat the world as a mosaic of distinct existences is specifically to make mysterious the apparent fact that there is *constraint* exercised by one part of the world on another—the fact that we can attend to one part of the world and learn about another, make things happen here by intervening there. These

are things that a pure mosaic view does not capture. Armstrong's view shows the consequences of introducing, within the same general framework, extra entities to re-establish the connections that have gone missing.

Within the usual dialectic, one common response is to think that the mosaic view and Humean supervenience really do push us towards surprising conclusions about constraint, inference, and intervention. This, I suggest, is the wrong train of thought. Humean supervenience and Armstrong's anti-Humeanism are both toy models of the universe. They are informative in specific contexts. They show the in-principle possibilities implicit in certain kinds of structures and ingredients. Some kinds of commerce between human thought and language and the rest of the world can be illuminated by treating the universe as a mosaic. Other phenomena, however, are obscured or lost. When we push up against the limits of the mosaic view in its handling of connectedness, we are seeing the limitations of an otherwise useful model.

For another analogy, think again about Lewisian possible worlds. Lewis locates modal facts by giving us extensions—an organized array of real things for modal talk to be about. The problem arises that, for any array of real things, we can apparently ask a modal question: Why not more, or less, of whatever this is?[11] In the case of Lewis's array of worlds, as in some earlier theological cases, we are not supposed to ask this modal question. But the instinct is to ask it anyway. Here again, what we see are the limitations of a model, showing up in regress-like form. Whether this is a problem depends on how the model is being treated.

To make the issue very stark, consider the following series of moves, which involve an obvious mistake: We deliberately idealize away from some features of a system for the purpose of some specific task, then, forgetting about the idealization, we argue for the non-existence of the things that have been idealized away from! Things are not as bad as that, in the case of Humean supervenience. But the situation is not so far off. It is something like this. We develop a model for some purpose X, making whatever idealizations seem appropriate. We then ask: Is the resulting model also able to deal with goal Y? Not entirely, but with lots of work we can come fairly close . . . . We then infer that our X-oriented model is a full description and goal Y must be deflated.

What are these "goals" in the case of Humean supervenience? If we go back to Hume himself, it is not so much a question of "goals." Hume thought that for psychological and epistemological reasons, no other framework was available. But those arguments in psychology and epistemology have now left the scene. Where else might we look? One cannot say that Humean supervenience has been designed to fit with modern physics—here Lewis's discussions are primarily defensive.[12] Sometimes the preference seems aesthetic; Lewis, as noted earlier, calls non-Humean entities "rubbish." In his 1994 discussion he presents the view as a bulwark against unscientific additions to our ontology from the side of philosophy, but Humean supervenience surely has too much positive content for that to be

the whole story. In Lewis and elsewhere, I suggest that the real motivation for modern Humeanism is more positive than negative.

In Lewis and others, the mosaic model and Humean supervenience are attractive because of a particular kind of problem-solving and systematizing power. If we think of the universe in this way, we can bring to bear a set of combinatorial and quantificational tools. These tools can be used to describe patterns and develop analyses of a kind that modern philosophers find illuminating and satisfying. I am not arguing against those projects. Some questions about the world and our dealings with it can be fruitfully handled this way.

My argument is against a certain kind of *response* to the usefulness of the mosaic model—the response that infers that this is a promising candidate for being a literal description, something to be affirmed or, at philosophical expense, denied. In particular, we should not say things like this: "If I reject Humean supervenience, I will have to then believe some other view in the same category. If I have to believe one of these things, I will believe Humean supervenience. It is the best of the bunch. I will deny rich concepts of constraint and connectedness, even if I court skepticism and other problems thereby." These are the wrong kinds of conclusions to draw from the products of model-building work. φ

## *Notes*

Acknowledgment: This paper was presented at Daniel Stoljar's conference on philosophical methodology (ANU, 2005), and also at Adelaide University. I am grateful to those present for comments and criticisms. I am also grateful to Greg O'Hair and Dmitri Tymoczko for especially useful comments.

[1] An example is a workshop at the ANU in 2005 focused entirely on the question of whether "ontological questions have determinate answers."

[2] Laurie Paul is a partial exception, though she does not agree with the details of my analysis. In a work in progress, Paul is developing her own version of the idea that metaphysical work involves an essential role for models.

[3] For *(i)*, see Armstrong (1987). For *(ii)*, see Lewis (1986b). For *(iii)*, see Dretske (1977), Tooley (1977), and Armstrong (1983). For *(iv)*, see Lewis (1986a). For *(v)*, see Lewis (1973) and (1986a).

[4] Suppe (1977) is a standard source for this research program. For criticism of relevant kinds, see Downes (1992), Thompson-Jones (2006), and Godfrey-Smith (forthcoming).

[5] For a more detailed treatment, see Weisberg (2003) and Godfrey-Smith (forthcoming).

[6] See ch. 3 of Dyson (1985), "A Toy Model." Dyson might the originator of the term. See also Szathmary (1994), "Toy Models of Simple Forms of Multicellularity, Soma and Germ."

[7] This is not to say that the doctrine of physicalism has been made fully clear. But in making the doctrine clearer, the right way is probably not via the notion of a model.

[8] It is important not to require an isomorphism between the two, as some discussions in the philosophy of science have done. The similarity can be more subtle and holistic than that. See also Giere (1999).

[9] Mackie's classic treatment of causation (1974) is another good example.

[10] If a Humean view is more nominalist than Lewis's is, there is also the problem of natural kinds.

[11] Rosen (1990) discusses this in connection with fictionalism.

[12] In his 1994 discussion Lewis says that Humean supervenience is "inspired by classical physics" (p. 474), and assumes it can be adapted to fit in with the deliverances of "better" physics.

## *References*


Anscombe, G. E. M. 1971. Causality and Determination. Repr. in *Causation*, ed. E. Sosa and M. Tooley (Oxford: Oxford University Press, 1993), pp. 88–104.

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

———— . 1987. *Universals*: *An Opinionated Introduction*. Boulder: Westview Press.

Carnap, R. 1956. Empiricism, Semantics, and Ontology. In *Meaning and Necessity*. 2nd ed. (Chicago: University of Chicago Press), pp. 205–221.

Collins, J., E. Hall and L. Paul, eds. 2004. *Causation and Counterfactuals*. Cambridge: MIT Press.

Downes, S. 1992. The Importance of Models in Theorizing: A Deflationary Semantic View. In vol. 1 of *PSA 1992*, ed. D. Hull, M. Forbes and K. Okruhlik (East Lansing: Philosophy of Science Association), pp. 142–153.

Dyson, F. 1985. *The Origins of Life*. Cambridge: Cambridge University Press.

Dretske, F. 1977. Laws of Nature. *Philosophy of Science* 44: pp. 248–268.

Fisher, R. A. 1930. *The Genetical Theory of Natural Selection*. Oxford: Clarendon Press.

Gentner, D. 2002. Mental Models, Psychology of. In *International Encyclopedia of the Cognitive and Behavioral Sciences*, ed. P. Bates and N. Smelser (Amsterdam: Elsevier), pp. 9683–9687.

Giere, R. 1988. *Explaining Science: A Cognitive Approach*. Chicago: Chicago University Press.

———— . 1999. Using Models to Represent Reality. In *Model-Based Reasoning in Scientific Discovery*, ed. L. Magnani, N. J. Nersessian, and P. Thagard (New York: Kluwer / Plenum) pp. 41–57.

Godfrey-Smith, P. Forthcoming. The Strategy of Model-Based Science. *Biology and Philosophy*.

Goodman, N. 1955. *Fact, Fiction, and Forecast*. Cambridge: Harvard University Press.

Jackson, F. 1998. *From Metaphysics to Ethics*. Oxford: Oxford University Press.

Johnson-Laird, P. 1983. *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Cambridge: Harvard University Press.

Kauffman, S. 1993. *The Origins of Order*. Oxford: Oxford University Press.

Levins, R. 1966. The Strategy of Model-Building in Population Biology. *American Scientist* 54: pp. 421–31.

Lewis, D. L. 1973. Causation. Repr. with postscripts in Lewis (1986a), pp. 159–213.

———— . 1983. New Work for a Theory of Universals. *Australasian Journal of Philosophy* 61: pp. 343–377.

———— . 1986a. *Philosophical Papers*, vol. 2. Oxford: Oxford University Press.

———— . 1986b. *The Plurality of Worlds*. Oxford: Blackwell.

———— . 1994. Humean Supervenience Debugged. *Mind* 103: pp. 473–490.

Mackie, J. L. 1974. *The Cement of the Universe*. Oxford: Oxford University Press.

Maynard Smith, J. and E. Szathmary. 1995. *The Major Transitions in Evolution*. Oxford: Oxford University Press.

Mitteldorf, J. and D. S. Wilson. 2000. Population Viscosity and the Evolution of Altruism. *Journal of Theoretical Biology* 204: pp. 481–496.

Rosen, G. 1990. Modal Fictionalism. *Mind* 99: pp. 327–354.

Suppe, F., ed. 1977. *The Structure of Scientific Theories*. 2nd ed. Urbana: University of Illinois Press.

Szathmary, E. 1994. Toy Models of Simple Forms of Multicellularity, Soma and Germ. *Journal of Theoretical Biology* 169: pp. 125–132.

Thompson-Jones, M. Forthcoming. Models and the Semantic View. *Philosophy of Science* (Proceedings of PSA 2004).

Tooley, M., 1977. The Nature of Laws. *Canadian Journal of Philosophy* 7: pp. 667–698.

Weisberg , M. 2003. *When Less is More*. PhD diss., Stanford University.
Yablo, S. 1998. Does Ontology Rest on a Mistake? *Proceedings of the Aristotelian Society*, supplementary volume 72: pp. 229–62.