Multiple Realizability and the Semantic View of Theories*

Colin Klein
Department of Philosophy, M/C 267
University of Illinois at Chicago
601 S Morgan St.
Chicago, IL 60607
cvklein@uic.edu

*0812 draft 1De1; Compiled March 27, 2010.

1 Introduction

1.1 A Puzzle

In a famous passage, Hilary Putnam has us imagine two distinct explanations of the failure of a peg to pass through a hole. One explanation uses a complex derivation from the properties of the atoms of in the peg; the other just cites the shape of the peg and the hole. The latter explanation, he claims, is superior because

> In this explanation certain *relevant structural features of the situation* are brought out. The geometrical features are brought out. It is *relevant* that a square one inch high is bigger than a circle one inch around. And the relationship between the size and shape of the peg and the size and the shape of the holes is *relevant*. It is *relevant* that both the board and the peg are *rigid* under transportation. And nothing else is relevant. The same explanation will go in any world (whatever the microstructure) in which these higher-level structural features are present. In that sense *this explanation is autonomous* ([Putnam, 1975] 296)

Many find this intuition, or something like it, a compelling reason to be a nonreductive physicalist. Putnam sketches a picture in which distinct sciences give different explanations of the same event. Some sciences give explanations that abstract away from messy details. These explanations are that very reason, superior to the explanations given in lower-level sciences. Further, reductive physicalists often *accept* this characterization. They respond that the complex lower-level explanation is nevertheless acceptable or preferable, perhaps because the relevant notion of explanation is an in-principle one, or because despite its verbosity, the lower-level story possesses counterbalancing explanatory virtues [Kim, 1998, Sober, 1999].

Yet I find this all very puzzling. All parties in this debate appear to accept what Putnam implies about explanation in physics: to wit, that the only explanation available to physics is the one that involves tediously tracking the micro-constituents of the peg and the hole. That is obviously false. Physicists give explanations that involve shape properties all the time. Pick any introductory physics text and flip to the section on Gauss' law: explanations in terms of shapes abound. Similar examples can be found in mechanics, fluid dynamics, and just about everywhere one looks: the shape of objects is important in physics, and important precisely because it allows for explanations that abstract away from tedious details. The peg-and-hole case does not show a conflict between physical explanation and the explanation given by some other science: it shows a simple conflict between two *physical* explanations, one of which is obviously superior to the other.

Perhaps the example is just ill-chosen. Block, for example, mentions in passing that the peg-and-hole case is flawed in the way I've noted, but says that we could certainly give better cases if we wanted to ([Block, 1997] 130fn2). I think not. I think that the failure of the peg-and-hole story shows that philosophy of mind is still carrying around some baggage in the form of an outdated late positivist philosophy of science. I will argue that the notion of multiple realizability (MR) is one of the more pernicious bits of that baggage.

1.2 Multiple Realizability

At a first pass, multiply realizable properties are those whose realizers are 'too diverse' from the perspective of physics to figure in unified explanations. So Eric Funkhouser, in a recent review of the MR literature, argues that "...differences that cannot be discerned from the perspective of a particular science ground their MR." ([Funkhouser, 2007] 313) Theories which talk about MR properties are, for this reason, thought to be *irreducible* to lower-level sciences. As MR properties do not admit of a unified treatment by physics, physics cannot give a proper explanation of theories that contain these properties.

But in fact, definitions of Multiple Realizability rarely deliver anything that strong. Instead, most conditions on MR only ensure that the sets of realizing properties be different from each other in *some* physically relevant way. Aizawa and Gillett's recent formulation is especially clear in this regard. They argue that the conditions on MR are:

> (Multiple Realization) Instances of a property G are multiply realized if and only if, (i) under condition \$, an individual s has an instance of property G in virtue of the powers contributed by instances of properties/relations $F_1 - F_n$ to s, or ss constituents, but not vice versa; (ii) under condition \$* (which may or may not be identical to \$), an individual s^* (which may or may not be identical to s) has an instance of a property G in virtue of the powers contributed by instances of properties/relations $F_1^* - F_m^*$ of s^* or s^*s constituents, but not vice versa; (iii) $F_1 - F_n \neq F_1^* - F_m^*$ and (iv), under conditions \$ and \$*, $F_1 - F_n$ of s and $F_1^* - F_m^*$ of s^* are at the same scientific level of properties. ([Aizawa and Gillett, 2009] 188)

The relevant condition is (iii), which ensures that realizing properties are distinct (and, presumably, distinct in some physically relevant way). But that's just to say that physics *can* fail to give a unified treatment of those properties, not that it *does*. There's a missing step in the argument.

What plugs the gap is a (rarely articulated) principle:

(S) The only legitimate explanations in a science ϕ are the ones that describe the domain of ϕ in the most specific way that ϕ allows.

If [S] were true, then we could move from the fact that we have *an* extremely specific physical explanation to the fact that tediously specific explanations are *all* physics can give us. That would do the work of blocking reduction (or let's suppose).

[S] is entirely implausible, however. First, it is descriptively implausible. Physicists can and do give multiple explanations of the same phenomenon using different vocabulary. Sometimes they care about tedious explanations based on the most specific description of individual items. Other times, physicists invoke equivalence classes of items—as, for example, they do when they appeal to shape. Physicists also appear to treat the latter explanations as valid and illuminating.

Note that this point is not restricted to physics. It also holds in neuroscience, which is surely the most relevant science when we talk about reduction of mental states. The most specific descriptions neuroscientists can give would be in terms of individual neurons and their spatial distributions. In many cases, however, such descriptions would be unutterably complex.[1] So neuroscientists also talk about Brodmann's areas (defined by similar cytoarchitectural features), gyri and sucli (anatomical landmarks based on the cortical convolutions), volumes located via a standard coordinate system (e.g. Talairaich coordinates), and functional subdivisions (often defined by homology with other mammals).[2] None of these are the most specific descriptions that neuroscientists could give. They do not seem to mind. So [S] is false.

Second, [S] is philosophically objectionable. Explanations serve, at least in part, to give information. Scientists, no less than ordinary folks, are bound by Gricean conversational maxims when they give information.[3] The maxim of Relevance requires speakers to give only such information as is relevant for the topic at hand [Grice, 1989]. Giving an explanation of the peg-and-hole in terms of individual atoms would imply that the details of atomic movement are relevant to the hearer's explanatory interests. But that would normally be false. Hence [S] counsels scientists to do something that they obviously shouldn't, and should be rejected.[4]

[1]Logothetis estimates that a $55\mu l$ volume of brain (corresponding to a standard unfiltered fMRI voxel) "contains 5.5 million neurons, 2.25.5 3 1010 synapses, 22 km of dendrites and 220 km of axons" ([Logothetis, 2008] 875).

[2]For a similar argument against Multiple Realizability, couched in terms of 'descriptive grain,' see [Bechtel and Mundale, 1999].

[3]Lewis suggests an argument along these lines in his [Lewis, 1986]. For a recent defense geared towards philosophy of mind, see [Bontly, 2005].

[4]Note that this argument is compatible with currently fashionable 'ontic' views of explanation. Ontic accounts claim that *explananda* are real objects—causes and the like [Craver, 2007]. The present argument does not deny this. It says only Gricean constraints affect how explanations are formulated, and so (indirectly) our estimation of the goodness of particular explanations.

1.3 A Diagnosis

[S] seems to have little support. Yet something like it is necessary to get arguments employing MR off the ground. Why might something like it have seemed so attractive? What follows will argue for a particular diagnosis.The initial formulations of MR were developed within a late positivist axiomatic view of theories. The axiomatic view of theories places significant restrictions on the theoretical vocabulary admissible in scientific explanations. Assuming an axiomatic view thus makes [S] plausible, and with it a certain picture of how MR works.

The problem is that the axiomatic view of theories is implausible. It has largely been supplanted by the so-called *semantic* view of theories. On the semantic view, [S] is (correctly) implausible, because the semantic view places few restrictions on the language in which theories and explanations are formulated. As such, I'll argue, it is very difficult to motivate anything like the standard view of MR. Thus we have good reason to be suspicious of MR, at least in its classical form.

A brief word about the scope of the argument. I will not argue that MR is an intelligible notion, but that in practice no scientific terms fit it.[5] My claim is stronger: that there is no intelligible notion of MR to be found. MR is a technical notion, introduced to do a job: properties that are MR are supposed to be irreducible because they can not appear in unified explanations in physics. I'm going to argue that the main ways of cashing out this idea only sound plausible because of the peculiar limitations of an axiomatic view of theories. When we move to a more plausible, expressively powerful *semantic* view of theories, these accounts of MR fail to do the job they are supposed to do. This should cast doubt on the very idea of MR. After making this point, I'll conclude by considering ways in which a weaker notion of MR might be salvaged.

[5] As for example Kim does on one reading of [Kim, 1992].

2 The Axiomatic Conception of Theories

2.1 Theories

Let's begin with the philosophical baggage, and in particular the late positivist conception of theories as developed in the writings of Carl Hempel and Ernest Nagel.[6] This is sometimes called the 'received view' or 'standard view' of theories, though this is now something of an anachronism. I shall refer to it as the *axiomatic* view of theories—because on this view, theories are conceived of as the best axiomatizations of a domain of phenomena.

On the axiomatic view, a theory consists of two parts. The first part consists of a set of *theoretical postulates*: a finite set of sentences, constructed from a basic vocabulary containing a fixed set of names and predicates, and augmented with the resources of the first-order predicate calculus. Speaking loosely, the predicates in the standard vocabulary are the properties and relations that the theory attributes to the world. The universally quantified statements among the theoretical postulates are the *laws* of a theory. The laws, together with statements of particular fact, allow us to derive particular consequences that predict and explain phenomena.

The second part of a theory is the *coordinating definitions*, which supply a semantics for the theory by connecting at least some of the terms in the basic vocabulary to the world. By the time of Hempel, it was widely agreed that this connection would not be (for example) an exhaustive characterization of theoretical terms via observational terms. Instead, in Hempel's formulation (later imported into philosophy of mind by [Lewis, 1970]), coordinating definitions link theoretical terms to other terms we already have a handle on, often because they occur in natural language. The coordinating definitions, together with the interrelations that the theoretical postulates propose, provide a *partial interpretation* of the theoretical terms. This partial interpretation ideally allows us to connect the predictions of the theory to the world, and so give our theories empirical content.

[6] See [Nagel, 1961], as well as [Suppe, 1989] for a contemporary reconstruction and discussion.

2.2 Multiple Realizability

The axiomatic view of theories—and in particular, the notion that the properties and relations of a theory are just those that correspond to the predicates that appear in the basic laws of the theory—should look familiar to philosophers of mind. In a well-known passage Fodor writes:

> Every science implies a taxonomy of the events in its universe of discourse. In particular, every science employs a descriptive vocabulary of theoretical and observation predicates, such that events fall under the laws of the science by virtue of satisfying those predicates.
> ...
> If I knew what a law is, and I believed that scientific theories consist just of bodies of laws, then I could say that 'P' is a kind predicate relative to S if S contains property laws of the form '$P_x \rightarrow \ldots y$' or '$\ldots y \rightarrow P_x$': roughly, the kind predicates of a science are the ones whose terms are the bound variables in its proper laws.([Fodor, 1975] 13-14)

This is simply the intuition behind the axiomatic approach imported into the philosophy of mind: the discourse of a science is reflected in its basic vocabulary, and the laws of a science are first-order constructions out of that basic vocabulary.

I will turn to the flaws of the axiomatic approach shortly. First, I want to show that the axiomatic view makes MR come cheap, and Putnam's peg-and-hole story seem natural. Let us grant that physics is at least sometimes concerned with atoms and their dynamics (where 'atom' is shorthand for the mereologically simplest bits of the world). On the axiomatic view, this means that predicates like 'x is an atom of type y' must be among the basic vocabulary in which the laws of physics are formulated. Assuming that physics is complete, the laws of physics must include the resources to derive the atomistic explanation of the peg-and-hole. Question: can physics then *also* give—that is, derive—the geometric explanation?

It would seem not. On the one hand, if physics doesn't also contain a predicate like 'x is a circle', it would seem that the geometric explanation is out of reach. Talk about circles would appear to require second-order quantification or the resources of set theory or both (For each x such that x is a simple of some type or other, the location of x is within the set of points equidistant from some single point, and the set of such simples fills the region of points. . .). But ex hypothesi, the theoretical postulates of physics are restricted to a first-order language. So if physics does not already have the resources to talk about circles, the axiomatic view doesn't seem to provide them.

On the other hand, there's a good argument that physics can't contain a predicate like 'x is a circle' in addition to the predicates that apply to atoms. For what could this *add* to our theory as far as empirical power goes? Again, ex hypothesi, the first-order sentences of physics are sufficient to deduce the behavior of the peg and hole. So the addition of predicates about circles only makes the theoretical postulates more complex; it does not allow us to derive anything new from them. Plausible simplicity constraints force us to prefer minimal sets of axioms. The minimal axioms of physics should thus be circle-free. Conclusion: either way you slice it, physics can't talk about circles, and can't give explanations that involve circles, and so some other science must supply the geometric explanation.

This argument can be repeated for *any* term P that is introduced into physics via a definition that does not make reference to the basic vocabulary of physics and which does not explicitly introduce P as a new term within physics. On the axiomatic view the truths of physics are just those that follow deductively from the axioms of physics. Because P does not appear anywhere within those axioms, truths about P do not follow from the axioms. Nor does it seem possible to introduce P into the language in order to allow for such a derivation. Since physics is restricted to first-order quantification, it does not have the logical resources to introduce P into the language directly (at least in most cases). This leaves so-called 'bridge laws' that link P to the terms of the language via linkage to (apparently arbitrary) disjunctions of possible physical realizers. Disjunctive bridge laws, while available in principle, are (let us grant) scientific muddles. Conclusion: facts containing P cannot be derived from physics, and sciences containing P are autonomous, at least insofar as they talk about P.

Should this sound insufficiently empirical, it's also easy to *discover* that P-things are MR. Suppose we find out that the theory-relevant features of a property depend, in Putnam's phrase, on "structural features of the situation." That is to say, we find out that what really matters about a type of thing are the relations between the properties that it has, or the relationship between it and some other things in the world. This is extremely common: one often finds that it is the *shape* of a protein that matters in biology, or the *causal links* between brain areas that matters in neuroscience. Structural properties in this sense are arguably second-order properties from the perspective of physics: they are the property of having some properties or other that stand in the relevant structural relationships. By the reasoning just given, facts that hold in virtue of these structural relationships will not be deducible from the basic axioms of physics. Again MR is had on the cheap.

3 The Semantic View of Theories

The axiomatic view fell out of favor for a number of reasons.[7] Two in particular are worth noting. First, as Suppes notes, first-order formulations of theories are inadequate for many scientific purposes. Any theory that requires, say, the real numbers will be difficult to capture in first-order language. As we saw above, any theory that relies on geometric concepts would would be difficult to capture using only the first-order calculi permitted by the axiomatic view. Further, axiomatizing both the theory and the accompanying math would be, in Suppes' words, "awkward and unduly laborious" ([Suppes, 1967] 58). By this, I take it that Suppes means that even if we can axiomatize the relevant math, it would be inappropriate to include mathematical apparatuses in the theory itself—certainly it is more natural to talk about set theory as something that we use to talk *about* various theories, not something that happens to be part of many distinct theories.

This is surely part of the strangeness of the peg-and-hole example. Putnam

[7] See chapter 2 of [Suppe, 1989] for an extended discussion of problems with the axiomatic account. [Salmon, 1998a], especially [Salmon, 1998b], also contains a number of useful critiques of deductive-nomological view of explanation associated with the axiomatic view.

appeals to geometric facts about the peg and the hole. But geometry does not belong to any particular theory: it's something we use to talk about theoretical entities and draw conclusions from them. Further, to insist that any good theory *contain* the axioms of geometry in order to do such reasoning seems like a mistake, or at least an unexpected and bizarre inconvenience.

Second, the axiomatic view requires theories to be axiomatizable. Theories that can be axiomatized turn out to be rare, and theories that are actually treated as a set of axioms rarer still. This was bad enough in disciplines like biology and psychology, where it was hard to find things that counted as laws. But it seemed to be true even of physics: as van Fraassen notes, many useful treatments of quantum mechanics are non-axiomatic in form [van Fraassen, 1970]. Even if we are confident that theories *could* be identified with sets of axioms, then, it seems like a stretch to claim that the axiomatic view has captured how *scientists* use theories.

From these criticisms, an alternative naturally follows. The *semantic view* of theories claims that theories are to be identified with sets of *models*, rather than sets of sentences. These models are real structures—abstract entities like sets or state-spaces in Suppes and van Fraassen, concrete objects in more recent treatments [Giere, 1988, Godfrey-Smith, 2006]). These structures are meant to be isomorphic to the world in some respect. Though the models of the theory are often *described* using language, the important linkages hold between models and the world, not between any set of descriptions and the world. So on the semantic view, a theory consists of two parts: a set of models, and a postulation of isomorphism between certain respects of models and parts of the world.

The semantic view seems to fit better with scientific practice; many disciplines present models of some target phenomenon and then reason about them. This is most obvious in fields like cognitive psychology. Models of facial recognition, say, are never presented as sets of laws. Instead, one is presented with a model mechanism and an assertion that *this is what the brain does*—that is, an assertion that the brain is isomorphic to the model in some relevant respects. Similarly, as Lloyd has shown, many of the central claims of evolutionary theory can be interpreted as models of systems under selection [Lloyd, 1994]. Newtonian mechanics can be interpreted as the postulation of certain models, the permissible Newtonian spaces [van Fraassen, 1970]. And so on.

In addition to fitting the apparent practice of science, the semantic view also provides a neat solution to the role of mathematics in science. Mathematics is something we use to reason *about* the models. Mathematics is not a part of any theory, but is available to all. Thus, as van Fraassen puts it, physics first sets up a framework of models and then, having done so, "The theoretical reasoning of the physicist is viewed as ordinary mathematical reasoning concerning this framework" ([van Fraassen, 1970] 338).

With that in mind, return to the question of the peg and hole. The axiomatic approach had trouble with predicates like 'x is a circle.' On the one hand, it seemed there was good reason to keep them out of physics. On the other, it was obvious that physics talked about circles. This turns out to be a pseudoproblem on the semantic approach. Physics postulates certain models for the action of (say) atoms. Some of those models contain circular structures. The physicist can talk about those circular structures, deduce geometric facts about them, and derive the consequences that those geometric facts might have for the behavior of his model. The physicist can talk about circles because *anyone* can talk about circles: the language of geometry (or of set theory, or of abstraction) are resources available to every science. Geometric properties have no unique home. Puzzle solved.

4 MR and the Semantic View

That's all very quick; there are many ways to make the semantic view more precise. I want to continue by arguing that accepting *any* version of the semantic view should have serious consequences when we think about MR. Indeed, it's a bit striking that most philosophy of mind goes on as if the axiomatic theory was still widely accepted, when it is not. Perhaps part of the problem is the apparent formal equivalence of the semantic and axiomatic account: strictly speaking, any set of models can be defined as the structures that satisfy a set of axioms, and any set of axioms can be given a corresponding definition in terms of models. For this reason, it was not obvious early on that either approach to theories was superior to the other, at least on formal grounds ([van Fraassen, 1970] 326).

The formal power of the approaches aside, however, the *interpretation* of scientific theories is quite different under each of the two approaches. On the axiomatic approach, language is everything: theories must be formulated in a particular language and, as we saw, the basic names and predicates of a theory are thought to be a good guide to its ontological commitments. In contrast, on the semantic view

> ...the language used to express the theory is neither basic nor unique; the same class of structures could well be described in radically different ways, each with its own limitations. The models occupy center stage. ([van Fraassen, 1980] 44)

The language we use—in postulating our theories, working with them, and using them to explain—is a poor guide to a theory's ontological commitments.[8]

The practical effect of this is to make a convincing account of MR much harder to come by. First, you can't get MR just by definition (or, I'll assume equivalently, by pumping for intuitions about alternative realizations). Take some term P, the the definition of which makes no reference to the physical stuff that has to be in place to be a P-thing. Putnam's peg-and-hole works well here—the definition of 'square peg' does not place restrictions on *what* is square and peggy. On the axiomatic view, remember, unrestricted definition alone meant that P plausibly designates an MR property. Since many terms are analytically MR in this sense, MR is common.

This fails once we move to a semantic view of theories. On the semantic view, terms do not belong to a particular science (though, of course, some sciences may use particular terms more than others). The fact that a term is defined in a realization-independent way does not prevent its use in a science that covers potential realizers. So long as a model, or a portion of a model,

[8]That is not to say that theories aren't ontologically committal. They are. Rather, it is to say that linguistic formulations of a theory don't show you what those commitments might be. The truthmakers for claims about model-world isomorphisms need not be transparent features of the language in which theories are formulated.

satisfies the conditions for proper use of the predicate, one can talk about the model using that predicate.

Put another way: On the semantic view, the question "what, from the perspective of physics, do all square pegs have in common?" has a simple answer: they all have a square cross-section and are rigid. The same story can be told for pains, same for species, and so on. If parts of some physical models are (or are isomorphic to) squares or pains or species, then physics itself can talk about squares or pains or species. It can do so because *any* science can do so: on the semantic view, language is something that we use to talk about theories, not something that limits the domain of theories. This means that any truths about items that hold in virtue of definition, or that hold in virtue of other necessary features of the item (like those that follow from geometry, etc.) are available to any science.

A similar, slightly more subtle point can be made about 'structural' forms of MR. Putnam's idea of 'structural features' of a situation grows naturally out of the functionalism with which MR is often associated. One discovers that the explanatorily relevant features of some property (say the property of being an axon) are its functional ones—what it does, how it is hooked up—and that many and varied things can come together to impart axonhood on individual entities. As I argued above, the axiomatic view makes structure-based properties effectively MR, because, lacking the resources of higher-order quantification, individual sciences don't have the descriptive resources required to derive structural facts.

Note, however, that different instances of axons have at least one thing in common (indeed, *must* have it in common): they are isomorphic to one another.[9] They all bear a structural relationship to a cell body plus a certain capacity for transmitting action potentials down to the synapse.

This isomorphism is of little consequence on an axiomatic view, since physics does not contain the resources to form statements about it. But on the semantic view, physics *must* contain the resources to make assertions about isomorphisms. What gives physical theories empirical content is the postu-

[9] Or, more precisely, they are either isomorphic to each other (on a concretist view) or their state-spaces are isomorphic (on an abstract, state-space view).

lated isomorphisms between models and various parts of the world. Physicists must be capable of discussing isomorphisms, and there is no reason why their ability to talk about isomorphism cannot be extended to isomorphisms between models. So contrary to the axiomatic view, there is something common to all axons that physics can appeal to: their isomorphism. This trick can be repeated for most any 'structural' property you please—shapes, causal relations, or whatever.

In conclusion, the axiomatic view contains two quick routes to accounts of MR—MR by definition and MR by structure. Neither view about MR will work on the semantic view. The fact that the definition of P makes no reference to realizers, or the fact that P-ness depends on structural properties, is *not* sufficient to show that physics cannot talk about P-ness. On the contrary, there is every reason to suppose that it can. This strongly suggests that the plausibility of MR is a simple artifact of the axiomatic view of theories.

5 Conclusion

Holding [S] (or any similar gap-filling principle) is tantamount to asserting that physics doesn't care about abstract descriptions of things within its domain. But of course, physics does care about abstract descriptions. Indeed, much of the power of physics comes from its ability to abstract away from the details of individual things, and then talk about these abstractions in a way that lets us say something true about every thing. Take, say, special relativity, which allows us to say true things about about *any* object in *any* inertial reference frame—no matter its size, composition, or whatever. It is this abstraction that is at work in Gauss' law: abstraction away from particular details lets the really important features of situations (the structural feature of shape) shine through. On the axiomatic view, it is difficult to see how or why abstract re-descriptions keep you within the same science—a change of vocabulary just is a change of theory. On the semantic view, the answer is simple: abstract re-descriptions belong to the same theory just in case they are used to describe the same set of models (or a subset thereof).

What does this mean for the notion of multiple realizability? If I am correct, the following two theses cannot be jointly motivated:

MR1 Multiple realizability is explicable as a scientific notion.

MR2 Theories which refer to multiply realizable properties are, for that very reason, scientifically irreducible.

If we accept MR1, then, given that sciences have extraordinary flexibility in how they discuss the details of their domain, MR2 is false (or at least unmotivated).[10] Conversely, if we accept MR2, we will need to turn to some non-scientific explication of MR, and so abandon MR1.

Nevertheless, it might be possible to deny either thesis, and so preserve some weaker notion of MR. Both options are worth considering. Many discussions of MR implicitly treat it not as a scientific but as a metaphysical thesis, thus denying MR1.[11] Whether any such an account is compatible with MR2 is open to debate; most likely, it depends on the particular view of scientific reduction assumed. In developing such an argument, it will be critical to avoid accounts of intertheoretic reduction (including Nagel's) which themselves were developed within an axiomatic view of theories.

On the other hand, one could treat MR as an interesting scientific notion, and develop the intuitions that underly the standard accounts independently of any notion of reduction. Aizawa and Gillett's discussion of MR, for example, explicitly treats it as a problem of explicating an important notion in the special sciences ([Aizawa and Gillett, 2009] 182). If I am right, any such treatment will inevitably require denying MR2 in an unrestricted form. Or to put it another way, the intuition that such accounts try to capture rely on accepting only the weaker

(S*) There exists a legitimate explanations in a science ϕ that describes the domain of ϕ in the most specific way that ϕ allows.

[10]This point is defended at length in [Klein, 2009].

[11]See for example [Bealer, 1994]; in general, discussions of MR that lean heavily on modal metaphysics (as opposed to arguments about particular scientific terms) have this feature.

S^* is sufficient to capture the idea that some special science entities *could* be treated differently by physics. That in turn might be enough to capture the variable-composition intuition that underlies many putative examples of MR.

If I am right, however, the mere fact of variable composition will not be sufficient to establish scientific irreducibility; hence MR2 will inevitably fail. I tend to think that abandoning MR2 is too high a price to pay. But this is a matter of taste: explaining what's going on in the standard examples of MR is a useful project, and what we call the resulting relationship will not matter if we are careful.[12] But careful we must be: we will equivocate if we re-introduce this more limited notion of MR into older arguments which assume that MR properties are automatically irreducible.

Either way we go, we must keep in mind that sciences are more flexible than philosophers often make them out to be. The axiomatic view of theories limits the expressive power of science: working with a limited vocabulary and limiting yourself to what can be derived from a minimal set of axioms does make it very hard to say anything general—in physics or in any other science. That limitation is what makes MR sound like a plausible idea, for it is only when we artificially limit the expressive power of a science can we *force* it to give awkward disjoint descriptions of the things in its domain. But that means that the plausibility of MR rests on the least plausible features of the axiomatic view. All the more reason to abandon the axiomatic view, and with it the full-blooded sense of multiple realizability.[13]

[12] Thanks to Carl Gillett for convincing me on this point

[13] Thanks to Karen Bennett, Carl Gillett, Esther Klein, Tristram McPherson, Tom Polger, and an audience at a Society for the Metaphysics of Science meeting at the 2009 Central APA for helpful comments on a previous draft.

References

[Aizawa and Gillett, 2009] Aizawa, K. and Gillett, C. (2009). The (multiple) realization of psychological and other properties in the sciences. *Mind & Language*, 24(2):181–208.

[Bealer, 1994] Bealer, G. (1994). Mental properties. *The Journal of Philosophy*, 91(4):185–208.

[Bechtel and Mundale, 1999] Bechtel, W. and Mundale, J. (1999). Multiple realizability revisited: Linking cognitive and neural states. *Philosophy of Science*, 66:175–207.

[Block, 1997] Block, N. (1997). Anti-reductionism slaps back. In Tomberlin, J., editor, *Philosophical Perspectives: Mind, Causation, World*, volume 11, pages 107–133. Blackwells, Oxford.

[Bontly, 2005] Bontly, T. (2005). Proportionality, causation, and exclusion. *Philosophia*, 32(1):331–348.

[Craver, 2007] Craver, C. (2007). *Explaining the brain.* Oxford University Press, USA.

[Fodor, 1975] Fodor, J. (1975). *The Language of Thought.* Crowell, New York.

[Funkhouser, 2007] Funkhouser, E. (2007). Multiple realizability. *Philosophy Compass*, 2(2):303–315.

[Giere, 1988] Giere, R. N. (1988). *Explaining Science: A Cognitive Approach.* The University of Chicago Press, Chicago.

[Godfrey-Smith, 2006] Godfrey-Smith, P. (2006). The strategy of model-based science. *Biology and Philosophy*, 21:725–740.

[Grice, 1989] Grice, H. P. (1989). *Studies in the Way of Words.* Harvard University Press, Cambridge.

[Kim, 1992] Kim, J. (1992). Multiple realization and the metaphysics of reduction. *Philosophy and Phenomenological Research*, 52:1–26.

[Kim, 1998] Kim, J. (1998). *Mind in a Physical World*. MIT Press, Cambridge.

[Klein, 2009] Klein, C. (2009). Reduction without reductionism: A defence of nagel on connectability. *The Philosophical Quarterly*, 59(234):39–53.

[Lewis, 1970] Lewis, D. (1970). How to define theoretical terms. *The Journal of Philosophy*, 63(13):427–446.

[Lewis, 1986] Lewis, D. (1986). Causal explanation. In *Philosophical Papers*, volume 2. Oxford University Press, New York.

[Lloyd, 1994] Lloyd, E. (1994). *The structure and confirmation of evolutionary theory*. Princeton Univ Pr.

[Logothetis, 2008] Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature*, 453:869–878.

[Nagel, 1961] Nagel, E. (1961). *The Structure of Science: Problems in the Logic of Scientific Explanation*. Harcourt, Brace, and World, Inc., New York.

[Putnam, 1975] Putnam, H. (1975). Philosophy and our mental life. In *Mind, Language, and Reality*, pages 291–303. Cambridge University Press, London.

[Salmon, 1998a] Salmon, W. (1998a). *Causality and Explanation*. Oxford University Press, New York.

[Salmon, 1998b] Salmon, W. (1998b). Deductivism visited and revisited. In [Salmon, 1998a], pages 142–177.

[Sober, 1999] Sober, E. (1999). The multiple realizability argument against reduction. *Philosophy of Science*, 66:542–564.

[Suppe, 1989] Suppe, F. (1989). *The semantic conception of theories and scientific realism*. Univ of Illinois Pr.

[Suppes, 1967] Suppes, P. (1967). What is a scientific theory? In Morgenbesser, S., editor, *Philosophy of Science Today*, pages 55–67. Basic Books, New York City.

[van Fraassen, 1970] van Fraassen, B. (1970). On the extension of beth's semantics of physical theories. *Philosophy of Science*, 37(3):325–338.

[van Fraassen, 1980] van Fraassen, B. (1980). *The Scientific Image*. Oxford University Press, New York.