THE STRATIFICATION OF NATURE

by

KRISTIAN KEMTRUP

A dissertation submitted to the Graduate Faculty in Philosophy in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York 2011

This manuscript has been read and accepted for the Graduate Faculty in Philosophy in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Professor Jesse Prinz_____

Date_____ _____
Chair of Examining Committee

Professor Iakavos Vasiliou_____

Date_____ _____
Executive Officer

Professor Michael Levin_____

Professor Stephen Neale_____

Professor Barbara Montero_____

_____
Supervisory Committee

Abstract

THE STRATIFICATION OF NATURE

by

Kristian Kemtrup

Adviser: Professor Michael Levin

Herein, I suggest that contemporary nonreductive materialism, the view originated by Fodor (1974) and Putnam (1975), and traditional British emergentism, the view advocated by Alexander, Morgan, and Broad, share a commitment to the existence of higher level properties. I identify all of the arguments and evidence cited in favor of belief in higher-level properties, including evidence culled from composition, multiple realization, projectable predicates, and higher-level *ceteris paribus* laws. Finally, I argue that all of the evidence cited in favor of the existence of higher-level properties can be explained without positing higher-level properties as long as we accept some plausible assumptions about predicates and properties, most importantly that singular predicates can pick out clusters of properties and that singular predicates can pick out different properties in different objects.

Acknowledgments

First and foremost, I would like to express my gratitude to my wife, Maryam Eskandari for patiently supporting me, for having faith in me even when I didn't, and for encouraging me to write what I authentically felt I needed to write. Similarly, I am grateful to my mom for giving me the confidence to pursue life as a philosophy graduate student.

I owe a great debt to John Heil, whose work inspired me to tackle the widespread belief in higher-level properties. His thoughts on higher-level properties, multiple realization, "the picture theory of language," and his take on C.B. Martin's metaphysics have served as a sort of north star for the arguments I give herein.

I also wish to thank Barbara Montero, Jonathon Adler, Nickolas Pappas, and John Greenwood for providing helpful feedback and encouragement during the dissertation proposal stage. In particular, I would like to thank Professor Montero for suggesting the "abductive strategy."

Most importantly, I wish to thank my adviser, Michael Levin who put in untold hours giving me detailed and constructive feedback and suggestions on a litany of drafts of this dissertation and dissertation proposals. He has been patient and supportive, but he was never willing to accept anything but my best work. The process of working with him has taught me how philosophical work is really done. Professor Levin's suggestions were particularly essential in forming the material on type theory in chapter 5 and the material on composition in chapter 2.

Finally, I would like to thank my teacher and mentor Charlie Martin, who inspired me to try my hand at philosophy. I have tried to write something he would have been proud of and sincerely hope that his influence can be seen in everything I have written.

for Soren,

one third of the trio

Table of Contents

Chapter 1: The Levels Doctrine and a Semantic Alternative

I. Thesis and Introduction to the Levels Doctrine

Richard Feynman once wrote, "If our small minds, for some convenience, divide this glass of wine, this universe, into parts — physics, biology, geology, astronomy, psychology, and so on — remember that Nature does not know it."[1] (Hey and Walters, 1987, p. 313) Herein, I will argue that Feynman was right.

Many philosophers have argued for the opposite claim, i.e. that nature is stratified into a hierarchy of levels roughly corresponding to the hierarchical branches of science, e.g. physics, chemistry, biology, and psychology.[2] For example, Jerry Fodor (1974) suggests that "there are special sciences not because of the nature of our epistemic relation to the world, but because of the way the world is put together." (p.113) He (1997) also suggests that "The world, it seems, runs in parallel, at many levels of description."(p.162)

These quotations from Fodor exemplify what I call "the levels doctrine," which is, in the broadest possible terms, the belief that the world is stratified into irreducibly higher and lower levels of properties. According to the levels doctrine, the world possesses microphysical features, like shape, charge, and solidity, but it *also* possesses biological, chemical, mental, and social properties, too. More precisely, the levels doctrine holds that these higher-level properties are distinct from but dependent on their lower-level counterparts in a unique and special way.

---

[1] Charlie Martin introduced me to this quotation in a lecture on this topic over a decade ago. This quote and his stance on levels and the levels doctrine can be found in his *The Mind in Nature*. (2008) John Heil's *From an Ontological Point of View*, (2005) influenced by Martin's work, is also a major influence on my argument against levels In particular, Heil's argument deploys the sparse theory of properties and a similar set of semantic claims against the levels doctrine in a similar way.

[2] Fodor (1974) calls these branches of science "the special sciences." For the sake of simplicity, we can also include folk-descriptions about macroscopic objects and folk-psychology as special sciences as well.

Interestingly, there is some disagreement amongst proponents of the levels doctrine about exactly how many levels there are. Some proponents reserve a level for macroscopic properties[3] or even moral properties,[4] and at least one advocate, Samuel Alexander (1918), argues for the existence of a divine level of properties.[5] Some compositional materialists, like Pereboom (2002), even go so far as to argue that every composed object exists at a higher level than its parts. On Pereboom's (2002) view, nature is splintered into innumerably many levels: a new level for every new combination of parts into a more complex whole.

But no matter how you slice the levels, and no matter how you parse the view, belief in levels draws strength from a couple of intuitions which lay behind familiar claims like, "there must be more to nature than the mechanical properties of physics" and "a world without higher-level properties is too brutishly eliminative, too reductive." Before we examine the details of the levels doctrine, we must examine the content of these intuitions a little further.

It is important to note that there are actually two distinct intuitions behind these anti-reductive claims. Keats (1884) expresses one of these two intuitions in poetic form:

> …Do not all charms fly?
> At the mere touch of cold philosophy
> There was an awful rainbow once in heaven:
> We know her woof, her texture; she is given
> In the dull catalogue of common things.
> Philosophy will clip an Angel's wings,
> Conquer all mysteries by rule and line,
> Empty the haunted air and gnomed mine-
> Unweave a rainbow…[6] (lines 229-236)

[3] See Lynne Rudder-Baker (2007)
[4] See Richard Boyd (1989) and Russ Shafer-Landau (2003), especially part II, chapter 4
[5] The theologian Philip Clayton (2006) also has an updated version of this view, where God is an emergent, pantheistic being.
[6] This quote from *Lamia* (1884) is, of course, the inspiration for Richard Dawkins' (2000) *Unweaving the Rainbow: Science, Delusion, and the Appetite for Wonder*. Therein, Dawkin's does a good job dispelling

Keats is suggesting that any scientific attempt to identify the mystery and grandeur of natural phenomena with more basic, physical entities is to "conquer all mystery by rule and line," leaving the world without something of value, something worthy of reverence, which it must possess.

But the force of this intuition, the fear that something might be lost in reduction, which Keats has expressed beautifully, ought to be resisted. Keats's fear has proven to be groundless repeatedly, throughout history. It was over 300 years ago that Newton explained how prisms and small drops of water refract white light into the spectrum we call a rainbow. But rainbows are still, to this day, beautiful, awe-inspiring, and deserving of reverence. Reductive philosophy has taken none of "nature's charms."

At any rate, it is unclear what such a vague intuition could really prove. If the reductive truth about the natural world really is as bleak a picture as Keats is worried it is, that doesn't mean it isn't the truth. Unfortunately, there is no rule that says the truth can't be bleak. Indeed, we might just as well take Keats's intuition to be telling us that truth itself, not reductivism, is cold and empty, and that we may have reason to prefer views that aren't true. But even if the truth is bleak, that is not relevant here, for we are after the truth about questions regarding reduction the levels doctrine. Bleakness is a separate matter.

The second anti-reductive intuition is more substantive and is, perhaps, the driving force behind all of the arguments for the levels doctrine. The force of this second anti-reductive intuition is something like this. If we identify a putative higher-level entity such as, say, the property of being a rainbow with a base-level phenomena like the

---

the intuition expressed in this quote by Keats by explaining that the scientific picture of the world can be just as rich and wonderful as the pre-scientific picture of nature Keats favors.

refraction of light, we eliminate something that plainly exists, something that we need to posit in order to explain what we know about rainbows. In other words, if we deny the existence of a macroscopic level, a biological level, a chemical level, and so on, it might seem as if we end up with less world than we want.

But what reason do we have to conclude that what is here presenting as an intuition is really anything more than a bias against a uni-level view? Why not say that the microphysical world, the so-called base-level, is metaphysically rich? After all, it contains a host of entities, including bosons, mesons, charge and spin, all caught up in an unimaginably complex web of relational properties, and all of which are situated in the fabric of space-time. Herein, I will argue this base-level is all the world we need. Or more precisely, I want to argue that base-level, microphysical properties are sufficient to explain everything we need to explain about the natural world. Thus, perhaps this second intuition behind the levels doctrine is little more than a failure of imagination, a misguided attempt to label the rich world of the microphysical as too crude, too simple to account for what we see and what we know to be true about the world. If so, then perhaps the base-level is as much world as we could want.

Just to clarify, I do see how a particular ontology could be too eliminative. The Churchland's eliminativism[7] is a prime example. Minds exist, after all. So do believing and sensing, and therefore denying the existence of minds is an obvious mistake. I wish to argue that we can do away with the levels doctrine, i.e. with the claim that the world possesses features other than those at the so called base-level, without eliminating things that plainly do exist. In particular, by denying the levels doctrine, I do not mean to argue that mental properties do not exist; rather they *must be identical t*o microphysical

[7] See, for example, Paul Churchland (1981) and Patricia Churchland (1986)

phenomena. And since being identical to something microphysical is not the same thing as not existing at all, the argument I offer will not imply that mental properties don't exist. But explaining and arguing for all of that is part of the burden of the argument that follows.

Before we get to the argument, though, we need a brief outline of some of the details of the ontology of the levels doctrine. The first thing to note about belief in levels is how it differs from other non-reductive views like Cartesian substance dualism and Spinozistic property dualism. The levels doctrine is unique in that it claims that higher-level properties are dependant on the existence of material, physical, base-level properties in a way that these other nonreductive views do not. For example, on the levels doctrine, whether or not a higher-level mental property like being in pain is instantiated (in someone's mind) depends on whether the right lower-level neurological properties are instantiated in that person's brain. By contrast, for, say, an interactive substance dualist like Descartes, it is clear that a mental substance's existence does not depend on whether there are brain states, or brains, in existence. For Descartes, the mind and brain interact, but they are entirely independent entities, which just happen to causally interact.

This dependence relation, which is unique to the levels doctrine, is also its conceptual core. But this dependence relation can be and has been explained and described in different ways over the years by different proponents of levels. Accordingly, it goes by different names for different authors: "realization," "supervenience," "upward causation," and "emergence" are the most common. Unfortunately, it is hard to tell whether these different ways of describing the levels doctrine are just that, i.e. just different names for the same thing, or whether there are subtle but important

metaphysical differences between whatever it is that is picked out by terms like "realization" and "emergence." In the following chapters, I address this controversy, arguing that it is very difficult to pin down clear metaphysical differences between the supposedly different versions of the levels doctrine. And I conclude that even if there are such differences, the various incarnations of the levels doctrine still share a common commitment to an abundance of real properties related by a unique sort of one-way determination relation.

The various incarnations of the levels doctrine can be traced back at least to G.H. Lewes (1875) and possibly to Mill (1843).[8] But C.D. Broad (1925) is, perhaps, the most famous early proponent of higher-level properties. He describes his view of emergent levels of properties as follows:

> "At each stage in this process [the combining of simple objects into ever more complex configurations] we shall get things with new and irreducibly characteristic properties and new intra-ordinal laws, whilst there will probably remain certain complexes of all the lower orders. (Chapter 2, Section 4)

Of course, Broad's emergentism eventually faded, replaced by Nagel's (1961) reductionist framework in *The Structure of Science*, and by Feigl (1968) and Smart's (1959b) versions of the identity theory. (Moreover, the anti-metaphysical spirit of both positivism and the linguistic turn focused the attention of philosophers away from such metaphysically loaded concepts as emergence.) But levels reemerged in the 1970's with seminal papers by Fodor (1974) and Putnam (1975) centered on the concept of multiple

[8] See especially the distinction between heteropathic and homopathic causation in the chapter on "the composition of causes. (*A System of Logic*, Book 3, Chapter 6) It isn't particularly relevant here, but it is not clear whether Mill himself believed in the levels doctrine. His particular form of empiricism would've kept him from accepting anything so *oddly* metaphysical. However, Mill introduced the distinction between homopathic and heteropathic laws, which in turn influences Lewes to think there is a difference between emergents and resultants. So, in a sense, Mill is the delinquent father of the levels doctrine. In chapter 3, I will point out, briefly, that Spinoza's doctrine of parallel attributes was also surprisingly influential on early formulations of the levels doctrine. In fact, I find it odd that otherwise commendable work on the history of British emergentism by McLaughlin (1992) largely ignores the influence of Spinoza. This is particularly odd, given how often Broad and Alexander reference Spinoza as an influence.

realization. This newer view, or some version of it, now holds sway with a large number, perhaps a large majority, of contemporary philosophers.

To be more precise, there have been three waves of the levels doctrine: the early 20th century British Emergentism of Morgan (1923), Alexander (1918), and Broad (1925); nonreductive materialism, which was pioneered by Fodor and Hilary Putnam in the 70's and held sway until Kim's (1989 and 1992) causal exclusion argument; and the most recent form of the levels doctrine, which Louise Antony and Joseph Levine (1997) call "reduction with autonomy." (This most recent view is supposed to be immune to Kim's causal exclusion argument.)

Each version of the levels doctrine is designed to be better than the last, i.e. to be immune to some problem that plagued the older version or supported by some new piece of evidence. In that sense, the levels doctrine is resilient in the way a weed is resilient. When you chop off the flowering stem of a dandelion or some other weed, the plant itself survives and simply sends up a new flower. As I will explain in chapter four, the reason that the levels doctrine has been so resilient is that its opponents have attacked it by criticizing some detail of some version of the view. For example, Kim (1989 and 1992) attacks Fodor and Putnam's version of the levels doctrine by saying that the view implies that what happens at the physical level is "causally overdetermined" and in violation of the "causal closure of the physical." But this critique has not been successful; all it has done is to spawn a new form of the levels doctrine, the aforementioned third wave of levels. This new version of the levels doctrine posits novel higher-level properties that do not have their own, unique higher-level causal powers, but which are not mere epiphenomena either, and is thus immune to Kim's critique.

The resiliency of the levels doctrine speaks in its favor. Indeed, it is somewhat tempting to conclude that because belief in levels is so hard to refute, and it has had so many arguments in its favor, it is therefore likely to be true. Of course, this is not how philosophy should proceed. We shouldn't just count up how many arguments a view has had in its favor and how many arguments against a view have failed. But, at the very least, the resiliency of the levels doctrine shows us that it demands respect as a philosophical position. By analogy, any good gardener has a healthy respect for the strength of weeds, even if she thinks they don't belong in her garden. So too must we have a healthy respect for the levels doctrine, even if we think it doesn't belong in our ontology. Moreover, a gardener might also appreciate the subtle beauty of flowering weeds just as we too should be impressed by the sophistication and intuitive appeal of the various versions of the levels doctrine.

I would like to propose a different strategy. I will try to dig the levels doctrine out at its roots, to consider every argument and intuition supporting the view and subject it all to criticism. This is difficult work and it can only be done if we proceed by trying to interpret the arguments for the levels doctrine as charitably as possible.

Unfortunately, investigating all the forms of the levels doctrine is difficult, because every time the levels doctrine is reinvented, the metaphysical claims it makes become a little more hidden, i.e. each subsequent reiteration of the doctrine contains a less metaphysically bold, less clear assertion that there are real, higher-level features in the natural world. (To return to our weeds analogy, it is as if each new flower the weed sprouts stays a little closer to the ground, making it harder and harder to cut them off.) This makes it difficult to argue against the levels doctrine because doing so requires first

proving that each expression of the levels doctrine really is committed to belief in metaphysically higher-level properties and then showing why that belief is unjustified given the evidence presented. As a result, what I have written is sometimes divagating and frustratingly full of tangents. A better writer could have avoided all this. I can only ask my readers to forgive me and to remind them that the ever shifting and sophisticated ontology of the levels doctrine makes it difficult not to write in this way.

II. A Semantic Alternative to Levels of Properties

To argue against the levels doctrine, I offer a more parsimonious but explanatorily adequate alternative, which I call "the semantic view of levels," which is built up out of the sparse theory of properties, pioneered by David Armstrong, (1989) David Lewis, (1983) and C.B. Martin. (2008)[9] Put simply, the sparse theory of properties is just the claim that only some, but not all, descriptive predicates "carve nature at the joints." That is to say, on the sparse view of properties, there are properties, i.e. features of objects, but not every predicate picks out just one property in an object or the same property in all of the objects it describes. For example, we might say that being made of carbon-graphite is a property.[10] And we can say that being L-shaped is a property. And we can say being 1.5 meters long is a property. But, we don't need to say that being a hockey stick is a property in addition to these other properties too, because it may very well be that the predicate "is a hockey stick" picks out the properties of L-shaped, made of carbon

[9] The theory of sparse properties itself seems to be derived from the Lockean distinction between real and nominal essences. See Locke (1979) *An Essay Concerning Human Understanding.* (Book 3, Chapter 3, Section 15.)

[10] I do not actually think these macrophysical properties are base-level properties for reasons that will become clear in the course of the dissertation. (The base-level consists of quantum mechanical properties.) I use it here as an example of a property for heuristic purposes.

graphite, and being 1.5 meters long. The sparse theory simply states that the mere linguistic fact that we have a useful descriptive predicate "is a hockey stick" that describes many individual objects does not imply that all the objects share the single property of being a hockey stick.[11]

Just to clarify, I am intending to use the term "property" as broadly as possible. There are a number of metaphysical disputes about the nature of properties regarding which I remain officially neutral for the purposes of this dissertation. For example, it is often debated whether properties are *ante rem* universals, as Plato thought; *in rebus* universals, as Armstrong (1989) thinks; "wholly particular ways" objects are modified, as Martin (2008) thinks; or even Lewisian (1983) "natural properties." When I say "property," I mean it to be as neutral between all these conceptions as possible. Really, all I mean by "property" is a feature of a thing, in the way that painfulness is a feature of the experience of having a migraine headache. The arguments offered herein are meant to be compatible with a wide variety of different views of properties and could be revised *mutandis mutandis* to fit with any number of metaphysical systems. However, because the uni-level alternative to the levels doctrine is formed out of the sparse view of properties, the uni-level view is not compatible with the total denial of properties *tout*

[11] Again, a similar view can be found in John Heil's (2005) *From an Ontological Point of View*. The view I argue for herein is largely in agreement with Heil's, and differs only on certain details. For example, I don't deploy the truthmaker thesis against levels as Heil does. This is not because I disagree with the truthmaker thesis. Rather, I don't think we need anything as strong as the truthmaker thesis to provide a semantic alternative to the levels doctrine. We need only assume that there is a difference between the predicate and what the predicate describes, which can be the case regardless of what you think about the rather thorny issue of truth. Put a little differently, Heil focuses on the semantics of truth, whereas I focus on the semantics of reference and designation. One advantage of my view is that even if some deflationary view of truth turns out to be correct, my arguments are untouched. However, in defense of Heil, the truthmaker thesis is much less controversial than it is often supposed. C.B. Martin (2008) intended the truthmaker thesis to be a minimal thesis: there are true statements and then there are things that those sentences are about. In a way, even some of the most ardent deflationists about truth don't deny the thrust of Martin's minimal thesis since they don't deny that there is a difference between the true sentence and the referents of the words in the sentence.

court. The uni-level view states that the only properties which exist are the properties at the base-level. So, if there are no properties whatsoever, then the uni-level view will come out false.

Herein, I will not offer a fully worked out argument for the existence of properties. This is partially because the levels doctrine, which is the focus of our investigation, concedes the existence of properties; it ascribes to both higher-level and base-level properties. The topic at hand is whether there are higher-level properties in addition to base-level properties, not whether there are properties at all. Similarly, the topic at hand is not whether we can know anything about the external world or whether we can know anything whatsoever. (Notice that philosophical skepticism about everything and skepticism about the external world are also incompatible with the sparse view of properties and the levels doctrine.) For the sake of argument, we assume that there are things we can know. Addressing skepticism about the world would take us too far afield. By analogy, we can also assume that nature has features or properties and note that a robust debate about the reality of properties would take us too far afield.

However, I will say that I find the claim that there are no properties untenable simply because introspection affords us examples of properties. For example, I experience a migraine and I say to myself "I am in pain." That is to say, there is an object, i.e. my mind or my experiences, and there is a feature of my experiences, i.e. my being in pain. Moreover, when I have a migraine again, next week, I will feel *the exact same* pain then. And I can at least guess something like, "You have a migraine so you probably have *the same* pain, too." I can wonder whether the feature of experience

present in your mind is identical with the property of being in pain which is present in my mind. Pain is thus a paradigmatic example of a property.

Someone could, I suppose, accept that pains and perhaps other mental properties exist but deny that the physical, external world has properties or that we can't know about the properties of external objects. But such a denial does not establish that nominalism about all properties is true, because it doesn't deny the existence of phenomenal features of experience. If anything, accepting that the mind possesses properties, but the world does not, would establish that some kind Berkeleyian idealism is true. It would be tantamount to claiming that we can know only about our minds and their properties, but the world between your mind and my mind is either completely without any features at all, or the features the external world does possess will remain forever hidden. It is unclear whether the threat of Berkeleyan idealism constitutes a reason to reject the strong forms of nominalism which deny all properties. But it does show that it is harder than it might at first seem to deny the existence of properties, without also accepting something radically unintuitive.

Armstrong (1989) explains the sparse view nicely and provides some foreshadowing regarding how it relates to the levels doctrine:

> "There is no automatic passage from predicates (linguistic entities) to universals [or properties more generally]… I do not think that there is an infallible way of deciding what are the true universals [or properties]… those who argue to particular universals [or properties] from semantic data, from predicates to a universal [or property] corresponding to that predicate, argue in a very optimistic and unempirical manner. I call them *a priori realists.* Better, I think, is *a posteriori realism*. The best guide to just what universals there are is total science. For myself, I believe that this puts physics in a special position. There seem to be reasons (scientific, empirical, a posteriori reasons) to think that physics is *the* fundamental science. If that is correct , then such properties as mass charge, extension, duration, space-time interval, and other properties envisaged by physics may be the true monadic universals [or properties]… If this is correct, then the ordinary types –the type red, the type horse, in general the types of the manifest image of the world- will emerge as preliminary rough-and-ready classifications if reality. For the most part they are not false, but they are rough and ready." [italics in the original] (p.202)

Of course, the sort of sparse theory that Armstrong is describing can be objected to along the following lines. We say something is correctly described as being a hockey stick depending on what kind of features or properties it possesses. And therefore, if an object is rightly called a hockey stick it has *to be* a hockey stick, and if it *is* a hockey stick, then it has got to possess the feature or property of being a hockey stick. There is a sort of abductive argument here that is not entirely unreasonable. After all, how can we explain that "is a hockey stick" is a successful description of something if that something doesn't possess the property of being a hockey stick? Positing the property of being a hockey stick seems to be the best and certainly the easiest way to explain the application conditions of the predicate "is a hockey stick."

If we generalize this abductive strategy for determining the existence of properties from predicates, we are forced to accept that every predicate "is an X" describes an object in virtue of its possession of the property of being *X*, and that every object so described must share X. That is to say, if we generalize this strategy, we would end up accepting a one predicate-one property view of the semantics of predicates, and we end up with an abundance of properties: a property for every predicate[12].

And once you have an abundance of properties in your ontology, you have to explain the relationships all these properties and kinds of properties bear to one another. And that road leads to the levels doctrine, almost inevitably. After all, if there is a property of being a hockey stick in addition to the properties of being L-shaped, being made of carbon-graphite, and being 1.5 meters long, then there has to be some relationship between the property of being a hockey stick and these other properties. And

[12] Armstrong (1989) discusses something similar in his arguments regarding the "'fido'-fido" fallacy and the "argument from meaning."

the most natural way of conceiving of that relationship seems to be the levels doctrine, simply because the other alternatives don't make much sense. For example, we certainly don't want a dualism of hockey sticks and 1.5 meter long, L-shaped carbon-graphite things, nor do we want some odd hockey stick epiphenomenalism for a variety of familiar reasons. (For one thing, if being a hockey stick is a mere epiphenomenal property, then it won't make any causal difference in the world, and if it doesn't make any causal difference, then how could we ever detect, either directly or indirectly, to confirm its existence? Moreover, if we are to believe in properties of objects in the external world, we should do so because they help us explain some observed behavior or natural phenomenon. But since epiphenomenal properties, by definition, don't contribute anything causally, they don't do any work explaining why observed behaviors or natural phenomena happen.)

To be a bit more precise, perhaps I should say that the levels doctrine accepts that every pragmatically useful, descriptive, projectable predicate picks out a unique property. That is to say, it may be that some predicates might not pick out properties for reasons that have nothing to do with the sparseness of properties. For instance, the use of the predicate "is good" may be best accounted for by noncognitivism, and would thus not be a descriptive predicate at all, except superficially. If so, there is no reason to believe in a property of goodness. Moreover, we may be able to rule out non-projectable predicates, and mere Cambridge properties as not picking out properties for similar reasons. Furthermore, if a predicate doesn't truly describe any actual objects, i.e. if it isn't ever involved in true descriptions of actual objects, it also cannot pick out a property. Thus, for example, those who accept the abundant theory of properites are not committed to the

existence of the property of having a certain amount of caloric because biologists no longer believe that descriptions like "has such and such amount of caloric" are true. They are, however, committed to the existence of a unique property of being a protein which is picked out by the predicate "is a protein." At any rate, the levels doctrine is committed, rather obviously, committed to a relative abundance of properties even if it is not committed to a maximalist version of the abundant theory of properties where there is a property for *every* predicate, including non-projectable predicates and predicates that don't actually correctly describe any existent objects.

By contrast, the sparse theory of properties arises from the idea that there is no reason to generalize the abductive strategy for inferring the existence of properties from some predicates to all predicates, i.e. there is no reason to think that *all* predicates (or all pragmatically useful, projectable predicates, which are deployed in at least some true descriptions of the actual world) line up neatly with properties in a one to one relation.[13] Rather, the sparse theory of properties suggests t it is entirely plausible that some predicates pick out a multiplicity of different properties: either a cluster of properties all at once, different properties in different objects, or even different clusters of properties in different objects. Indeed, there are a variety of possible semantic relationships between descriptive terms and the features of the world they describe. There is no good reason to take a one size fits all approach which assumes that all descriptive predicates must relate to properties in the same way.

Perhaps that is too strong. There is one intuition which might tempt philosophers to accept a one-predicate-one-property view. This intuition is similar to the intuition

[13]Once more, I'm not saying anything new here. Heil (1999) makes a similar suggestion. So too does L. Shapiro, (2000) arguing that multiple realization based forms of the levels doctrine are question begging.

which I suspect lead philosophers in the early 20th century to think that the classical view of quantification –as opposed to what is now called "plural quantification[14]"- was the natural way of thinking about semantics and existential commitment. On the classical view of quantification, the noun phrase "some critics" refers to a set, and that set contains critics as members. Thus, the phrase "some critics," on the classical view, seems to commit us to the existence of sets.

Boolos (1985) finds this commitment to sets troublingly unnatural, and perhaps rightly so. It doesn't seem natural to say that the phrase "some critics" commits you to the existence of sets. If I want to punch some critics for their bad reviews, I shouldn't aim my fist at a set. I should aim at individual critics. Thus, Boolos (1985) thinks it at least seems more natural to say that the phrase "some critics," refers to individual critics, i.e. the noun phrase plural-quantifies over individual critics. And thus, on Boolos's account, the phrase "some critics" ontologically commits us only to the existence of critics, not to a set.

I remain officially neutral about the question of plural quantification, especially how it relates to interpreting second order logic and the set-theoretic paradoxes. I simply want to suggest that the classical view of quantification and the one-predicate-one-property view of predication appear to rely on similarly mistaken intuitions about semantics and existential commitment. Both views seem to fail to realize that we don't need to be immediately "familiar" with an object to refer to it, nor do we need to be familiar with a property to have a predicate that picks it out. (If my suggestion about the classical theory of quantification turns out to be false, it won't matter, since it will have

[14] See Linnebo (2003) for a nice overview.

helped us clarify our thinking and intuitions about the semantics of predicates and properties.)

Post Kripke (1980), we are now fully cognizant of the fact that it is possible for someone to refer to an object which that person has never observed, which she has little information about, or about which she has false information. (The real value of the Goedel-Schmidt thought experiment is that it shows us the extent of *how* unfamiliar we can be with an object and still refer to it.) According to the causal theory of reference, in order to successfully refer to something, we need only be connected to someone -in the right way- or connected to someone who is connected to someone, and so on, who is capable of referring to the object. In the face of this problem, semantic internalists and hybrid theorists (following Evans (1973) and Searle's (1983) discussions of problem cases for the causal theory, like the Madagascar case) have found ways to explain why, even if the reference of a term is determined by some internal psychological state, we still don't need to be fully familiar with the referent which is designated the internal thought.

For example, suppose I intend to refer to the person who Jones often refers to, who I know only as "the guy Jones mentioned." (Maybe I want to tell Jones to quit talking about the guy, because it is getting boring.) According to semantic internalists, there must be an internal, intentional state occurring in me which fixes the reference of "the guy Jones keeps mentioning." But that is a very bare thought I have about Jones's acquaintance. I know only that he is sometimes mentioned by Jones and that is all.

It is hard to call this minimal amount of information which is found in the content of the thought that is expressed in phrase "the guy Jones mentions" familiarity. I am clearly not familiar with Jones in this case. Thus, even semantic internalists can agree that

we are not always familiar or at least fully familiar with the person we are successfully referring to. And even if semantic internalism is true, that only implies that we have a thought about the person we are trying to refer to. Of course, it might be objected that semantic internalists require we be familiar with the object that we are referring to since we have to be familiar enough with it to form a thought about it.

Again, that's a very weak kind of "familiarity" at best. By that broad standard of familiarity, I am familiar with any X you can substitute in the sentence "I am unfamiliar with X." Since anything I am unfamiliar with can be substituted for X yielding a true sentence, whose terms successfully refer, then, I would be familiar with everything I am not familiar with. Thus, according to our best semantic theories, regardless of the dispute over internalism and externalism, reference without familiarity has to be possible. It is just the details of how reference to objects with which we are not familiar which need to be worked out.

Returning to the concept of plural quantification, notice that the phrase "some critics" doesn't tell us which critics. That is to say, the noun phrase successfully refers to some critics even when we don't know or aren't familiar with the actual critics it refers to. This is likely to have thrown off earlier logicians and philosophers of language if they thought familiarity is required for reference. As a result, they may have thought something like this: "Well, the phrase "some critics" has to refer to something we are familiar with, and we are not familiar with the individual critics, though we are familiar with the fact that there is a collection or a set of critics. So that must be what we are referring to." Perhaps, then, the fact that some philosophers in the classical tradition found it natural to think that the phrase "some critics" refers to a set or creates an

ontological commitment to sets is really an expression of their thought that it is natural to say we need to be familiar with that to which we refer. But *if there is no requirement that we need to be familiar with what we refer to when we use a collective referring expression like "some critics," then the need for sets or collections of critics falls away.*[15]

I argue the same is true with the semantics of predicates and properties. Suppose, for example, that we don't know what material constitutes some particular hockey stick. It could be wood, or carbon-graphite, or plastic. Of course, even if we don't know what it is made of, we can still know that it is a hockey stick. We then might be tempted to think something along the following lines. "I must be familiar with that feature of this object in virtue of which it is correct to call it a hockey stick. And I am not familiar with how it is composed. Therefore, the property which explains why it is correct to call this "a hockey stick" must be something distinct from its composition and shape, because I'm not familiar with its composition and shape."

But once we realize that we don't need to be fully familiar with the properties that are actually picked out by a predicate, then the need to posit a property of being a hockey stick falls away. Of course, that doesn't mean that the levels doctrine itself can now be rejected. If only it were that easy. As we will see, there is a lot more to say about whether we need to posit higher-level properties. Proponents of the levels doctrine have more evidence to offer. Nonetheless, there doesn't seem to be anything in what we know about the semantics of predicates that cuts against the sparse theory of properties or favors the levels doctrine. And the sparse theory of properties suggests a new way of explaining

[15] There may be other reasons to posit sets. Perhaps we need to posit sets to explain mathematics or some feature of logic. Again, I want to remain neutral on this. I only wish to point out that there's nothing semantically implausible about plural quantification or plural reference, and by analogy it is quite possible that there are plural picking out relations, i.e. a one-predicate-many-properties view,

away the abundance of properties at the core of the levels doctrine. That is to say, perhaps there are no higher-level properties, only higher level predicates which plural quantify over multiple properties.

Here is another way of putting the same point. It is standard in philosophy to think of the noun "Boolos" as picking out an individual and the noun phrase "some Cheerios" as picking out a set. However, advocates of plural reference rightly note that it seems "unnatural" to say that Boolos eats a set. Rather, he eats the individual Cheerios. Thus, it seems "natural," or at least semantically plausible, to say that the phrase "some Cheerios" commits you to the existence of the individual Cheerios that are eaten, not the set to which they belong. The upshot of plural reference is that it would seem to skirt an ontological commitment to sets. Linnebo (2003) calls this the "Ontological Innocence" of plural quantification. The value of Boolos's work is that he is shown that it is possible to explain how a noun phrase which appears to refer to a composite-set actually refers to the individuals in the set, and that this kind of reference is, as Linnebo (2003) says, "cognitively primitive" and "logically pure."

However, I argue we should apply the same sort of reasoning to predicates. It is at least possible that a single predicate which appears to pick out a conjunctive set of properties actually picks out each property individually. This is what I mean when I say it is at least possible that a single predicate "plural quantifies" over multiple properties.

A brief look at a couple of examples shows that the sparse view of properties is *prima facie* semantically plausible. We can easily form disjunctive predicates like "is red or blue." Obviously, this disjunctive predicate describes any object as long as it possesses one of two different properties, i.e. redness or blueness. If we say that a predicate picks

out a property if and only if it describes an object in virtue of the object possessing one of these properties, then clearly our disjunctive predicate *picks out* different properties in different objects. Another way of putting all of this is to say that the set of objects described by the predicate is not ontologically homogeneous in any way relevant to the predicate; the set is not unified by some common property or feature. It is a disjunctive kind or disjunctive type. (I am not here claiming to be able to define "homogeneous," just pointing out what seems an undeniable intuition. It is an intuition shared by anyone who sees something fishy in Goodman's (1983) disjunctive predicate "grue," which applies to an object x at a time t if t is before midnight Dec. 31 2010 and x is green, or t is after that date and x is blue.)

To take another example, we have a predicate "has a handle" and we have a predicate "has a blade." Suppose we begin encountering a lot of objects that are well described by both predicates. We might get tired of using both predicates every time we wanted to describe such an object. It is easy enough to imagine that in order to make things easier on ourselves we might come up with a short hand way of saying "has a handle and has a blade," e.g. something like "has a blandle." But how, then, is this shorthand conjunctive predicate "has a blandle" any different than the ordinary predicate "is a knife?"[16] If there is no relevant difference, as I suspect, i.e. if the predicate "is a knife" is something like a shorthand way of describing an object as having two different

[16] There may be some differences between the predicate "is a knife" and the simple artificial predicate "has a blandle." For example, we might be able to imagine a knife that has no handle. Or perhaps we could imagine a knife that cuts even though it has no blade, perhaps with a laser. But these hypothetical knives only show that the single predicate "is a knife" can't be identical to a simple conjunction of two predicates. They do not prove that the predicate "is a knife" isn't going to be identical to a more complex predicate, presumably one that is disjunctive, e.g. "has a blade and a handle, or has a laser cutter and a handle, and so on."

properties, then why not say that the predicate actually picks out two different properties at once?

Why can't a simple predicate like "is a knife" bear the same relationship to the property(s) it picks out as "has a handle and a blade?" Why can't "has color" pick out the same disjunction of properties as the disjunctive predicate like "is red, blue, yellow, green, brown, orange, or purple?" The point behind these rhetorical questions should be clear. There is no *prima facie* reason at all to reject the semantics of the sparse view of properties. It is a wholly plausible view for a wide range of descriptive predicates and must remain so until evidence and argument to the contrary are presented.

If anything, the idea that the extensions of all predicates (or all pragmatically useful, projectable predicates used in true descriptions of the actual world) are ontologically homogeneous sets has an air of implausibility. There is a kind of epistemic overconfidence in denying the sparse view of properties, because doing so implies that the classifications we come up with to describe the world line up one-to-one with ways the world is. It is exactly this sort of overconfidence that Feynman cautioned us against in the quotation at the very beginning of the paper. That is, even though we find it useful to divide predicates into different kinds like biology, chemistry, and physics, we should not assume that the world reflects these distinctions.

Here is another way of expressing how the levels doctrine seems epistemically overconfident. If the predicates of every branch of the sciences pick out their own unique set of properties, then it would be very hard to make sense of the fact that we have been forming more precise, "fine grained"[17] descriptions of the world as we discover new

[17] The term "fine grained" is used by Bechtel and Mundale (1999) for somewhat similar purposes. They write, "Thus, one diagnosis of what has made the multiple realizability claim as plausible as it has been is

lower-level descriptions. On the levels doctrine, if we were to learn that there is a level of description even more basal than quantum mechanics, we would not have learned that there was any vagueness or error in what we previously thought to be the most fine-grained description of the world; we would have only learned about some new properties that the world possesses in addition to all the other properties the world possesses, e.g. its quantum mechanical properties, its atomic properties, its chemical properties, its biological properties, and its psychological properties. Thus if the levels doctrine is true, it seems very hard, perhaps too hard, to discover that all these properties don't exist.

Again, I don't see this as a decisive refutation of the levels doctrine, but the epistemic overconfidence of the abundant view of properties and the problem of improved descriptions of the world is reason enough to at least cause us to investigate a little further.

We can parse the sparse theory of properties into two different semantic claims about the relationship between predicates and properties, both of which I have already hinted at in the examples above. The first claim is that a particular predicate can pick out different properties in different objects. I call this idea "semantic disjunctivism."[18] There is nothing particularly controversial about semantic disjunctivism. If we accept even a soupcon of realism about properties, i.e. if we accept that there is any difference at all

---

that researchers have employed different grains of analysis in identifying psychological states and brain states, using a coarse grain to identify psychological states and a fine grain to differentiate brain states. Having invoked different grains, it is relatively easy to make a case for multiple realization. But if the grain size is kept constant, then the claim that psychological states are in fact multiply realized looks far less plausible. One can adopt either a coarse or a fine grain, but as long as one uses a comparable grain on both the brain and mind side, the mapping between them will be correspondingly systematic. For example, one can adopt a relatively coarse grain, equating psychological states over different individuals or across species. If one employs the same grain, though, one will equate activity in brain areas across species, and one-to-one mapping is preserved (though perhaps further taxonomic refinement and/or delineation may be required). Conversely, one can adopt a very fine grain, and differentiate psychological states between individuals, or even in the same individual over time." (pp.202-203)

[18] I use semantic disjunctivism to create an alternative to multiply realized properties in Chapter 4.

between descriptive predicates and the features they describe,[19] we have agree to accept that every predicate describes objects in virtue of their possessing certain properties. But, again, we need to remember that this doesn't imply that every property picks out *the same property* in every object so described. Rather, it is quite plausible that some predicates or even most predicates describe objects in virtue of their possession of some one of a number of different properties.[20]

The second semantic claim I want to put forward is that a single predicate can pick out a cluster of properties in a single object. Returning to the example of a hockey stick above, it seems entirely plausible that the predicate "is a hockey stick" describes an object in virtue of its possessing the property of being L-shaped, the property of being 1.5 meters long, *and* the property of being made of carbon-graphite. The predicate picks out all of these properties at once, i.e. it describes an object in virtue of it is possession of all of these properties. I call this second claim "semantic conjuctivism."[21]

Adding semantic disjunctivism and conjunctivism together, we get the claim that some predicates pick out different clusters of properties in different objects. This latter claim forms the basis of the semantic view of levels, which is just the idea that there are no higher-level properties, only higher-level predicates, and higher-level predicates are, by definition, predicates that refer to different clusters of base-level properties in different objects. A little more precisely, the semantic view of levels defines "higher-level predicate" and "lower-level predicate" in the following way:

---

[19] Again, I mean properties in the broadest sense as just features of things. I do not mean to advocate for universals over tropes.

[20] My two semantic claims are heavily influenced by John Heil's (2005) rejection of what he calls "the picture theory of language" and the one predicate-one property view.

[21] I use semantic conjunctivism to criticize British emergentism in the second and third chapters.

1. A base-level predicate picks out the same unique property in all the objects it describes. In other words, the categories or kinds of objects as taxonomized by base-level predicates are metaphysically homogeneous in some interesting and relevant way. Base-level predicates line up with properties in a one-to-one relationship. The properties picked out by base-level predicates are the sparse properties, and vice versa. (I will say a little more about the relationship between base-level predicates and predicates deployed in physics below.)

2. A higher-level predicate, by definition, picks out either a cluster of properties in a single object, different unique properties in different objects, or different clusters of properties in different objects. Whether a higher-level predicate correctly describes an object is determined by whether the right lower-level predicates describe it.

Just to give a quick example, we can say that a higher-level descriptive predicate in biology, for instance "is a protein," picks out different clusters of base-level properties in different objects described by the predicate. Whereas a base-level predicate like "is a tau neutrino" picks out the same unique property of being a tau neutrino in all the objects it describes.

On the semantic view of levels, then, the only things that are at all "higher-level" are descriptive predicates. Put simply, if the semantic view of levels is true, then there is a hierarchy of different ways of describing the world, not a hierarchy of features of the world. The world, *pace* Fodor, does not "run in parallel at many levels of description." Rather, *there is a hierarchy of parallel descriptions of the one, flat world.*

One important and defining implication of the semantic view of levels is that it provides a new way for explaining the relationship between the higher-level descriptions

in the "special sciences" and the more base-level descriptions in physics. I suggest that predicates in physics describe the world not just correctly, but also, as I will explain below, *precisely.* By contrast, predicates in the special sciences describe macroscopic objects correctly describe objects, but they do so in a sort of imprecise or unrefined way.

Perhaps "imprecise" isn't exactly the right word. (It is an imprecise word.) Maybe "refined" is a better word. Regardless, I define "precise" and "refined," as technical terms, as follows. The fact that a higher-level predicate describes an object only implies that the object possesses some one of a number of different clusters of properties. That is to say, merely knowing that a particular higher-level predicate, like "is a protein," describes a particular object doesn't tell you precisely what property or properties that object possesses, or even that it has just one property. Higher-level predicates aren't refined enough to individuate single properties. Higher-level predicates are "coarse grained" predicates.

To invoke a metaphor, higher-level predicates are like fuzzy pictures with poor resolution.[22] Just as a very fuzzy picture of a person can represent any number of different persons, so too can a higher-level predicate describe an object in virtue of its possessing one of a number of different clusters of properties.

Here is an analogy that helps explain how a disjunctive predicate describes the world in an imprecise way. (We could modify the analogy to add semantic conjunctivism, but that isn't necessary.) The analogy is related to the old puzzle that a

[22] I do not mean to ascribe to what Heil (2005) calls "the picture theory of language." This is just a metaphor. Moreover, explaining how terms and concepts do or don't represent the world is a thorny issue in its own right. Fortunately, explicating the "semantic" view of levels does not require an excursion into the question of how concepts or the content of noun phrases represent the world. Both parties in the debate over levels agree that words can describe the world and that predicates can pick out properties. If it turns out that words do not refer or pick out properties, then the semantic view of levels and the levels doctrine are both false.

stopped clock is right twice a day, but a clock which is 5 minutes slow, which never indicates the exactly right time, is more correct than the stopped clock. The way the puzzle is stated makes it seem that the stopped clock is more correct than a slow one, even though that just can't be. (That is why it is a puzzle.) Now, the most obvious way of solving this puzzle is to note that a slow clock keeps time correctly but imprecisely. That is just common sense, after all. Indeed, a slow clock is still correct and very useful. And it is useful and correct, *because* it tells me that the current time is some one of a number of different times; it just doesn't tell me exactly which one of the those times is the current time, i.e. a slow clock is not a *precise* clock.

Someone might object that our slow clock registers a precise time, just a time that is off by precisely 5 minutes. So, perhaps a better analogy is one of those modern-looking analog clocks, which has had its numbers and the second-hand removed for aesthetic reasons. With these modern analog clocks, the minute hand indicates vaguely or imprecisely what the minute is. That is to say, the hands don't tell you exactly what time it is, only that the current time is some one of a range of different times. Nonetheless, numberless, analog clocks are still more useful than a stopped clock which is exactly right only twice a day.

I would like to suggest that higher-level predicates are correct but imprecise in roughly the same way as our analog, numberless clock. To refine the analogy for our purposes, suppose I look up at the clock and its hands are pointing roughly vertically to the position where the marker for 12:00 should be. (Remember the clock has no numerals.) In this case, I just don't know exactly what time the clock indicates, but the clock is still telling the time correctly. In that situation, the numberless analog clock

indicates that the time now is some one of a number of different times near 12:00. But it doesn't tell me exactly which one of those times it is. Remember, our clock puzzle is only a problem if you think that the clock has to be precisely correct to be correct at all, i.e. if it has to tell you the exact time. But that isn't how it is with clocks. We look at a normal clock which we are unsure is slow or fast, we say that as long as the clock is roughly accurate, the time is likely to be some one of a number of different times around the time indicated. Of course, before the age of digital clocks, even a clock with hands and numerals on its face was a bit vague, because you usually couldn't see the precise minute the minute hand was pointing to. And even a digital clock that displays only minutes is vague, because the time it displays is consistent with 60 different seconds. But still such clocks are correct and useful. When the clock says it is about noon, it *is* about noon, in that it is some one of a number of times near noon.

I argue that the same is true of higher-level predicates. If a higher-level predicate correctly describes an object, this tells us that the object possesses some one of a number of different properties. For example, the predicate "is a protein" describes certain molecules correctly even though different kinds of proteins are very different in terms of their microphysical properties. If a predicate is like the hands on an imprecise clock that tells you that the current time is some one of a number of times, then a predicate can tell you that an object possesses some one of a number of different micro-chemical structures, each of which is really just a cluster of microphysical properties. The imprecise-clock-imprecise-predicate analogy explains how imprecise, disjunctive, higher-level predicates can nonetheless be successful and accurate descriptions of the world

despite their imprecision, which in the case of predicates means their not denoting any one property

Moreover, the clock analogy is useful in dispelling a number of possible initial objections. For example, someone might object that the levels view can explain why we are right to describe a set of objects as being the same -i.e. because the objects described are the same in that they share a higher-level property- while the semantic view of levels cannot, because the semantic view of levels doesn't posit some ontological sameness between all the things that are being described as the same with some higher-level predicate. In other words, the semantic view I am offering states that there are predicates that correctly describe sets of objects as being the same, even though the objects don't share any relevant property that makes them the same. So, how then do we explain that we are correct to describe the objects as being the same, when they aren't really the same?

To answer this by analogy, suppose someone asks, "Why do we think of the clock as indicating the right time, when it indicates about 12:00 and it is actually 12:01?" Again, a numberless clock without a second hand may have its minute hand pointing up, but we know that this only implies that the time is some one of a number of times near 12:00. Nonetheless, the clock is perfectly correct; the current time is one of a number of times near 12:00, and that is exactly what the clock is indicating. Picking out a multiplicity of properties isn't a problem for the clock, so it isn't a problem for higher-level predicates either.

Someone might object along similar lines that some higher-level predicates describe vaguely defined sets or even impossibly large sets of objects, and these large

sets are too large and too open to be unified under one predicate. But let's think of our imprecise clock again. Because time is infinitely divisible and dense, there are an infinite number of times between even 12:00 and 12:01. At best, then, we know that any clock, even one calibrated in nanoseconds, indicates that the current time is some one time from an infinitely large set. The fact that this set is infinitely large is not a reason to say that the clock doesn't accurately indicate roughly what time it is. Again, it is not a problem for the clock. So it is not a problem for predicates either.

Furthermore, different people in different contexts associate different, though partially overlapping, sets of possible times with an analog numberless clock which has hands pointing straight up near where the marker for 12:00 should be, when they have been told that the clock "tells time reasonably well." I might look at such a clock and conclude that the current time is some time between 11:55 and 12:05. But perhaps someone else might think that the clock's hands pointing to 12:00 implies that the time is between 11:57 and 12:03. The explanation for the difference in the ranges is not mysterious. How we react to the clock depends on how we have been trained to think about clocks and about how loosely we interpret the phrase "tells time correctly." (This is a matter of empirical psychology and linguistics, or perhaps speaker meaning. Of course, there are competing accounts of all these topics. Nonetheless, there is no unsolvable philosophical puzzle here.) No matter what how we explain the fact that different people associate different sets of times with the hands of the clock pointing straight up, it is clear that the fact that different people think of different ranges of possible times picked out by the hands of the clock pointing up doesn't imply that the clock doesn't pick out some one of a number of different times. We only know that different people will associate

different ranges of possible times with the hands pointing symbolically up, and that they will do so depending on their conception of "tells time correctly," i.e. with how they were trained to react to the phrase "tells time correctly."

Once again, higher-level predicates are just like the hands on the clock. The fact that different people might associate different, though partially overlapping, sets of objects as being described by a predicate is no reason to think that the predicate doesn't pick out different properties in different objects. For example, I might think a tiny Barbie table is described by the predicate "is a table." But you might think that it isn't properly described with "is a table," because it is just a toy. To use a more grown-up example, some scientists might think that a virus falls under the category living and should be described by "is living," while others will disagree. Now, all speakers will agree that some objects belong in the set, e.g. that I belong in the set described by the predicate "is alive." So the sets overlap. But the fact that people *sometimes* disagree about what goes in the set does not imply that the there is some problem with predicates that pick out different properties in different objects. It only shows that different speakers react to words differently according to how they've been trained. Ultimately, these problems with the vagueness and the openness of certain categories are to be resolved by examining speaker meaning and meaning more generally. But no matter how these disputes about meaning are resolved, the referential semantics of the sparse theory of properties will not be threatened.

Before we move on, I should say a bit more about the semantics of base-level predicates. I have already mentioned that on the semantic view of levels, if you know that a base-level predicate describes an object, then, by the definition of "base-level

predicate," you know that the object possesses some unique property. And I have assumed that our best candidates for base-level predicates are those in use in contemporary microphysics. That is to say, on the semantic view of levels, our best contemporary physics gives us the most precise possible description of what the world is like. Or, at the very least, the predicates in physics are and always will be the best candidates we have for lining up with properties in a one to one relationship.

Obviously, there are good reasons to see contemporary physics in this way. (These same considerations drive levels proponents to associate physics with base-level, bedrock properties.) For one thing, whether a particular predicate in physics describes an object is *not determined* by whether some other predicate describes the same object. The same cannot be said of predicates at the higher-levels, like "is a protein." Whether or not something is correctly described as a protein is determined by whether it has the right sort of atomic elements arranged in the right sort of way. The predicates in physics don't have anything below them as determiners in the way biological and other special science predicates do. So, there can be no evidence that these predicates are shorthand for disjunctive or compound predicates, i.e. there can be no evidence that such a predicate picks out multiple properties or different properties in different objects.

I suppose it is always possible that our current and best physics will be revised and we will discover that there is a more basic level of description than what we have now. This is presumably what happened in the quantum mechanical revolution early in the 20th century. But the mere possibility of this sort of revolution, the mere possibility that predicates in physics will turn out to be determined by something more basic yet, only shows that we can never be entirely certain that the predicates which we think

describe the world in a maximally precise way actually do so. There might always be a more precise way of describing the world. We can never be sure. But just to be clear, this is a feature of the semantic view of levels, not a bug. After all, human inquiry is open to error. What we conclude empirically via induction can always turn out to be mistaken.

It is therefore an implication of the semantic view of levels that physics is the search for the most precise description of the features of the natural world and for the most precise description of causal laws. Or, put a little more succinctly, an implication of the semantic view of levels is that *precision is the hallmark of physics.* For example, at the macro-level we have an imprecise, vague description of how a ball travels when thrown. But a description in physics will tell us precisely how the ball will travel in a parabolic arc, precisely how far the ball will travel given the precise force it is propelled with, and the precise force of gravity pulling it down. Biology will describe the parts of a cell, e.g. the cell walls and the nucleus, but physics aims at describing the smallest parts of the cell, the electrons, protons, and other quantum mechanical bits that make up the cell in the most minute, exacting detail.

Just to avoid any confusion, it should be noted that saying that precision is the hallmark of physics is not a complete definition of "physics" or "physical" in the sense that is needed to resolve "Hempel's Dilemma," which Stoljar (2009) explains as follows:

> "…if physicalism is defined via reference to contemporary physics, then it is false — after all, who thinks that contemporary physics is complete? — but if physicalism is defined via reference to a future or ideal physics, then it is trivial — after all, who can predict what a future physics contains? Perhaps, for example, it contains even mental items. The conclusion of the dilemma is that one has no clear concept of a physical property, no concept that is clear enough to do the job that philosophers of mind want the physical to play." (Section 11.2)

It may turn out that a full fledged definition of all the criteria for what counts as the domain of physics will rule out the existence of mental properties or the ontological

building blocks of mental properties, or it may turn out that some future physical predicates will describe some purely mental feature of nature which is there at the base-level along with position, motion, charge, and spin. The semantic view of levels and the claim that the predicates deployed in physics are maximally precise, fine grained predicates is compatible with either attempt to escape from Hempel's Dilemma. The only consequence of the semantic view of levels' picture of physics is that *if* there were some predicate in physics that described mental properties, then that predicate would have to be maximally fine grained. This certainly doesn't seem to be a controversial consequence.

III. Strong and Weak Reduction

Before we finish explicating the semantic view of levels, it is important to note that the semantic view of levels is *strongly reductive*, by which I mean that, if the semantic view of levels is correct, then every property is identical to a microphysical property. In other words, the microphysical properties are all the properties there are, and in that sense, the semantic view of levels is an ontologically minimalistic view. By contrast, there are rival views which are often called "reductive" which posit, or would appear to posit, more properties than the semantic view of levels. Such *weakly reductive* views may be somewhat ontologically deflationary in that they may deny the existence of some particular kind of higher-level entity, e.g. multiply realized properties or emergent properties, but less deflationary than the semantic view of levels in that fey don't deny the existence of all higher-level properties. Remember, the matter of ontological deflation and sparse versus abundant theories of properties is relevant to the debate over levels, because (as we already discussed in section I.) an abundant view of properties, i.e. a view

which posits properties in addition to microphysical properties, leads almost inexorably to some form of the levels doctrine. Once we admit the existence of properties that are distinct from the microphysical properties, we have to posit some relation, other than identity, between the different kinds of properties and the microphysical properties, and the most natural relation is a one-way metaphysical determination relation: emergence, realization, or supervenience. Thus, weakly reductive views will always end up being mild versions of the levels doctrine, dressed up in reductive clothing.

A good example of a merely weakly reductive view is Andrew Melnyk's (2003) realization physicalism. Melnyk (2003) describes his view as follows:

> "... According to realization physicalism, then, the world contains a certain distribution of physical tokens, a certain distribution of physical property instances... these tokens realize and so *the world also contains*, a certain number of tokens of various functional types." [Italics mine] (p.27)

The word "also" in this quote is either an unfortunate bit of loose talk or a major concession to the levels doctrine and a move towards weak reductivism. By contrast, on the semantic view of levels, the token property instances of functional, special scientific, and other higher-level types *just are* physical tokens arranged into different taxonomical types. There is no "also" about what properties there are; there are just the physical or microphysical properties described by different predicates.[23]

It might be objected that this is an unfair reading of Melnyk and that Melnyk's view is as strongly reductive as the semantic view of levels. But Melnyk can't really be a strong reductivist and maintain that his view is, as he puts it, "retentive," i.e. not at all eliminative, regarding realized properties. According to Melnyk (2003), a view is retentive if and only if "it does not require denying the existence of tokens of special

[23] In chapter 4, I will discuss types and tokens in greater detail.

scientific or honorary scientific types." (p.32) Thus, it appears that an ontological view is retentive if and only if it doesn't eliminate *token-instances of properties* that aren't physical or microphysical properties. If so, then a retentive view is committed to more properties than the uni-level, semantic view, even if these extra properties are anchored by the realization relation to physical properties.

Further evidence that Melnyk's (2003) view is only weakly reductive comes out in his discussion of the "Truthmaker Intuition."(p.33) Melnyk (2003) claims that all true descriptions of the world in the higher-level sciences *are made true by* the "distribution of physical tokens given physical laws."(p.33) This sounds right as it is and it sounds like the sort of thing a strong reductivist would say. But Melnyk also states that his retentive realization physicalism "honors" the "Truthmaker Intuition" in a sort of indirect way. Melnyk explains that the "thick physical realizers," i.e. the physical token-properties in conjunction with the physical laws that govern them, *necessitate* what other token-properties will be realized. And, according to Melnyk, it is these realized, I would say higher-level, token properties that are the truthmakers for true descriptions in the higher-level sciences. So, on Melnyk's view, there is a kind of ontological middle man here. The physical things in conjunction with physical laws necessitate the existence of nonphysical but realized properties, and these non-physical properties in turn are the truthmakers for higher-level descriptions. But these higher-level entities are not identical to anything at the lower-level, even though their existence is necessitated by the existence of lower-level property tokens and physical laws.

These higher-level ontological middle-men that Melnyk deploys are higher-level entities just as much as emergent properties, and the argument I offer is directed at them as much as it is at any other higher-level property.

IV. The Abductive Strategy

Having laid out a brief outline of the levels doctrine and a uni-level semantic alternative, I can now explain how I will go about arguing against the levels doctrine. The argument that I offer in favor of the semantic view of levels is abductive. That is, I argue that the semantic view of levels explains everything that higher-level properties were designed to explain. And, rather obviously, the semantic view is, by definition, more metaphysically parsimonious than the levels doctrine. The semantic view of levels does not multiply metaphysical entities beyond necessity. So, if we accept that when we are choosing between two equally explanatorily adequate explanations we should always choose the more metaphysically parsimonious view, then the semantic view wins on grounds of parsimony.

At this point, it might be objected that I have not offered any rigorous definition of parsimony or simplicity, and that until I give criteria for what counts as a more simple ontology, my argument lacks a solid foundation. This is an important objection. I have to agree that it is not always clear whether a metaphysical view is simpler than some competitor. But my argument against the levels doctrine is really simpler -no pun intended- than that, and it doesn't require a sophisticated statement of Ockham's razor or any well defined principle of parsimony.

Rather, I argue that there is no reason at all to believe in higher-level properties and that belief in levels has arisen because of a series of unsound arguments which rest on false assumptions about semantics. In fact, I try to show how the metaphysical intuitions behind the levels doctrine, which someone might count as some sort of evidence for the view, can be dispelled when we compare the levels doctrine to the semantic view. In other words, I don't rely on some complex notion of parsimony as a theoretical virtue so much as the much simpler idea that you shouldn't believe in a metaphysical entity unless you at least have some minimally compelling reason to do so. And there is no compelling reason to believe in levels.

So, to proceed with this abductive strategy, it is necessary to compare the semantic view with the three waves of the levels doctrine described above, i.e. to look to see whether there is any way in which the levels doctrine can gain a leg up. The dissertation is organized accordingly. I go through all of the supposed evidence for the existence of irreducible levels of properties contained in the different versions of the levels doctrine, and I show that all of this evidence can be better explained on the semantic view of levels.

V. Chapter Outline

The argument that runs throughout the dissertation proceeds chronologically, following the various incarnations of the levels doctrine.

In chapter two, I introduce the basic concept of emergence, and explain that it is a particular form of the levels doctrine, which relies on an argument centered on the metaphysics of composition. To deal with this argument, I attempt to explain the

phenomena of composition without positing emergent properties. In doing so, I explicate semantic conjunctivism, which is just the idea that a single predicate which appears to pick out a higher-level property belonging to a composite object actually picks out a cluster of properties belonging to the parts of the object and the relational properties that bind the parts together. This view of composition is similar to, but distinct from, Merricks' (2001) view of composition. I also suggest that the mere fact that composite objects appear to possess novel properties can be explained without positing emergent, higher-level properties. And to do that, I put forward two models of how relational properties bind parts together into composite objects. I then explain how on either model we can explain the existence of novel properties without positing emergent properties.

In chapter three, I dig into the more sophisticated arguments and metaphysical claims of the British emergentists, Samuel Alexander, C.L. Morgan, and C.D. Broad in particular. I explain that Morgan and Alexander disagree about what emerges, i.e. about the exact nature of emergent higher-level properties. Morgan believes in emergent relations and Alexander believes in emergent qualities. I argue that Morgan's view of "emergent relations" collapses into either a uni-level view or Alexander's view. Emergent relations appear to be *prima facie* indistinguishable from microphysical, base-level relations, and if that is so, then what Morgan (1923) calls "emergent relations" are in no way emergent at all. Indeed, they're just base-level relational properties in emergent clothing. The only way for Morgan to save his view from a collapse into reductivism, then, is to posit that emergent relations differ from microphysical relations with respect to their qualities. But, if that is the case, then Morgan's view is just Alexander's. The point of all of this is to show that emergent properties are ontological extras, which we should

only believe in if we have some special reason to do so. I then show that what I call “the prediction argument,” is the only reason given to accept emergent properties. But the prediction argument isn't sound. It fails to recognize that there are *a posteriori* identity statements like “heat is mean kinetic energy,” These theoretical identifications serve to explain why we can’t make certain predictions *a priori*. Thus, the *a posteriori* discovery of these theoretical identities or bridge laws should not be taken as evidence for emergence. I then consider some possible objections to this attack on the prediction argument. I conclude that the prediction argument is based on a very flawed and misleading intuition that a difference in how we conceive in things, i.e. a difference in Fregean sense, implies a real ontological difference in those things.

In chapter four, I discuss Fodor and Putnam’s nonreductive materialism. I begin by explaining how the realization relation posited by Fodor and Putnam is not clearly metaphysically different from emergence, despite the fact that there seems to be some supposed difference. Moreover, I argue that nonreductive materialism has to be construed as a metaphysical theory that is committed to the existence of higher-level, multiply realizable properties. Next, I show how nonreductive materialists, Fodor in particular, obfuscate about the metaphysics of higher-level properties by oscillating back and forth between talking about “predicates” and “properties.” I then explain how Kim’s causal exclusion argument fails to knock down the levels doctrine because it relies on a strong reading of the claim that properties must come packaged with causal powers, and because a weaker reading which won’t serve Kim’s purposes is just as warranted. I round out this chapter by pointing out that semantic disjunctivism may allow us to explain away multiply realizable properties.

In chapter 5, I consider three arguments involving the nature of higher-level laws and projectable predicates designed to supplement the argument from MR properties. In each case, I will show that we can explain the putative evidence for higher-level properties without higher-level properties. To that end, I provide my own account of laws and projectable predicates, and I offer an informal type theory which explains why we don't need to posit disjunctive properties picked out by disjunctive predicates that are distinct from any of the properties referred to by the disjunct-predicates.

\

Chapter 2: Emergence, Composition, and Semantic Conjunctivism

I. Overview

In this chapter, I explain that belief in emergent properties is a particular form of the levels doctrine that is partially rooted in the concept of composition. I argue that there is no sound argument for emergent properties arising from evidence culled from the concept of composition. We can explain everything we need to explain about composite objects, while positing only the existence of the base-level properties of the parts and the base-level relational properties that bind the parts into a composite object. And to do so, all we need to do is make the very plausible semantic assumption that a single predicate can pick out multiple properties in a particular object. At the end of the chapter, I discuss the existence of novel properties in composite objects. I show how we can explain the existence of some novel properties and the apparent existence of other novel properties without positing emergent properties. My argument involves two different models of how relations bind the parts of a composite object. On one model of relations, there are relational properties that are entirely novel in composite objects, i.e. there are some relational properties that aren't instantiated until a particular composite object comes into being. However, I will show that we can make do with base-level, microphysical relations; there needn't be anything higher-level or emergent about the relations that bind a composite object together. On the other model of relations, the relational properties that bind an object are not novel; they only appear to be novel.

The ultimate goal of this chapter is to show that if there are emergent properties, then they are otiose ontological extras, and they are not necessary to explain composition. This sets up the decisive point I make in the next chapter, where I argue that the

prediction argument, which is the argument the British emergentists rely on to prove the existence of emergent properties, is also unsound.

II. A Sketch of Emergence

Let's start with an overview of the concept of emergence. As we will see, the British emergentists disagree about how to formulate their view, but for our purposes, the following quote from Broad (1925), comparing a world with emergence to a world without it serves to define emergence:

> On the emergent theory we have to reconcile ourselves to much less unity in the external world and a much less intimate connexion between the various sciences. At best the external world and the various sciences that deal with it will form a kind of hierarchy… we should have to recognise aggregates of various orders. And there would be two fundamentally different types of law, which might be called "intra-ordinal" and "trans-ordinal" respectively. A trans-ordinal law would be one which connects the properties of aggregates of adjacent orders. A and B would be adjacent, and in ascending order, if every aggregate of order B is composed of aggregates of order A, and if it has certain properties which no aggregate of order A possesses and which cannot be deduced from the A-properties and the structure of the B-complex by any law of composition which has manifested itself at lower-levels. An intra-ordinal law would be one which connects the properties of aggregates of the same order. A trans-ordinal law would be a statement of the irreducible fact that an aggregate composed of aggregates of the next lower order in such and such proportions and arrangements has such and such characteristic and non-deducible properties. (Chapter II, Section 2d)

In the simplest terms, what Broad is suggesting here is that when the right "aggregation" of properties are instantiated in the parts of an object, a novel emergent property, which is not identical to any of the old properties in the arrangement nor their mereological sum, comes into existence and is instantiated in the composite object.[24] For example, on Broad's view, when proteins and amino acids are combined into the right sort of aggregate, the composite object now possesses not only the older, more basal chemical properties but also the new property of being alive. And this latter property is,

[24] Of course, "the right arrangement" is not defined any further. This is intentional. The emergentists all argued that we cannot know ahead of time, *a priori*, which arrangements will result in emergent properties and which won't. Rather, we must empirically discover the laws of emergence just as we empirically discover the laws of causation.

rather obviously, something that no individual protein or amino acid possesses. And so it would seem that this latter property only comes into being, i.e. it only emerges, when the composite object is formed.

The advantage of emergence, ontologically speaking, is that it purports to explain the relationship between higher-level properties (the emergentists were concerned with mental and biological properties in particular) and base-level, physical properties, i.e. what Broad (1925) calls "mechanical" properties. Thus, the concept of emergence is an attempt to root mental, biological, and other higher-level features of the world in a physical world without eliminating either. Such a view is coherent and has strong intuitive appeal.

Emergentism is also very clearly meant to be more amenable to a scientific picture of the world than other non-eliminative ontologies like Cartesian substance dualism or the belief in biological entelechies.[25] All of the emergentists were clearly aware that substance dualism was doomed given the well worn causal interaction problem and given the absence of empirical evidence for the existence of either Cartesian souls or living entelechies. Moreover, the emergentists must have been aware that there was an ever growing body of research indicating that changes in the mind could be affected by changes at the neurological level, suggesting that what happened at the level of mind was determined by what happened at the lower-level.

So, the emergentists didn't want to eliminate mental properties. That porridge was too cold. But they didn't want to posit mental and biological entities that were wholly distinct and separate from the physical word either. That porridge was too hot. The

[25] See, for instance, Broad's remarks in *Mind and Its Place in Nature* on Hans Driesch and vitalism more generally. (1925, Chapter 2 Section 2)

concept of emergence offered a solution to this dilemma. It allowed that there could be a kind of ontological middle ground between substance dualism on the one hand and eliminative materialism on the other. Emergence, then, is meant to be a sort of metaphysical Goldilocks doctrine. Interestingly, as I will mention again in chapter 4, this same desire for a Goldilocks doctrine, this search for a metaphysical porridge that is neither too eliminative nor too ontologically robust, repeats itself again with contemporary nonreductive materialism, a doctrine which is supposed to fall in the conceptual ground between emergence and eliminativism.

It is important to note that emergentism avoids being too eliminative by allowing that a higher-level property can be determined to exist by the presence of a lower-level property(s), but insisting that the higher-level property is not identical to the lower-level property(s) which determines its existence. What this means is that *emergence must be a kind of determination relation*; when the right component parts are present, that *determines* the coming into being of a new property.

It is also very clear that emergence is supposed to be a very special, unique kind of determination relation, different from other more familiar kinds of determination. For example, we know that identity is a determination relation. The fact that someone is Mark Twain *determines* that said person is also Samuel Clemens, and vice versa. Or rather, the fact that someone is Mark Twain determines what properties he will have: all and only the properties of Mark Twain, and the fact that a person is Samuel Clemens determines that he will have exactly the same properties as Mark Twain.

Thus, even though emergence is a determination relation like identity, it can’t be the same relation as identity. This point is obvious enough, but it needs to be reiterated,

because some emergentists make the relation of a property to the properties it emerges from so intimate that it is hard to distinguish it from identity.[26] If an emergent property were identical to the properties it emerged from, then there would be no metaphysical difference between the emergent property and the base it emerged from, and emergentism would be ontologically indistinguishable from a view that denied emergence. The levels would collapse. Again, this is obvious since the whole point of emergentism is to claim that there are some properties, particularly those of living things and those belonging to the mind, that *aren't identical* to microphysical, base-level properties nor their mereological sums.[27]

Moreover, rather obviously, identity is a symmetrical relation. By contrast, and this brings out an important point, emergence is asymmetrical. With emergence, once the lower-level properties are present, then the higher-level property is determined to exist, but the converse is not true. Now, there are two very different reasons we might suspect that this relationship is asymmetric. First, although an emergent property can't exist without a lower-level base from which it emerged, there is no explicit requirement that an emergent property always has to emerge from the same lower-level base properties. For example, a putative, emergent higher-level property, say being an amoeba, might emerge from different microphysical base properties on different occasions, e.g. one amoeba might be made of different chains of molecules than another. To use a term that postdates

---

[27] Unfortunately, the emergentists, Alexander (1918) in particular, sometimes seem to forget that emergence can't be the same thing as identity. Or, at the very least, they're guilty of loose talk about identity. This makes some of their claims about the mind-body problem exceedingly hard to interpret. For example, in *Space, Time, and Deity*, Alexander (1918) states: "We thus become aware, partly by experience, partly by reflection, that a process with the distinctive quality of mind or consciousness is in the same place and time with a neural process that is with a highly differentiated and complex process of our living body. We are forced, therefore, to go beyond the mere correlation of the mental with these neural processes *and to identify them.*"[italics mine] (Book III, Chapter 1, Section A)

British emergentism, it certainly seems that emergent properties are "multiply realizable." (I say "it seems" that way because -and perhaps this is surprising- I can find no convincing, explicit confirmation in the writings of the emergentists that they believed emergent properties were multiply realizable. This shows us a possible difference between emergence and contemporary nonreductive materialism. The former bases its arguments for the existence of levels in composition, while the latter bases its arguments for levels in multiple realizability.)

The second reason why we should think of emergence as an asymmetric relation is best understood by noting that, according to emergentism, an object possesses a higher-level property only because it possesses the right aggregation of lower-level properties, but not vice versa. For example, emergentism suggests that the reason I am conscious is that processes in my brain have certain properties, but emergentism doesn't imply that the processes in my brain have those neuro-chemical properties because I am conscious. In other words, the presence of an emergent property is explained by the presence of level properties, but the presence of the lower-level properties is not explained by the presence of the higher-level property. Thus, emergence is certainly asymmetric as an explanatory relation.

Morgan (1923) attempts to explain the asymmetry of emergence by asserting that an emergent property "involves" the lower-level properties it emerges from, but lower-level properties don't "involve" the property that emerges out of them. (Lecture I, Section iii) For Morgan, "involves" clearly means something like "contains," in the way that a whole or composite entity contains its parts. What this shows is that involvement is an asymmetric relationship in the same way that composition is asymmetric. In other words,

according to Morgan, a composite *whole involves its parts but the parts don't involve the whole*, and thus there is an asymmetrical relation between emergent properties and what they "involve."

Summing up, emergence must be a one-way or asymmetric determination relation where, when the right aggregation of lower-level base entities are brought together, the existence of an emergent property is determined; but the emergent property does not determine that the lower-level base properties exist or are aggregated. Indeed, one way to explain the difference between the properties of an aggregate and the properties of its constituents is that the aggregate could not exist unless the constituents did, but the constituents can be dispersed and continue to exist although the aggregate does not. So this sort of ontological dependence is asymmetrical.

Here we reach a problem. We already know about one sort of asymmetric determination relations: plain old causal connections. With causation, the cause determines the effect, but the effect doesn't determine the cause. For example, drinking too much will damage your liver, but damage to your liver will not cause you to drink too much. Causation flows in one direction and not the other. Perhaps, then, emergence is conceptually indistinguishable from causation. Certainly, it is beginning to look like emergence is not a special, unique kind of determination. Rather, it looks like just another name for causation.

Indeed, there seems to be some confusion amongst the emergentists about how and to what extent emergence is supposed to be distinct from ordinary intra-level causation. However, there are a few points that might help illuminate the question. For one thing, emergence is a synchronous relationship, i.e. the higher-level property exists at

the same time as the lower-level properties it emerges from. Causation, on the other hand, is (at least usually) diachronous, since the effect occurs at a different time than the cause. For example, striking a match causes a flame, and the striking of the match occurs before the flame comes into existence. By contrast, when you've got an emergent property, the composite parts are supposed to remain in existence in the composite, i.e. they are still -to use Morgan's (1923) phrase- "involved" in the emergent property. The properties of the composite parts aren't burned up, so to speak, in the act of forming an emergent property. Again, for the purposes of illustration, think of the case of striking the match, where the striking of the match does go out of existence, is burned up, at or around the time the flame comes into existence. (I don't mean to confuse matters any further by positing the existence of events as metaphysical entities over and above properties, e.g. the event of "striking the match." In fact, I have begun to vacillate between taking about cause and effect in terms of properties and in terms of events. More consistent terminology is possible, but it would make this passage unnecessarily stilted, given our purposes. In general, statements like "the event of striking the match ceases to exist before the event of the match burning comes into existence" can be restated in the language of properties by saying something like, "the features of the match and the causal power properties of the match which explain the occurrence of the flame cease to exist before the match is a blaze.)

At any rate, this attempt at a temporal distinction between emergence and causation doesn't really work particularly well, because there are instances of causation where the cause and the effect exist at the same time. For example, in *The Critique of Pure Reason*, Kant (1787, A203/ B248) discusses the case of a ball pressing down on a

pillow, causing a dimple to form. If you say the cause is the pressing down of the ball and the effect is the indentation in the pillow, then it certainly seems as if the effect occurs at the same time as the effect and remains for the same duration.[28]

At this point, it becomes difficult to see how emergence is supposed to be different from the kind of causation Kant cites, which is, as Kant himself admits, a special kind of causation. Perhaps there is nothing wrong with seeing emergence as a special kind of efficient causation. In fact, Broad (1925) may have implicitly accepted that emergence is a kind of causation by characterizing emergent relations between properties with terms like "trans-physical law" and "intra-ordinal law" (see chapter 2, section 1) Perhaps this is reading too much into Broad's use of the term "law," but as we know, in the simplest possible terms, a typical scientific law -e.g. Newton's laws of thermodynamics- is just that which governs what happens, causally, within a particular level. That is probably how Broad understood "law." For example, the laws of physics govern how physical properties like momentum and position will be affected by physical causes, and the laws of psychology determine how we think and feel. If ordinary scientific laws govern causality within a particular level, then, by saying that there are" trans-ordinal laws" of nature, it would certainly seem that Broad is accepting that the relations between higher and lower-level properties are just a special species of causal laws or a "cause-like," one way determination relation.[29]

[28] For a discussion of this point see Rosenberg's (2005) *Accessing Kant* (222ff)

[29] However, if emergence is just a species of causation, this creates all sorts of ontological problems, which Kim (1989 and 1992) has used against causally active, higher-level powers. Kim (1999) addresses the concept of emergence specifically in "Making Sense of Emergence." To put the point crudely, Kim is going to argue that if there are trans-ordinal causal laws of emergence, i.e. what Kim calls upward causation, and there is causation within each level, the inevitable result is a violation of the causal closure of the basal physical world, i.e. the idea that events at the physical level are wholly determined by prior events at that level. Now, I think that emergence can be immunized against Kim's attacks and that

We can sum all this up into a basic outline of the concept of emergence. No matter what else someone might think about emergence, the core concept of emergence is that an emergent property is a novel property which *is instantiated in an object when the right composition of more basic properties is also instantiated in that object*.

None of what we have said so far is a complete account of the metaphysics of emergent properties. There is substantial disagreement amongst the British emergentists about exactly what it is that emerges, and this is no small matter. For example, Morgan (1923) argues that emergent properties are novel "relatednesses." (See lecture 3, section 11 of *Emergent Evolution.*) For Alexander (1918), although there are emergent higher-level qualities, there is no such thing as higher-level, emergent causality. Broad is a little less specific, suggesting that both "properties" and "secondary qualities" are emergent. As we will see in the following chapter, we need to decide on how to conceive of emergent properties before we can evaluate whether the prediction argument is sufficient to prove their existence.

III. Composition and Emergence: Preliminaries

Having defined emergence, we may turn to some of the arguments and evidence for it. In this connection, the first thing I want argue for is that there is no evidence for emergent properties in the phenomenon of composition. We can explain everything we need to explain about composite objects without positing higher-level, emergent properties.

Alexander's view in particular survives Kim's criticisms, but only if what emerges is not something with no novel causal powers.

This is an important thing to note as there does seem to be some confusion within emergentism about whether or not we can account for the existence of composite things without emergence. For example, Broad (1925) claims that:

> Put in abstract terms the emergent theory asserts that there are certain wholes, composed (say) of constituents A, B, and C in a relation R to each other; that all wholes composed of constituents of the same kind as A, B, and C in relations of the same kind as R have certain characteristic properties; that A, B, and C are capable of occurring in other kinds of complex where the relation is not of the same kind as R; and that the characteristic properties of the whole R(A, B, C) cannot, even in theory, be deduced from the most complete knowledge of the properties of A, B, and C in isolation or in other wholes which are not of the form R(A, B, C). The mechanistic theory rejects the last clause of this assertion. (Chapter 2, Section 2, b)

To be fair to Broad, the argument he is actually making here is supposed to go hand in hand with what I call "the prediction argument," which we will discuss at the end of the next chapter. But certainly, this passage can be interpreted as suggesting that there is evidence in the nature of composition for emergent properties, for it does suggest that composite objects have something existent in them that can't be explained, or "deduced" if all that exists in such composite objects are the lower-level properties belonging to the more elementary parts, i.e. Broad's "A, B, and C."

As a side note, just to be fair to the British emergentists, Broad, Morgan, and Alexander are usually fairly careful in distinguishing emergence and composition as distinct phenomena and distinct concepts.[30] In fact, the emergentists have to accept a strong distinction between emergence and composition, because the whole concept of emergence owes a lot, at least historically, to G.H. Lewes's (1875) distinction between

[30] Proponents of the levels doctrine don't often make the claim that composition gives evidence for emergence. Derk Pereboom (2002) has argued that composition is always multiply realizable and as a result composition implies higher-level properties. But I don't think Pereboom is guilty of failing to distinguish the concepts of composition and emergence or higher-level properties and properties of composite wholes. Actually, I think that Pereboom is half right: if there are multiply realizable properties, then the properties of composite entities are going to be prime examples of multiply realizable properties. I just think that there are no multiply realizable properties. In chapter four, I argue against multiple realizability. Obviously, if those arguments are sound, Pereboom's arguments about composition being multiply realizable are moot.

emergent properties and mere "resultant" properties. That is to say, following Lewes (1875), all of the British emergentists distinguish occasions where simpler objects are combined into more complex composites and all the properties of the complex are mere "resultants" -i.e. where all the properties of the composite object already existed in the parts before the parts were combined- from occasions where some of the properties of the composite object are novel, emergent properties. And if that distinction is valid, then composition can occur even when there are no emergent properties in the composite object, and thus emergence must be distinct from composition. To put the argument another way, if composition and emergence weren't metaphysically distinct relationships, then creating composite objects would, virtually by definition, always result in emergent properties. There would be no such thing as a composite object with merely resultant properties. But the whole point of emergentism is to distinguish emergent from resultant properties. What this suggests is that the emergentists *should* recognize that the mere fact that an object is a composite object is no evidence that the object possesses higher-level, emergent properties. However, we still need to examine whether there is some evidence for the existence of higher-level properties in composite objects, partly because the emergentists sometimes suggest that they accept such an argument, and partly because someone could offer such an argument for higher-level properties.

But regardless of what Broad intended to argue, there certainly seems to be some intuitive pull behind the idea that there must be properties of whole composite objects which are not properties of any of the parts of the composite. This intuitive pull yields a crude and familiar argument as follows. A macaw parrot is blue and white. But a macaw is composed of atoms. And no atom is blue or white, since no single atom is large enough

to reflect enough light to be visible. Therefore, there must be higher-level properties, e.g. properties of the whole parrot, not found amongst the more base-level parts of the parrot. That is, the parrot possesses the property of being blue only as a result of the aggregation of properties that are instantiated by the molecular parts of the parrot. Thus, blueness must be a property instantiated in composite objects, which emerges when more elementary parts with the right properties are put together in the right way.

In order to deflate this argument by providing a uni-level account of composition, we need to get clearer on composition itself. Unfortunately, accounting for composition is a tricky thing. There are a number of deep philosophical problems involving composition, most of which we will have to sidestep, including problems regarding identity over time and "the problem of the many."[31] Fortunately, there are a few relatively uncontroversial axioms and definitions about composition and mereological summing that can be stated simply, which are sufficient for our purposes.

We can start with some simple terminology. Suppose we have a collection of objects $x_1, x_2, \ldots, x_n$. Let us suppose that there is an object x which is the mereological sum of $x_1, x_2, \ldots, x_n$. If there is such an object x, then we need a mereological summing function s which maps the n-tuple of objects $x_1, x_2, \ldots, x_n$ onto the composite object x. To say that x is a mereological sum of $x_1, x_2, \ldots, x_n$, then, we will write $x = s(x_1, x_2 \ldots, x_n)$. Furthermore, we can think of the relation that holds between the elements of a mereological sum as the summing relation S, which holds of $\langle x_1, x_2, \ldots, x_n, x \rangle$.

Now, it seems quite obvious and intuitive that there is only one sum and only one summing relation binding together any given n-tuple of objects. For example, there is only one sum of two apples, and there is only way to sum up these two apples into a sum.

[31] See Unger (1980)

So, we need an axiom stating that there is at most one sum for any given n-tuple of objects. We can state that axiom as follows:

Uniqueness Axiom (UA): $\forall x_1 \ldots \forall x_n([S(x_1, x_2, \ldots, x_n, x) \ \& \ S(x_1, x_2, \ldots, x_n, y)] \rightarrow x = y)$.

Moreover, it also seems, at least to me, that any n-tuple of objects, no matter how spatially or temporally distant and no matter how different in kind form a mereological sum.[32] And the elements of that sum are held together by a summing relation. For example, Plato and Socrates form a mereological sum. We can name that sum "The two greatest philosophers" or "Platocrates." Moreover, Socrates and the sun form a sum. We can even name that sum "Suncrates" if we like. And abstract objects can form sums too. For example we can look at a bunch of propositions in a bit of prose, say all the third sentences of Shakespeare's plays, and we can speak of their mereological sum. So it would seem that for any n-tuple of objects there is a mereological sum x, where $x=s(x_1, x_2 \ldots, x_n)$ and a summing relation S binding the elements of the sum into a whole. We can state an axiom capturing this intuition as follows:

Formation Axiom (FA): $\forall x_1 \ldots \forall x_n \exists x S(x_1, x_2, \ldots, x_n, x)$.

And when we put FA and UA together we get the claim that there is at least one and no more than one unique sum x for any n-tuple of objects. We can express this axiom as follows:

Summing Axiom (SA): $\forall x_1 \ldots \forall x_n \exists! x S(x_1, x_2, \ldots, x_n, x)$

[32] It might be objected that some n-tuples of axioms are too scattered to be considered a mereological sum and that therefore FA is too strong as stated. That may be. However, I suspect that these worries about whether some n-tuples are too scattered to count as sums are better thought of as worries about whether some n-tuples are too scattered to be parts of a composite object. (I explore the difference between a composite object and a mere sum below.) Regardless, we can accept FA as stated for the sake of simplicity.

Now, it is absolutely crucial to note that *SA does not say that every n-tuple of objects is a composite object.* It only says that every n-tuple of objects has a single mereological sum. This difference between a mere mereological sum and a composite object is at the heart of the uni-level account of composition.

A composite object is something concrete, which is held together by concrete, existent relations. For example, a car is a composite object.[33] My body is a composite object. A molecule is a composite object. When you dissolve the relations, holding together a composite object, you remove the object from existence. For example, the parts of a molecule of hydrochloric acid are held together by covalent bonds. If we were to pull those bonds out of the molecule, the composite object called "the hydrochloric acid molecule" would cease to exist.

By contrast, a mere mereological sum cannot be destroyed by dissolving its relations, since the only relation between the elements of the sum is S, and S is not the sort of thing that can be destroyed or removed. For example, let's say there is a sum *p* of my shoe *m* and your shoe *y.* To use the notation from above, *p=s(m,y).* Notice that if we change the relations holding between the shoes, e.g. the space between the shoes, it remains the same sum. Indeed, there don't seem to be any concrete, real relations binding your shoe and my shoe together as elements of the sum *p*, other than the summing relation *S*, in the sense that there is no concrete, existent relation we can subtract from the sum that would yield the destruction of the sum. Yet there is still a sum. That, then, is the ontological difference between a mereological sum and a composite object. In a composite object, there must be real, existent relations binding real, existent parts. And the summing relation itself cannot be added to or removed from sums in the way that

concrete, real relations like a covalent bond can be removed from –thereby destroying- a composite object like a water molecule.

It might be objected that abstract objects like, say, the ontological argument can be composite objects too, even though they are not concrete in the same way that a car or a molecule of hydrochloric acid is concrete and existent.[34] Of course, I don't mean to imply that abstract objects can't be composite objects. That would be untenable. Rather, I argue that abstract composite objects are bound together by what we could call "abstract object forming relations," and these relations will themselves be abstract entities, unlike the concrete relations which bind together the parts of a concrete object, e.g. the parts of a car or the atoms in a molecule of hydrochloric acid. For example, the ontological argument has propositions as parts, i.e. premises and a conclusion. But those parts are not all there is to the argument. The abstract parts are related by an abstract relation, in this case the logical structure of the argument or the fact that the premises are offered as support for the conclusion. Notice that if we remove the abstract relation between the propositions that compose the argument, i.e. the logical relation between the premises and the conclusion, the argument ceases to exist. It becomes a mere sum of some propositions. In other words, the argument is a composite abstract object, but it is more than a mere sum of a series of propositions. So, an abstract object can be a composite object, but it too will have to have abstract relations analogous to the existent concrete relations that bind concrete composite objects like a car or a hockey stick, and these abstract relations are what make the object a composite object and not just a mere sum of some propositions,

[34] I wish to remain as neutral as possible on the ontological status of abstract objects. The view uni-level view I am arguing for is meant to be as compatible with as wide a variety of metaphysical views as possible, including any number of stances on the existence of abstract objects.

An entirely different objection could be generated as follows. It seems like some composite objects can survive some losses or additions in the relations that bind them in the way that composite objects can, apparently, survive the loss of some of their parts. For example, my body is still a composite object, perhaps the same composite object, even as the cells which compose it rearrange themselves, say, as my body changes shape after too many donuts. This is, of course, one of the roots of the problem of identity over time. And it seems like the view I am offering is in danger of requiring an answer to the problem of identity over time. After all, I have claimed that when you remove or change the relations binding the parts of an object together, you destroy the object. And if that view is correct, then how will we explain how a composite object ever survive changes or loss to the relational glue that binds its parts?

To defeat this objection, we need to remember that there is a difference between removing some of the relations holding together a composite object and removing all of them, in the same way that there is a difference between removing some of the cells of my body and all of them. No one can deny that if you remove all of the relations binding together the parts of a composite object then that object ceases to exist. If I scatter all of the planks composing a ship to the four corners of the earth, the composite object called "the ship" is gone. However, that doesn't mean that the ship can't survive some changes in the relations that hold its planks together. If you move one plank of the ship's hull five centimeters to the right, thereby changing the spatial relations between that plank and its neighbors, it would seem like the ship still exists. But noticing this raises a question: how many and what sorts of changes in the relational properties binding the parts of the ship can the ship survive? I simply have no answer to this latter question. It might very well

not have an answer. What I wish to point out here is that the question "How many losses and changes of relations between its parts can the ship survive?" is different than "Can the ship survive without any relations binding its parts?" I don't know the answer to the prior question even though I am certain that the answer to the latter question is "no."

Here's a similar way of putting the same point. Think of a tree as an actual composite object composed of wood fibers. Notice that the mere sum of the fibers would still be the same even if the fibers scattered all over the Solar System. But if the fibers were scattered, the tree, as a composite object, would no longer exist. On the definition of composite objects I am arguing for, that the tree is those fibers, bound together in the right way, by the right relations. To use a phrase that Trenton Merricks (2001) favors, a baseball is not just a set of atoms, it is a set of atoms "arranged baseball-wise," and a tree is a set of fibers "arranged tree-wise."

Still, we are left with a problem: what are the "right relations?" I simply don't know the answer to that question. But here is what we can agree on. There are some concrete relations that are, as I will call them, "composite forming relations." For example, the covalent bond between hydrogen and oxygen yields a composite object. However, it might be the case -and it seems intuitive to say that it is the case- that some relations do not form composite objects. For example, Socrates and my shoe are related temporally; Socrates existed before my shoe. But it doesn't seem right to say that there is a concrete composite object composed of my shoe and Socrates. Of course, there is a mere mereological sum of Socrates and my shoe. But that sum is not a composite object.

What I wish to claim is that the mere existence of a mereological sum *x,* where $x = s(x_1, x_2, \ldots, x_n)$ does not imply the existence of a composite object c whose parts are x_1,

…, x_n. A composite object consists of parts suitably related. That is, a water molecule is not just the sum of hydrogen and oxygen; it is hydrogen and oxygen bound together by the right sort of covalent bond. Moreover, in many cases, a composite object's parts are bound together by a variety of relations. For example, a protein is composed of a variety of molecules and atoms tied together by different sorts of covalent and ionic bonds. If we want a formalized definition of a composite object, then, we need to introduce new variables referring to the relations of the parts. Let's call the n-tuple of relations binding together the parts of a composite object *c*, "$R_1, \ldots, R_m$" and we will assume in all cases that $R_1, \ldots, R_m$ belong to the class of composite forming relations. We can now define a composite object as follows:

Composite Object Definition (CD): *c* exists iff $\exists x_1, \ldots, \exists x_n[c = (s(x_1, \ldots, x_n) \& R_1(x_1, \ldots, x_n) \& \ldots, \& R_m(x_1, \ldots, x_n))]$

In slightly more formal terms, the point I wish to make here is that it follows from CD that the existence of a composite object *c*, is sufficient for the existence of a sum $s(x_1, \ldots, x_n)$. But the existence of a mereological sum doesn't imply the existence of a composite object. To simplify, let's say we are dealing with a case where there is only one composite forming relation R binding the parts of a composite object, i.e. $R(x_1, \ldots, x_n)$. We can state this formally in terms of necessary and sufficient conditions.

1. $R(x_1, \ldots, x_n)$ is sufficient but not necessary for for $\exists x S(x_1, \ldots, x_n, x)$
2. $\exists x S(x_1, \ldots, x_n, x)$ is not sufficient but is necessary for $R(x_1, \ldots, x_n)$.

IV. Semantic Conjunctivism and Composition

Now that we have some of the definitions of mereological summing and composition out of the way, we can turn to the metaphysics of composition and emergence. For emergentists, or for proponents of the levels doctrine more generally, composite objects serve as bearers of higher-level, emergent properties, while lower-level properties are confined to the more elementary parts of composite objects. Therefore, the emergence of ever higher-levels of properties will track, at least roughly, the combining of simpler parts into ever more complex, composite, higher-level objects.[35]

For those who favor a flat, uni-level universe, the relationship between properties and the objects that bare them is a little more complex, and explaining it requires a bit of semantics. Let's call the uni-level view of composite objects, "the semantic theory of composition," On the semantic theory of composition, there are no higher-level properties of composite objects; the only real properties are the properties at the base-level. These will be properties of the most elementary objects and the base-level relations or relational properties between these objects. In other words, the semantic theory of composition states that there can still be ever more mereologically complex, composite objects.[36] For example, the Eiffel tower is an object, it has properties, and can be described with certain predicates like "weighs 10,000 metric tonnes." The semantic theory of composition suggests that the Eiffel tower just is a composite object composed of atoms bound together by microphysical bonds, e.g. ionic and covalent bonds. Now, it is a result of the semantic theory of composition that the predicates which describe

[35] I say "tracks roughly" because, as Broad admits, some properties are "ordinally neutral" and some composite objects possess only lower level properties. For instance, both a building and the bricks and steel beams that compose it have mass, so possessing mass is, for Broad, an ordinally neutral property.
[36] The uni-level semantic view of levels is also logically compatible with denying the existence of composite objects. But if we deny that composite objects have higher-level properties, then we can be strong reductivists and still accept that they exist. So, a reductivist has no reason to deny the existence of composite objects.

complex composite objects like the Eiffel tower cannot pick out some higher-level property like being the Eiffel Tower. Rather, on the semantic theory of composition all of the predicates that correctly describe the composite object, in this case the Eiffel Tower, are going to pick out microphysical properties of the most elementary parts of the tower along with microphysical relational properties that bind those parts together. For example, the predicate "weighs 10,000 metric tonnes" describes the tower, but, according to the semantic theory of composition, it picks out a cluster of microphysical properties ultimately belonging to the atoms which make up the Eiffel tower, e.g. their masses and their momentum.

Therefore, the semantic theory of composition requires that some predicates which appear to pick out a single, higher-level property belonging to a single composite object, actually pick out multiple base-level properties, some of which belong to and others bind together the elementary parts of the composite object. On the semantic view of levels, the higher-level predicate conjunctively picks them out all at once. In the discussion of semantic conjunctivism in the first chapter, we determined that it is semantically plausible that a single predicate which might at first appear to pick out a single higher-level property actually picks out many base-level sparse properties all at once. Thus, the semantic theory of composition certainly seems to be a wholly plausible alternative to the emergentist account of composite objects and higher-level properties.

We can formalize this a bit as follows. Suppose a higher-level predicate like "is a statue of Lincoln, describes a composite object c. Let us say, "*Lc*" is an expression that means that the composite object *c* is well described with the macro-level predicate "is a statue of Lincoln." Imagine also that the statue is composed of molecular bits $x_1, \ldots, x_n$

which are in turn bound together by relations $R_{1...}$ R_m, which are, let's say, a series of covalent and ionic bonds. Moreover, let's suppose the molecular bits have base-level properties A_1, …, A_l like such and such a mass. (For the sake of simplicity, we will think of the properties of atoms and molecules, and not bosons and mesons, as the base-level of reality.) The semantic theory of composition defines "is a statue of Lincoln" as picking out the same properties as the following expression:

$$c(x_1, \ldots, x_n) \,\&\, A_1 x_1 \,\&\ldots,\, \&\, A_l x_n \,\&\, R_1(x_1, \ldots, x_n) \,\&\, \ldots,\, \&\, R_m(x_1, \ldots, x_n)$$

We could write this as an axiom schema along the following lines. For any higher-level predicate H which describes a composite object, that higher-level predicate picks out base-level properties of the parts of the object and the base-level relations that bind them. That would give us something like the following definition of a higher-level predicate "*H*" describing any composite object *c*:

$$Hc = c(x_1, \ldots, x_n) \,\&\, A_1 x_1 \,\&\ldots,\, \&\, A_l x_n \,\&\, R_1(x_1, \ldots, x_n) \,\&\, \ldots,\, \&\, R_m(x_1, \ldots, x_n)$$

Formalism aside, what I really want to claim is that all you need to explain the linguistic fact that some objects are correctly described with higher-level predicates are the base-level properties of the parts and their base-level relations, and the semantic claim that the *one predicate can pick out more than one property* in a particular object described by the predicate. That is to say, the picking out relation does the work of conjunction. This is why I describe semantic conjunctivism as "quasi-nominalistic." It does away with some properties, i.e. composite properties. But it doesn't eliminate microphysical, base-level properties. So it is not wholly nominalistic. It is not nominalism *tout court*

If the semantic theory of composition is at least plausible, then any attempt to argue that the fact that we describe such and such a composite object as a statue of Lincoln does not constitute a good reason to believe that there is a higher-level property of being a statue of Lincoln belonging to the composite object and not to any of it is parts nor their mereological sum. Rather, it is quite plausible that higher-level properties pick out the base-level properties of the base-level parts. (And we can explain how this works semantically formally in as great detail as we would like.) And thus, there is no evidence in the fact that we deploy higher-level descriptions of composite objects for the existence of higher-level properties of composite objects.

There are a few things that need to be said to head off some possible objections to the semantic theory of composition. First, the semantic theory of composition implies that there is no such thing as a "composite property." Or rather, if a composite property is something that exists over and above the properties of the component parts and the relational properties that holds those parts together, then there can be no such properties. On the semantic theory of composition there are only conjunctive predicates, i.e. predicates that describe composites in virtue of the parts of the composite possessing a variety of different properties.[37] I could say much the same thing about "conjunctive properties." The semantic view of composite predicates rests on the idea that there is no such thing as a conjunctive property, because a conjunctive property would have to be

[37] I should point out that this is one way my view of properties differs from Armstrong's (1983). Armstrong accepts conjunctive universals, but not disjunctive universals. I accept neither, nor do I think properties should be thought of as universals. Conjunction, disjunction, negation, and conditionals are logical or semantic operations. They are not real features of the world.

something distinct from any individual base-level property for the same reason that a conjunction is distinct from any of its conjuncts.[38]

Second, it might be objected that we can't identify being a statue of Lincoln with some specific conjunction of microphysical properties, because it is possible that two different objects can both be a statue of Lincoln, even if they don't share any of the relevant microphysical properties. This objection revolves around the problem of multiple realizability. The problem is very large, and I give it a great deal of attention in chapter four and, accordingly, I don't deal with it here. However, I do wish to note here, just so I don't seem to be ignoring the problem, that I ultimately do away with multiply realizable properties in much the same I do away with composite properties, i.e. by explaining away the apparent existence of such properties with a few minimal and, I hope, plausible claims about the semantic relationship between predicates and properties.

Now, I admit that there is a way in which this quasi-nominalistic, semantic theory of composition is controversially eliminative for some metaphysicians, i.e. those who accept the levels doctrine. But it is not at all radically eliminative, nor does it violate common sense, nor does it suggest that basic Moorean truths like "Here is a statue of Lincoln" are actually false.[39] What the semantic theory of composition suggests is that there is no property of being a statue of Lincoln and no reason to believe in such a property. But it does not imply that there are no statues of Lincoln. The linguistic fact that we have a predicate "is a statue of Lincoln," which applies to some composite objects and not their constituents, doesn't commit us to the existence of a single corresponding

[38] Here is one way in which I depart from David Armstrong (1989), who argues for conjunctive universals, but against negative and disjunctive universals. I argue against them all.

[39] The semantic view of levels denies the existence of a lot of sortal properties like being a statue or being a baseball. Again, the semantic view comes close to and is heavily influenced by Heil's (2005) views on composite objects.

property of being a statue of Lincoln that is born by the composite object so described. Instead, it at least *might* be true that the predicate "is statue of Lincoln" picks out more than one property in the objects it describes. It may well be true of an object if and only if each of those objects possesses parts which instantiate A's and are related by C's. Really, that is not controversial at all. It certainly doesn't offend against common sense.

In this way, the semantic theory of composition agrees with Merricks' (2001) view in *Objects and Persons,* but avoids a number of problems his view runs into. For example, Merricks is an elminativist about composite objects. He thinks there are no baseballs, statues, or trees; there are only atoms, some of which are arranged statue-wise or baseball-wise. By contrast, the semantic theory of composition is eliminativist regarding emergent properties and composite properties, but not about composite objects. This is a big difference. In focusing on eliminating macrophysical objects instead of macrophysical sortal properties, Merricks creates problems for his view that could have been avoided and are avoided on the semantic view of composition. For example, the semantic view of composition accepts is consistent with the claim that there are composite objects like baseballs, statues, and trees. These are objects which we can point at and describe with predicates and they shouldn't be eliminated from our ontology. The semantic view I am arguing for eliminates something different; it eliminates the sortal property of being a baseball, in virtue of which all objects which we call "baseballs" actually are baseballs. On the semantic view of composition, an object is correctly called a baseball in virtue of its possessing, or rather the parts that make it up possessing, a cluster of different properties.

Another advantage of the semantic view of composition over the view Merricks offers is that the former doesn't have to work as hard to fit with common sense as the latter. Common sense suggests, as Moore insisted, that my hand is a real thing. It is right here, after all. Therefore, Merricks has to do a great deal of work to make his denial of the existence of hands, and macroscopic composite objects more generally, fit with Moorean common sense. Either that, or Merricks has to deny Moorean common sense. In contrast, the semantic view of composition I've explicated doesn't deny the existence of hands. It just denies that all things described by the predicate "is a hand" share some unique property in virtue of which they are hands, and therefore the semantic view of composition does not conflict nearly so violently with common sense.

To see how the semantic view of composition accords with common sense, perform the following thought experiment. (You can make it a real experiment if you like.) Suppose you ask a non-philosopher, "Is your kitchen table the same as mine?" Presumably he will respond in a common sense way, "Yes, in some ways, but not others." You might ask, "How so?" A plausible response would be, "Well, yours is the same size and the same shape, but not the same material as mine." Suppose you then ask your non-philosopher friend, "But don't your table and my table have another feature in common, in addition to all of those, in that they're both tables?" I am willing to bet that your non-philosopher friend will answer with some mixture of "No" and "Huh?" He might even have the wisdom to tell you that your question goes beyond common sense and is a matter best left to philosophers. At any rate, I don't see that the property of being a table *seems to exist* because of common sense. However, that *is not to say* that a table seems not to exist either. It is just to say that there is no evidence in how things seem, i.e.

in our intuitions about the metaphysics of composite objects for the sortal property of being a table. Moreover, if you tell your non-philosophical friend that neither you nor he has a kitchen table, he would inform you, “Yes, I do have a table, I bought one at Ikea last month.” If you then say, “But you can’t have a table, and Ikea doesn’t sell tables, because all that exists are atoms arranged table-wise,” he will, if he thinks for a moment, respond that your view either conflicts with or goes beyond common sense.

Let’s examine our intuitions about macroscopic objects a little further. We can abstract away the hardness, the shape, and material the table is constructed from, i.e. the more basal properties of the constituents of the table and their relations. We can imagine removing these features from the table. After we do that, what’s left? More to the point, what does common sense and intuition tell us is left? Does common sense tell us there is a property of being a table? What is it to be a table or a particular table? Isn’t it nothing more than being a hard flat surface, with a certain shape, composed of a particular material? It seems clear that common sense suggests that tables exist but being a table is nothing more than being a hard, flat thing, or in the language of philosophers, that being a table isn’t a real property at all. *Rather, being a hard, flat surface of such and such composition are properties, and every object that happens to possess all of those properties is what we call a table.*

Someone might object in the following way. Suppose, for example, that we have a water molecule. Because water is composed of oxygen and hydrogen, the semantic view of composition implies that there is no property of being a water molecule. Rather, the semantic view suggests that there are the properties of oxygen and hydrogen and the predicate “is a water molecule” picks out those properties and the relational properties

(covalent bonds, in this case) that connect the atoms and will be true of the mereological sum. So, it would seem that the uni-level, semantic view of composition implies that there is a property of being oxygen and a property of being hydrogen. But, of course, oxygen is itself a composite object which is composed of protons, neutrons, and electrons.[40] So, if we reject the property of being a water molecule, then it seems like we have equally good reasons to reject the property of being an atom of oxygen, and the property of being a proton, and so on *ad infinitum.* Therefore, it appears that if the uni-level view of composition is true, there is nothing to stop us from sliding down a slippery slope to eliminating all properties. And that would be absurd. For example, suppose there is a property of being a proton. If protons parts have parts, and proton parts have parts, and so on *ad infinitum*, how can there be any properties at all? The worry here seems to be that the semantic view of composites drains away all properties, because every property seems to be a property of a composite object, and every object seems to be a composite object.[41]

Before we discuss this objection further, it should be noted that I am not wedded to the proposition that being oxygen is a property. I have been using examples like being an oxygen atom or being made of carbon graphite for heuristic purposes. If it turns out that there are no such properties, the semantic view of levels is in no way diminished. In fact, I would argue that being oxygen is not a real property; rather the predicate "is an oxygen atom" picks out the properties of the component parts of an oxygen atom, e.g. its

[40] Again, forget about the fact that oxygen comes in different ionic and isotopic forms, and is therefore multiply realizable. We'll get to that in chapter 4. For now, assume we're dealing with the property of being some particular ionic form of oxygen.

[41] Jonathon Schaffer (2003) raises an argument along these lines. Schaffer concludes that since you can always subdivide an object, i.e. that there is an infinite complexity going all the way down, that there can be no base-level of properties.

component protons, neutrons, and electrons, and the relational properties holding those parts together. The semantic view of compositions requires us to accept that the only real properties are the properties of the fundamental base-level, and we have learned, empirically, that the base-level of reality is best described by quantum mechanics, not 19th century atomism. We must, at some point, give up on 19th century atomism.

Unfortunately, if we have to discuss particle physics and quantum mechanics, this makes the already complex subject of higher-level properties even more complicated. Nonetheless, the semantic view of levels is not compromised if we are forced to give examples of what I argue are actual base-level properties, simple because we can give such examples, e.g. being a tau neutrino. But for the sake of simplicity, when giving examples, I will often pretend that the properties of atoms, like having atomic weight 78, are base-level properties. It is just too difficult to deal with quantum mechanical examples while also dealing with the philosophical issues at hand. So, we can pretend that there is a property of being an oxygen atom, even if that isn't strictly true.

But disavowing the property of being an oxygen atom doesn't touch the heart of this objection. Rather, it only raises a new version of the objection: on the semantic view of composition, or the semantic view of levels more generally, what properties are real? Suppose it is true that there are no objects that aren't composite objects, i.e. every object has parts. Obviously, the semantic view of levels' view of higher-order properties implies that composite objects can only have the properties that their parts have and the relational properties that bind those parts. In a way, then, the semantic theory of composition implies that the only real properties of a composite object are the properties of the parts. But, if every part is also a composite object with its own more elementary parts, then

what objects serve as the bearers of properties? And if there are no perfectly elementary parts, i.e. simple objects, then the semantic view would seem to imply that there are no properties. And again, that is absurd.

There are two responses. The first is to accept that there are simple objects, i.e. objects that do not themselves have parts. I am not really sure whether or not a neutrino has parts. Can you cut a neutrino or an electron in half? Can you even imagine cutting one in half? I certainly don't see any reason to believe that everything has parts or that everything is or even can be infinitely divisible except for the idea that since we can imagine cutting anything in half, every whole object is composed of two halves, which are the parts of the whole.

Moreover, it is unclear whether everything can be cut in half, even conceptually. For example, electrons are not extended through distinct regions of space in the way that macroscopic objects are; rather they occupy points in space. As a result, I am not sure if we can imagine cutting an electron in half. Similar things can be said for fields, waves, and forces. So I am not ready to grant the existence of a gunky world of ever more indivisible objects.

But suppose there are no simple objects, i.e. that all objects can be divided into parts. This leads to my second response. Imagine for a moment that it is true that at the base-level of reality, a particular property can belong to or be present in two objects at once. In other words, suppose that a single property can be spread through more than one object.[42] For example, let's suppose a neutrino isn't a simple object, that it has two parts: part A, which is the left half of the neutrino and part B, which is the right half. (Again, I believe that it is odd to think of dividing a neutrino into left and right parts. I am allowing

[42] I think these considerations way heavily against Schaffer's (2003) conclusions about infinite complexity.

my opponent this conceit so we can get a grip on what, after all, is his objection.) Why can't part A and B both co-instantiate one property at the same time? Can't we conceive of the property of, say, having a certain minimal charge, or the causal property of being disposed to travel at great speeds, existing in both the left and the right part of the neutrino? Extending the same line of thinking, we can say that if part A itself has two parts, say, a top part and a bottom part, then there is no reason to deny that the property possessed by A and B is possessed by those even more elementary parts.

As near as I can tell, there is no problem in conceiving properties as being spread out over multiple regions of space in this way. And if the spatial regions an object occupies can be thought of as parts of the object, then there is no problem in conceiving a property as being instantiated in multiple parts of an object all at once. And thus, to deny that one property can be instantiated in multiple objects which are part of a whole would be to deny that properties can have extension in space. And that denial is absurd. Properties can be spread throughout a region of space and can thus be possessed by all the parts, sub parts, sub-sub-parts, and so on of that region of space.

So, we can head off these worries that the semantic theory can't be true on the grounds that it would drain away all properties in a variety of ways, either by positing indivisibly simple objects or by positing properties that are instantiated in all the parts and sub parts of an object all at once.

There is more to say regarding the metaphysics of composition. But in all of the issues we might discuss, I don't see any reason to reject semantic conjunctivism or the semantic theory of composition. If anything, the discussion above suggests that the

semantic theory of composition can be made compatible with a wide variety of different positions on composition, and so it offends against none of them.

Remember, all we needed to prove was that the semantic view was plausible and that, having it in hand, we can account for composition without positing emergent properties. And I aim to show in the arguments that follow we can do without higher-level, emergent properties.

V. Novelty, Relations, and Two Models of Composition

Another argument for the existence of higher-level emergent properties, which is intimately related to the argument from composition discussed above, invokes the concept of a novel property to give evidence for higher-levels. For example, in his *Emergent Evolution* Morgan (1923) suggests that:

> "Under what I here call emergent evolution stress is laid on this incoming of the new… When carbon having certain properties combines with sulphur having other properties there is formed, not a mere mixture but a new compound, some of the properties of which are quite different from those of either component. Now the weight of the compound is an additive resultant, the sum of the weights of the components; and this could be predicted before any molecule of carbon-bisulphide had been formed. One could say in advance that if carbon and sulphur shall be found to combine in any ascertainable proportions there will be such and such weight as resultant. But sundry other properties are constitutive emergents which (it is claimed) could not be foretold in advance of any instance of such combination (Lecture 1, Section 1)

Morgan is here arguing that the existence *of novel properties* gives evidence for emergence. I think this is a slightly different argument than the argument from composition we discussed in the previous sections, and I will treat it a bit differently. This argument suggests that if there are no emergent properties in composite objects, then no feature of the composite object can be novel, i.e. all its properties must have existed before the composite object was formed. In other words, once a composite object is formed, it at least *seems* that something new has come into being. For example, once a

lump of clay is formed into a statue, then there *is* a statue. Before then, there *is no* statue; there are only molecules. So it seems that something new comes into being with the statue.

The flaw in this argument, as we have seen, is that it ignores the role of relational properties in composite objects. I grant that some composite objects may possess novel relational properties that aren't instantiated in any of the individual parts that make up the composite object, but there is no reason to conclude that these novel relational properties are anything other than base-level properties. And thus, the fact that composite objects sometimes possess novel properties does not constitute a valid argument for the claim that novel, higher-level properties come into being when we combine parts into composite wholes. Indeed, it appears Morgan has conflated relational properties with emergent properties and concluded that the need for relations implies a need for emergent properties.

Below, I will argue that relational properties needn't be emergent properties by answering the following two questions. First, how do relations serve to glue parts together into a composite whole? And second, can we explain the appearance of, or even the existence of, novel properties in composite objects without positing emergent, higher-level properties, i.e. by positing only the relational properties that bind the parts of a composite object and the properties of the parts?

To answer the first question, we need to differentiate two models of composition, which I call the "external glue" and "sticky parts" models of composition, respectively. The second question will be considered later in this section and the answer depends on the answer to the first question.

The external glue model is simple enough. It says that relations or relational properties exist independently of the things they relate. That is, the external glue model pictures a composite object as sort of union between two otherwise distinct sorts of entity: the parts with their wholly non-relational properties and the wholly relational properties that glue the parts together. To use a metaphor, this model suggests that a relation should be thought of as something that gets poured into the composite object the moment the composite object is formed.[43] On this model, even though all composite objects have parts and relations, the parts and relations can and should be conceived of as distinct entities, i.e. as a kind of external glue.

The obvious problem with the external glue model is that it is not entirely clear what meaning we can attach to the claim that "relations have an existence independent of the things they relate" After all, few relations -maybe even no relation- can be conceived of without conceiving the things they relate. (This is a problem that distinctively applies to relations, as opposed to monadic properties. The traditional problem for Platonism is how properties of any sort can exist apart from their instances. The issue I am considering takes it that the problem for relations is more acute.)[44] For example, try to conceive of one object exerting gravitational pull on another object without conceiving of the things

[43] Whether an external property exists prior to the moment of composition is a tricky question. As I will explain, I think spatial relations are a good example of external glue relations. Obviously, since the fabric of spatial relations exists prior to, say, my pencil, it would seem that the spatial relations between the parts of my pencil existed before my pencil did. However, it is entirely possible that some external glue relations are only instantiated at the time a particular composite object is created. Suppose, just for the sake of this argument, that gravitational pull is an external glue relation between two atoms. It is quite possible that the pull comes into existence only when the atoms do; i.e. the relational property of gravitational pull is not instantiated until the moment of composition.

[44] Here's what I mean. Platonic *ante rem* universals have to exist -by definition- outside of the world of *particular* physical things. They are in the realm of the forms or existing *sub species aeternitatum*. (Of course, the spatial phrase "outside of the world" is meant somewhat loosely.) By contrast, I am worried about whether wholly particular relational properties have an existence in the physical world -i.e. the world of particular objects and their particular features- that is in any way distinct from the particular categorical or qualitative features of particular objects.

doing the pulling or being pulled. When I try, I end up thinking of one small grey sphere or point pulling on another. But small imaginary spheres are things, too. Or at least you're conceiving of things when you conceive of spheres. The resultant problem is this: if we can't imagine a relation existing without the relata, then how can we attach any meaning to the claim that relations have a distinct, separate existence from the things they relate?

This conceptual problem shouldn't lead us to dismiss the external glue model so quickly. For one thing, there are some relations we do seem to be able to conceive of without their relata. Spatial relations seem to be of this sort. I can conceive of the relation of being 7 feet apart without conceiving of any two things that are seven feet apart at all. Indeed, it may be that the external glue model of relations in a composition works well for some important microphysical relations, spatial relations in particular.

So, how then do we deal with this conceptual problem for the external glue model? A defender of the external glue model can suggest that the reason we can't easily conceive of an existent relation, e.g. the relation of two atoms being covalently bonded, without conceiving their relata, is that there is a law of nature that determines that a relational property is never instantiated in the absence of the things it relates. That is, maybe it is not logically necessary that relations always exist with the things they relate, as it admittedly seems to be, but a matter of some contingent physical law. If so, maybe it is merely difficult, though not impossible, for us to conceive of relations existing without the things they relate. And perhaps it is only difficult to conceive of relata-less relations because it is hard to conceive of things we have never observed, and we have never observed, for example, a covalent bond existing without atoms that are bonded. Of

course, we do live in a world where covalent bonds aren't ever instantiated in nature without atoms that they bond. This possibility is admittedly highly speculative, but if we are going to do metaphysics, and take properties and relations seriously, the question of their instantiation-independent existence may be more of a contingent matter than a necessary one. I am not sure how seriously to take this view, but it is *prima facie* possible, and as long as it is possible, the external glue model must be taken seriously.

As an aside, I should point out that the external glue model also seems to work particularly well for explaining relational properties in quantum mechanics, like super-position, and perhaps here we have another example of a microphysical external glue relation. Again, this discussion of quantum mechanics is a bit tangential, but it might help a bit to explain exactly what I mean by external glue relations.

We are all familiar with the *metaphysical* worry arising from the fact that the individual entities described in quantum mechanics, like photons, electrons, and quarks, appear to behave in an indeterministic way that is very unlike the deterministic ways macroscopic objects behave. I suggest, with all due tentativeness, that this worry may dissolve if we think of the relations between quantum mechanical entities, e.g. the relational properties that bind individual bosons and mesons together into a larger object, as a kind of external glue that binds the quantum mechanical bits together.

For example, a baseball always has determinate motion, position, and momentum. But, as physicists have discovered empirically in the double slit experiment and elsewhere, an individual photon with a specific momentum does not have a definite, determinate position. Instead, the position of a particular photon can only be described by a Gaussian curve whose value for any point in space is interpreted as the probability of

finding the particle at that point. In other words, quantum mechanics as standardly interpreted suggests that there is a range of different positions for the individual photon. And all individual objects recognized in quantum mechanics, not just photons, display this same indeterminacy, not just in how they move about in space, but also in what sort of state they're in as well, e.g. having z-spin up.

Of course, it is tempting to say this sort of indeterminacy can't be part of the world, and that so-called quantum indeterminacy must be epistemic. We want to say that the photon has a position and that we just don't which position it has. But this epistemic solution has been rejected by quantum physicists. However, if we don't accept the epistemic solution, it seems like we are stuck with some unappetizing possibilities. One of the possibilities is to say that perhaps deterministic behavior just pops into the universe, seemingly out of nowhere according to some mysterious law of quantum-to-atomic emergence. In other words, since deterministic behavior must be a *novel* feature of atomic and macro-level composites, therefore deterministic behavior must have emerged in the composition of atomic and macro-level objects. At least, I have never quite understood why improbabilities don't accumulate and manifest themselves in the macrophysical world, so that the macrophysical world keeps drifting from what deterministic theories predict.

I suspect the philosophical worry over quantum mechanics evolves out of a particular form of the argument from novel properties that Morgan offers above. Moreover, I would also like to suggest, as a tentative conjecture, that these worries about quantum indeterminacy might dissolve for much the same reasons that Morgan's argument does. To that end, I would like to suggest that the relational properties that bind

individual quantum mechanical entities like bosons and mesons together -superposition may very well be such a relation- are a kind of external glue relation that binds sets of quantum mechanical entities and explains their indeterminate characteristics.

This solution to the quantum indeterminacy problem does require that at least one of the *relations that bind individual quantum mechanical objects together into composite wholes comes prepackaged with the causal powers that we associate with the indeterministic behavior of quantum-level objects.* If what we have said so far is correct, then there is nothing ontologically mysterious in the differences in behavior between individual quantum mechanical objects and other objects. Indeed, on my view, superposition and other quantum mechanical relations should be seen as existent properties, as much a real part of the quantum world as photons, and electrons. More precisely, I suggest that the relations between quantum mechanical objects should be seen as real in the same way physicists see spatial relations as real.

We are told that before the collapse of the wave function of a particle P upon observation, the particle's position-state is a superposition of a great many position-states. That is, there is a 25% chance that P is at point x, a 5% chance that P is at point y, a 17% chance that P is at z, and so on. What, ontologically speaking, is responsible for keeping P, in that stochastic state? Well, it maybe that there is an external glue relational property binding together the property of P-being-at-x with P-being-at-y and with P-being at z. This would be a relational property that holds the object together into an indeterminate composite. To imagine such a composite, first imagine that the property of being P is in all of x, y, and z at once. Then imagine a relation holding between being-P-at-X and being-P-at-z that determines that being P occupies places x, y, and z in an

essentially chancy way. To imagine this, we have to imagine quantum mechanical objects bound together by relations other than ordinary spatial relations, which exist along with the fabric of relational threads constituting space-time. I see no reason why this can't be the case. At any rate, all I mean to be doing here is speculating that perhaps there is a quantum mechanical glue that explains indeterminacy.

Let us move on to the sticky parts model. The core idea of the sticky parts model is that relations are not metaphysically distinct from the parts they relate, i.e. relations have no independent existence from their relata. Thus, the sticky parts model is the reverse of the external glue model. With the sticky parts model, we don't need to add anything to the component parts to glue them together; rather, the properties of the component parts serve to bind the parts together. That is, the sticky parts model pictures relations as present in the parts of a composite object, as being identical to some property of some component part.

Unfortunately, the sticky parts view has a conceptual problem not unlike the one we posed for the external glue model.[45] If relations exist, but they don't exist independently of the parts they relate, then relations must be identical, somehow, to the non-relational properties of the parts they relate. But in order to make sense of that identity between relations and non-relational, categorical properties of parts, we would need to say that the apparent difference between relations and non-relational properties is not real but exists only in our conception. How could this be?

[45] This suggests there is something like an antinomy of relational properties. It is hard to conceive relational properties as being distinct from the properties of their relata and it is hard to conceive of relational properties as being indistinct from the properties of their relata.

The best answer to this question invokes what C.B. Martin (2008), following Locke, calls "partial consideration."[46]Here's an example of partial consideration. Imagine conceiving of a triangle as being a rectilinear object with angles adding up to 180 degrees. Obviously, we can also conceive of a triangle as being a rectilinear object with three sides. Now, being a rectilinear object with angles adding up to 180 degrees is the same as being a rectilinear object with three sides. But we can conceive of being three sided and having angles that add up to 180 degrees as wholly separate and distinct things (or "things."). We can do this because we can pay attention to just the three sidedness without thinking of the 180-degree-angledness, or vice versa. Locke would call conceiving of a thing as being three sided, without paying attention to the fact that it also has angles that add up to 180 degrees, an example of partial consideration. And if partial consideration is possible, then it is possible to conceive of a relational property as being distinct from a non-relational property, even when they are identical.

Martin (2008) cites Jastrow's famous duck-rabbit picture to explain partial consideration.[47] He points out that we can think of the duck-rabbit picture as a picture of a rabbit only, or as a picture of a duck only. But the picture is both a duck picture and a rabbit picture. Ultimately, the difference between the picture being a duck picture and a rabbit picture is a difference in conception only. The fact that we have two different ways of conceiving of the picture does not imply that the picture has two different properties, i.e. being a duck picture and being a rabbit picture which are distinct in any way except in how we conceive them.

[46] See especially chapter 6, section 7.
[47] *ibid*

On the sticky parts view, then, we have to conceive of a relational property as being identical to a nonrelational property in the same way being a duck drawing can be identical to being a rabbit drawing. In other words, on this model, we should see relations and nonrelations as being distinct only in how we conceive of them, not in how they are.[48]

I think the sticky parts model works particularly well as an account of a wide variety of relations, including covalent and ionic bonds in and between molecules. Molecules are composite entities whose atoms are held together by the sharing of electrons. The covalent bond itself is a relation; it holds together the atoms in a composite molecule. But the relation can equally well be viewed as a feature of the electrons, in that the pull each electron exerts according to its charge is a non-relational property of that electron. The charge is a nonrelational property of the electron, but the charge *is* the pull. That is to say, charge can be conceived of as both a categorical, nonrelational property of each electron and as a relational property of the covalently bonded molecule.

Please note that the point I wish to make here is not that one of these models is right and the other wrong. I don't have a view about which model of composition is more correct. I prefer to think both kinds of relation exist, but the only external glue relations in the universe are spatial relations and some fundamental relations in quantum mechanics like superposition, all of which are best thought of as relations between base-level microphysical objects. At any rate, the point of this discussion is to give an account of how composite entities might be bound together by relations and how the apparently novel properties in composite objects can be explained without emergence.

[48] As a result, the sticky parts model does not deny the existence of relations, as has been done by some philosophers. See, for example, Fisk (1972).

We can turn now to our second question: whether we can explain composition without emergence. To answer this question, it is important to note that on the external glue mode, there can be base-level relational properties that come into existence which are nonetheless entirely novel. I have been at pains to show that external glue relations are properties that have never been instantiated before a particular composite object is formed, even though *these relational properties are not* higher-level, emergent properties. In other words, on the external glue model, if such and such parts have never been related in such and such a way until time t, then at time t *a novel relational property* comes into existence. The relational glue gets poured into the composite thing, so to speak, and comes into existence only in the pouring. For example, there was a first instance when two things had the *novel* relational property of being 1 mile apart, presumably some time after the big bang, but that spatial, relational property is not *emergent*; it didn't emerge out from something at a lower-level and it doesn't exist at any level other than the microphysical, base-level that all spatial relations exist at time t. Similarly, as we saw with philosophical worry about quantum mechanics, the existence of novel causal powers in quantum mechanical entities that exist in a superposition is no evidence for those novel powers being emergent; the powers may belong to the relation of being in a particular superposition. Thus, the existence of novel properties in composite objects doesn't imply the existence of emergent properties. Indeed, the novel properties which the emergentist points to as evidence for his theory may be nothing more than relational properties that have never been instantiated before.

By contrast, the sticky parts model appears to have trouble explaining the existence of novel properties which have never been instantiated before some composite

object is formed. The sticky parts model suggests that at the moment parts come together to form a composite object, the relational properties already existed before t; it is just that at t, the relations are now rearranged in such a way that it is easier to see them as relations and not just categorical properties. For example, the electrons of an oxygen atom have their charge before the oxygen atom forms a covalent bond with two hydrogen atoms. If the bond just is the charge of the electrons, the bond in some sense existed before the composition. I agree that the sticky parts explanation of what happens at the moment of composition is not easy to describe. But the idea is this: the charge of the electrons orbiting an atom is present even before the atom bonds covalently with another atom. A single electron's charge is, on the sticky parts view, the relation that holds the covalently bonded molecule together. It is, to coin an ugly phrase, a sticky-ness which exists before the composition does and remains in existence in the composition, along with the sticky-nesses of the other parts. It is a stickiness that is not yet stuck to anything. On the sticky parts model, then, *there is no novelty* in composite entities, only the appearance of novelty. And the illusory appearance of novelty in these cases is a failure to see that the relations that bind a composite object together were present even before the moment of composition. They just didn't look like relations pre-composition, because we were partially considering them as purely categorical properties.

To sum up, we can now deflate Morgan's argument from novel properties by saying that it is possible to have composite objects with novel properties, and it is possible to have the appearance of novel properties without having higher-level properties at all, because base-level relations can be novel or at least appear to be novel. Morgan is right that some composite objects may possess a truly novel relational

property, i.e. a property that has never existed until the composition was formed, But the fact these relations are novel is no evidence for their being emergent, though if they are truly novel properties, they do have to be external glue relations, because no sticky parts relation is ever novel. However, it is also possible to claim that what appear to be novel relations, on the sticky parts model, are not novel, but are merely sticky-nesses that have, so to speak, not previously been stuck to anything. So, no matter how you look at relations, you can have compositions and novel properties in your ontology without having "higher-level" properties.

It might be objected that I have here ruled out emergence *a priori*. Perhaps I have made it seem that there can never be emergent properties, even if there are novel properties. That is, it seems like I've argued as follows. If relations are the stickinesses of the parts, then nothing new comes into existence in composition. And if relations are external glue, then in some sense they are novel properties that are instantiated when composite objects come into existence, but they aren't higher-level, emergent properties. But relations have to be either external glue or sticky parts, and neither sort of relations are emergent, so it seems that relations *can't* be emergent, on either of the two models. And that conclusion is surely too strong.

This objection misses the point I am trying to make. I am arguing that relational properties are not *necessarily* higher-level, emergent properties. That is, relations don't have to be emergent. But that is not to say there couldn't be something inherent in some relational properties that makes them emergent. Indeed, as we will see in the next chapter, a good deal of the disagreement between Alexander and Morgan about what it is that emerges in emergence revolves around whether relations themselves are emergent, as

Morgan thinks, or whether some relations come hand in hand with emergent qualities, i.e. qualities not found at the microphysical level, as Alexander thinks. That is, *merely pointing to something novel in a composite object doesn't prove that there is something emergent.* However, this does not imply that there aren't other ways of trying to prove or disprove that a particular property is emergent.

Chapter 3: The British Emergentists and the Prediction Argument

I. Overview

In the previous chapter, I argued that neither the phenomena of composition nor the apparent existence of novel, relational glue properties entail the existence of higher-level, emergent properties. In this chapter, I contend that the more detailed accounts of emergence offered by Morgan, Alexander, and Broad don't fare any better. To make this case, I give a brief overview of what the emergentists actually say about the ontology of emergent properties. To that end, I first discuss Alexander's doctrine of emergent, higher-level qualities. I then explain Morgan's doctrine of emergent relations and argue that emergent relations can be interpreted in two ways. On the first interpretation of Morgan's emergent relations, emergent properties are emergent solely because they are relational. However, as we saw in the previous chapter, we can account for relational properties without positing emergence. Therefore, on this reading, Morgan's multi-level ontology collapses and ends up being indistinguishable from a uni-level view. On the second interpretation, Morgan's emergent relations are distinguished qualitatively from lower-level relations. But the idea that emergent properties are different in quality from their lower-level counterparts is Alexander's view. I end this chapter by discussing what I call the prediction argument, which is supposed to establish the existence of higher-level properties, higher-level qualities in particular. I show that the argument is unsound for a variety of reasons. For one thing, the argument is based on an invidious distinction between that which is predictable *a priori* and that which can only be predicted by *a posteriori* means. Moreover, once we remember that not all identity statements are known to be true *a priori*, we can see that the prediction argument is not valid. I finish

this chapter by investigating the cause of the intuitive appeal of the prediction argument. I explain that we can dispel this intuitive appeal by remembering Frege's distinction between sense and reference.

II. Alexander and Emergent Qualities

Let's start by looking at Alexander's (1918) view of emergent qualities, which is summed up in the following passage:

> Empirical things or existents are, it has been more than once suggested in accordance with our general conception, groupings within Space-Time, that is, they are complexes of pure events or motions in various degrees of complexity. Such finites have all the categorial characters, that is, all the fundamental features which flow from the nature of any space-time in an empirical form — each finite has its proper extension and duration, is built on the pattern of its specific universal, in a substance of a certain sort and the like. What remains to be described is its possession of quality. The facts can best be described as follows. New orders of finites come into existence in Time; the world actually or historically develops from its first or elementary condition of Space-Time, which possesses no quality except what we agreed to call the spatio-temporal quality of motion. But as in the course of Time new complexity of motions comes into existence, a new quality emerges, that is, a new complex possesses as a matter of observed empirical fact a new or emergent quality. (Volume 2, Book 3, Chapter 2b)

In other words, Alexander (1918) suggests that "…in the course of Time new complexity of motions comes into existence, *a new quality emerges*, that is, a new complex possesses as a matter of observed empirical fact a new or emergent quality." (Volume 2, Book 3, Chapter 2b) [italics are mine] If we look past the emotive terms and the abstruse technical terms, Alexander's idea is not that complicated. He is articulating a version of the levels doctrine wherein there are higher-level emergent "qualities," but there are no emergent "motions."

This distinction between qualities and motions is somewhat clear, but it demands a bit of explanation. I think Alexander's use of the word "quality" is not all that different from Locke's as Locke uses it in the term "primary quality." A quality is just a way some region of the world is modified, i.e. a qualitative property or feature like shape, color, etc.

And I think it is also clear that the sort of qualitative properties that Alexander has in mind here are intended to be contrasted with what we would call "causal powers." Qualities are ways an object is modified, i.e. its wholly categorical features. By contrast, an object's causal powers are the ways an object is disposed to behave in different contexts. So, we can update Alexander's position by saying he is arguing for the existence of emergent qualities but not emergent causal powers.[49]

There is a fair amount of textual evidence for this interpretation of Alexander's concept of qualities and motions. For one thing, he tries to explain the concept of a quality by giving us a paradigmatic example of a quality, which he does by pointing to what we now call "qualia." For Alexander, qualia are an example of qualities more generally, and so as long as we know what the qualities of experience are, we know what qualities are. He (1918) states,

> Mind is at once the case which most urgently forces on our attention the problem of quality and at the same time offers the readiest means for its solution. For our mind is experienced by us as a set of connected processes which have the character of being mental, possessing the quality of 'mentality,' or as I shall most frequently say, the character of consciousness. (Volume, Book 3, Chapter 1A)

Put simply, Alexander's idea is that when you introspect a pain you see that the pain has a certain qualitative feel. Or when you watch a hockey puck in motion, you have an experience of something with the phenomenal quality of looking black in your visual field.

Returning to the distinction between powers and qualities, these paradigm cases of mental qualities, Alexander argues, teach us that qualities of sensory experience appear to be conceptually distinct from the causal powers of our experiences. For example, I can

[49] This may sound as if it borders on epiphenomenalism. I will argue that it does not, and Alexander took pains to explain that it did not. Moreover, it is ultimately this characteristic of his view that saves it from Kim's causal exclusion argument.

and do differentiate the qualitative feeling of the pain that comes with a migraine headache from the changes the migraine brings about in my disposition to behave in certain ways, e.g. the disposition to reach for an Advil or to cover my eyes. For Alexander, all qualities are analogs to the qualitative feel of a migraine in this way. That is, the qualitative features of objects in the external world, like shape, position, motion, and solidity are conceptually distinct from the causal powers we might associate with these qualities, e.g. as the power to impart force is associated with solidity.

One concept that modern readers will find uncongenial, to say the least, is that of "pure events or motions." I take it that "pure motion" is supposed to be a succession of causes and effects that is wholly lacking in any quality whatsoever. To use more contemporary terms, to conceive of pure motion, we would have to conceive of the causal powers of an object without conceiving of any of its qualities. It is an odd idea, one that I have difficulty grasping myself, but by "pure motion" Alexander means something like an event occurring without any change in quality. Again, it is not easy to see exactly what Alexander has in mind here, but it helps to notice that Alexander thinks of space and time as qualities, and of being solid as a quality, too. So, on Alexander's view, an atom is nothing more than a bundle of qualities like solidity, charge, a certain shape, existing in a certain region of space. So, since Alexander sees an atom as a bundle of qualities, including the quality of having a certain spatial location, the event of that atom moving is just a change in its qualities.[50] For example, suppose at one moment, the quality of solidity is pervading this space over here, i.e. there is a solid body in such and such a space, and at the next moment the solid body moves, i.e. now the same space isn't

pervaded by the same quality, and some other region is. It is possible to imagine the motion of a solid body through space as a change in the relationships between qualities.

Perhaps formalizing this a bit will help explain it. Let's imagine the quality of solidity, let's call it S1, existing in such and such a region of space, call it R1, at such and such a time, T1, and at T1 the adjacent region R2 is devoid of that quality. Next, imagine that at a later time T2, S1 moves to an adjacent region R2, i.e. it is no longer at R1 but is at R2. We can represent that by claiming that the pair <R1,S1> signifies that solidity pervades the region R1 and <R1, ~S1> signifies that S1 does not pervade R1. And we can say the change we see just is the difference between {T1, <R1, S1>, <R2, ~S1>} and {T2, <R1, ~S1>, <R2, S1>}

Notice, the differences between these two sets can be understood even if we don't know what possesses the various elements in the sets or what qualities they refer to. And we can describe the change in the ordering of the elements of the set regardless of what the elements are. This is, I believe, what Alexander has in mind when he says "pure motion" or "pure events." i.e. pure motion is the change in the two sets regardless of what possesses the qualities are referred to by "S1," "R1," etc., or what in fact the qualities are. For example, let's imagine S1 is the property of being a baseball moving from R1 to another region R2. The difference of the two sets would describe that change. Conversely, imagine S1 is a collection of lower-level molecular properties which the property of being a baseball emerged from. The change in the two sets would describe the same event. Indeed, the event or motion of the molecules moving is the same motion as the baseball moving. According to Alexander, yhe only thing different in the first case is

that the qualitative property of being a baseball has moved whereas in the second case, the change or motion was occurring in base-level properties.

Of course, it may be that Alexander is wrong about the conceivability of "pure motions." Large portions of his ontology could still be correct, including the idea that we should conceive of motions or events independently of qualities.[51] Indeed, Alexander's view of emergence seems to make more sense if we take the motion of objects in space-time as a given and work from there. (It should be noted that Alexander wouldn't have agreed with that, since he wants to meld his concept of emergence with his concept of space-time from the very long first volume of *Space, Time and Deity*.) At any rate, we can put any problems arising with the concept of "pure motions" aside for our purposes.

Alexander seems to be aware that he needs to show that emergence as a relation between higher and lower-level qualities is at least plausible. He does this by pointing out what he thinks is an analogous relationship: the Aristotlean relationship between form and matter. He writes:

> To adopt the ancient distinction or form and matter, the kind of existent from which the new quality emerges is the 'matter' which assumes a certain complexity of configuration and to this pattern or universal corresponds the new emergent quality. But whereas up to the present we have been content to treat the quality as something which is correlated with a certain configuration of its basis, we can now, following the clue of the relation between mind and its body, identify the quality with its peculiar form of body. (Volume 2, Book 3, Chapter 2b)

I think the point of this passage is to suggest that an emergent quality is related *necessarily*, to, the "configuration" of its component parts, or, in our terms, the relations between the component parts, even though the parts and the whole aren't identical. For Aristotle, matter can't exist if it isn't arranged into a particular form, and a particular form can't exist without matter. But matter isn't form and form isn't matter. Each

[51] Morgan suggests that he, in general, agrees with Alexander's position on emergence but disagrees with Alexander's odd account of base-level properties, along with his odd advocacy of divine properties.

necessitates the other, but matter and form are ontologically distinct, despite the fact that they necessarily occur together. Alexander, therefore, seems to be suggesting that the emergent quality simply can't occur without "the configuration" at the lower-level, and that once that lower-level configuration comes into existence it necessitates that such and such a quality has to exist, too. Nonetheless, the configuration and the quality are different entities. Thus, the relation between emergent qualities and the lower-level configurations that they emerge from is a necessary, symmetric determination relation that is not identity. (Of course, the analogy only proves that emergence as Alexander conceives it is plausible if the Aristotlean relationship between form and matter is plausible.)

But this raises a problem similar to a problem we considered in the first section of the second chapter. Identity is also a necessary, symmetric determination relation. So, should we say that an emergent quality is identical to the configuration it emerges from or something in the configuration it emerges from? Of course, Alexander has to say that emergent qualities aren't identical to the relations or configurations that they emerge from, After all, emergent qualities can't be identical to anything at the base-level, because emergent qualities are supposed to be -this is the whole point- something more than lower-level properties.

What I want to point out is that it is at least possible, given all that Alexander has insisted on, that what Alexander calls "emergent qualities" don't exist. *Instead, perhaps all that exists are clusters of base-level qualities, including the base-level qualities associated with base relational properties* (like covalent and ionic bonding).

Moreover and more importantly, if we are going to believe in emergent qualities, we are going to need some reason to do so. Until we have such evidence, higher-level qualities should be seen as otiose ontological extras. Of course, what I call "the prediction argument" is supposed to do exactly that. We will consider this argument in the fifth section of this chapter.

III. Morgan and Emergent Relations

Morgan's view of emergence is, as he admits, basically in broad agreement with Alexander's. In the second section of the first chapter of *Emergent Evolution*, Morgan (1923) states,:

> The most resolute attempt to give a philosophic interpretation of nature as a whole, with adequate stress on the concept of emergence, is that of Professor S. Alexander in *Space, Time, and Deity*. In order to get at the very foundation of nature as it now is, he bids us think out of it all that has emerged in the course of evolutionary progress—all that can possibly be excluded short of annihilation. That gives us, as an inexpugnable remainder, a ground plan of ultimate basal events (pure motions) with naught beyond spatio-temporal terms (point-instants) in fluent relations of like order. This he calls space-time, ubiquitous, all-pervasive, and inseparably hyphened. From this first emerged "matter" with its primary, and, at a later stage, its secondary qualities. Here new relations, other than those which are spatio-temporal only, supervene. So far, thus supervenient on spatio-temporal events, we have also physical and chemical events in progressively ascending grades. Later in the evolutionary sequence life emerges—a new "quality" of certain material or physico-chemical systems with supervenient vital relations hitherto not in being. Here again there are progressively ascending grades." (Lecture 1, Section 2)

But Morgan disagrees with Alexander about what it is that emerges. That is, Morgan claims that he wants to deemphasize emergent qualities and replace them with "emergent relations." Morgan (1923) writes:

> "…we must first clearly grasp his use of the word "quality." He speaks of the emergence of new qualities. He would say that at some stage of inorganic evolution this or that so-called secondary quality, such as colour, emerged; that at some later stage of evolutionary process the quality of life emerged; and yet later the quality of consciousness. I shall often use the word "quality" in this sense. But my own interpretation runs rather on lines of what I call relatedness. The discussion of relatedness, to which I shall devote the third chapter, requires the consideration of the terms in relation within any given field of relatedness, and of the relations of these terms. Relatedness, in my sense of the word, includes both; not the terms only; not the relations only; for they can never be divorced if my usage of the word "term" be provisionally accepted. I shall speak of the relatedness which obtains wholly within any given system as *intrinsic*; and I shall distinguish the

> relatedness of this system to some other system, or systems, as *extrinsic*. A system of intrinsic relatedness I shall provisionally call an entity. In so far as the character of a natural entity is determined by intrinsic relatedness I shall speak of it as a quality which is an expression of that intrinsic relatedness. In so far as the character of a natural entity is determined by extrinsic relatedness to other such entities, I shall speak of it as a property which expresses that extrinsic relatedness (Lecture 1, Section 4).

He (1923) goes on to claim that "what is supervenient at any emergent stage of evolutionary progress is a new kind of relatedness -new terms in new relations- hitherto not in being."[52] (Lecture 1, Section 4)

But here Morgan has run head on into the following problem, which we discussed in relation to the sticky parts and external glue models of composition, and earlier, when discussing the existence of composite objects. The mere presence of a relation that glues the parts of a composite object together does not imply the existence of an emergent property in the composite object. For example, spatial relations exist at the microphysical level. They relate microphysical entities in space. And surely ionic and covalent bonds exist at the microphysical level, too. So why should we believe that so called "emergent relations" are emergent at all? What exactly is emergent about them? What reason do we have to believe that so called "emergent relations," aren't just microphysical relations by a different name?[53]

Moreover, suppose we can demonstrate with certainty that a composite object has properties that its individual component parts do not. For example, suppose a water molecule has novel properties, i.e. properties which the individual hydrogen and oxygen atoms composing it do not have. It is still possible to maintain that the novel properties of the composite object, the molecule, are just relational properties that have never been instantiated before, i.e. external glue relations. (This was discussed, at some lengths, in a

previous section of this chapter.) It is consistent to say of a statue of Lincoln, for example, all the following: (1) there are relational properties between the atoms that compose it that never existed until someone made such a statue, indeed, that the shape of the statue is a complex, previously uninstantiated spatial relation between the atoms. (2) That spatial property of being Lincoln-statue shaped may very well be novel, i.e. it may never have existed until someone made a Lincoln-shaped statue. (3) That spatial property is a relation which is identical to a base-level property. It is consistent to maintain that the shape of the Lincoln statue just is a microphysical property, the very complicated relational property that the atoms composing the statue bear to each other.

Indeed, even if we could show that composite objects must have novel relations, this doesn't establish that there are emergent properties. We must remember, as we saw in our discussion of the sticky parts model of relations, that some relations which appear to be novel relations may not be novel at all. For example, it appears that a covalent bond between atoms doesn't exist until the atoms are brought together into a composite molecule. But if the covalent bond just is identical to the charges of the electrons, the pull that the electrons exhibit, then the bond was in existence, in a way, before the molecule was composed, since the charge of the electrons was in existence before the molecule. The bond manifests what is already there. It is possible to hold that a predicate that appears to pick out a single emergent relation actually picks out a plurality of different microphysical relations all of which must be present for the predicate to apply. That is to say, semantic conjunctivism can be formulated for relational properties so that one relational predicate can pick out many relational properties. As long as this reworking is possible, the data Morgan cites does not warrant belief in emergence. Thus, when

pressed, Morgan's theory of emergent relations is consistent with a collapse into a uni-level theory.

One way to save Morgan's view from this collapse is to claim that there is something qualitative about some relations that makes those relations distinct from relations between microphysical, base-level entities. In other words, Morgan could say that emergent relations like the relations between neurons are qualitatively distinct from clusters of micro-chemical relations. But if Morgan were to take this escape route, then his view becomes conceptually indistinguishable from Alexander's, because Morgan would then be arguing that ultimately what emerges are higher-level qualities. And Morgan's view would then stand or fall with Alexander's. (I argue below that Alexander's view rests on, and collapses with, the prediction argument.)

Another way to try to save Morgan is to say that emergent relations possess distinct causal powers, i.e. causal powers that don't exist at the microphysical level. Morgan seems amenable to this ways of defending his emergent relations. He writes:

> In virtue of such new kinds of relatedness, not only have natural entities new qualities within their own proper being, but new properties in relation to other entities. The higher entities are not only different in themselves; but they act and react differently in presence of others. (Lecture 1, Section 4),

The latter part of this citation places Morgan's philosophy squarely in the crosshairs of Kim's (1989 and 1992) causal exclusion problem. But the former part of the citation suggests that novel qualities exist "in virtue of" new kinds of relatedness, which perhaps implies that Morgan would, if pressed, take back the claim that there are novel higher-level causal powers.

Regardless of what Morgan actually thought about qualities and relations, he has to make a choice. Either the novel relations that exist in composite objects are identical to

microphysical, base-level properties (of the components) or they are not. If they are, then there is no reason to say that there is anything emergent. If they are not, then there must be something about those relational properties that implies that such relational properties aren't identical to microphysical relations like covalent bonds, ionic bonds, and spatial relations. This emergent "something" must either be a quality, a causal power, or both. If the emergent ingredient in relations is a quality, then Morgan's view needs to be supplemented with an account of what sets this quality of relational properties apart from other qualities of relational properties. A very natural way to fill this need, and one which the British emergentists made famous, is the prediction argument.

IV. The Prediction Argument

The prediction argument is the weapon of choice for the emergentist. Unfortunately, it is not the weapon it was once thought to be. Indeed, I think the argument has been known to be faulty, since the 1950's. This is bad news for the British emergentists, since we have seen that, without the prediction argument, they cannot give sufficient evidence to demonstrate the existence of emergent properties in composite objects.

The argument has an interesting history. It evolves out of Mill's (1843) distinction between heteropathic and homopathic effects and takes form in Lewes's (1875) distinction between resultants and emergents.[54] But the argument is presented by Broad (1925) who explains the key premise of the argument as follows.

> It is clear that in no case could the behaviour of a whole composed of certain constituents be predicted *merely* from a knowledge of the properties of these constituents, taken separately, and of

[54] See McLaughlin (1992) for an overview.

> their proportions and arrangements in the particular complex under consideration. (Chapter 2, Section 2)

The key premise, the key piece of evidence in favor of emergent properties, is that we can't, even in theory, "deduce" that a certain property will be instantiated in a particular composite object, *a priori*, no matter how much we know about the lower-level properties of the parts. That is, even if we had a complete, God-like knowledge of base-level properties, we couldn't predict *a priori* that arranging such and such base-level properties into a composite object would result in the property in question. He then concludes that such unpredictable properties are emergent.

Broad suggests that this premise implies that the property in the composite which couldn't be predicted cannot be identical to anything that exists at the base-level, nor the mere mereological sum of things existing at the base-level, precisely because if it were identical to a sum of base-level entities, it would have been predictable *a priori*. At first glance the argument seems to be valid. After all, it is true that we can always predict *a priori* what a mereological sum of two entities will be like. To use an example Broad discusses, if we have one tap pouring 1 gallon of water per minute into a tub, and another tap pouring 1 gallon of water into the same tub, we can predict, *a priori*, that the result of turning both taps on will be that 2 gallons of water per minute will pour into the tub. (Actually, Broad is wrong about this. The conservation of matter is not known *a priori*. It is not *a priori* that, even if matter is not conserved, that water is conserved. It is not *a priori* that no odd interaction between water molecules occurs when the two taps are open. Indeed, it is just a fact of experience that masses are additive. But for the moment, we can allow Broad to use this example to illustrate the prediction argument.)

According to Broad, the total amount of water pouring into the tub is a mere resultant, a mere "mereological" sum of the amount of water flowing from each tap. What this shows is that, in general, if something is a mere resultant, then it is predictable *a priori.* So, if there is something about a composite object that couldn't have been predicted *a priori,* we can conclude that this something is not just a mere "mereological" sum, not just a mere resultant, of things that were already there. Or at least that is how the argument goes.

If this argument is sound, it would seem that the emergent property must be related to the base-level parts and the relations between the parts, not by identity, but in some other way, i.e. by what Broad calls "a transphysical law." Thus, an even simpler way of stating the argument is to say that, since transphyical laws are discovered empirically, then transphysical laws cannot express or report identities. Or more precisely, the things related by transphysical laws cannot be identical to each other, because if they were, the occurrence of the one could have been predicted *a priori* whenever and wherever the other existed.

At this point, the prediction argument has come into tension with a core tenet of contemporary philosophy: the belief that there are *a posteriori* identity statements. We know not just from Kripke (1980), but from Frege too, that there are *a posteriori* identities, and in the light of what Kripke and Frege have taught us, the prediction argument seems to be in serious trouble. Indeed, the fact that Kripke and Frege (who disagree about so much) agree about this means that Broad is almost certainly misguided here. The prediction argument asks us to accept that if we discover a relationship between two things, X and Y, *a posteriori,* then the relationship between X and Y can't be

identity. In the light of the externalist revolution in semantics, the distinction between rigid designation and reference-fixing, and Frege's seminal distinction between sense and reference, we simply have to conclude that this argument is invalid; it is based on a thoroughly debunked conception of the relationship between identity and *the a priori*.

Consider the well worn point that we cannot deduce *a priori* that Hesperus is Phosphorous. Nor can we deduce *a priori* that the morning star, the evening star, and Venus are all identical. No matter what we think about the semantics of such statements, of internalism or externalism, or reference-fixing versus sense, we now recognize that these truths must be discovered by empirical means. Of course, the fact that the sentence was discovered to be true empirically doesn't imply that "Hesperus is Phosphorous" isn't an identity claim because it may very well be an *a posteriori* identity.

In other words, as long as there *a posteriori* identity claims, the prediction argument is invalid. And we have a long history, again from Frege to Kripke, showing that a posteriori identity claims are commonplace. To take another familiar example, we can't predict *a priori* that "is gold" and "has atomic weight 79" pick out the same property; we had to learn that "gold has atomic weight 79" empirically. But that doesn't meant they're not identical. If it did, we would have to reject a good deal of what we think we have learned about the externalist causal theory of reference *and* the internalist Fregean notion that sense determines reference. And apart from these philosophical considerations, what *other* relation could there be between the morning star and the evening star, or being gold and having atomic weight 79, if not identity?

Indeed, I suspect one reason why Smart (1959b) and Feigl (1969) introduced the concept of theoretical identifications like "Heat is mean kinetic energy" and "Lightning is

electrical discharge" was to accommodate and clarify the fact that you didn't need transphysical laws to explain the relationships between the apparently different entities described by predicates in the different branches of science. Thus, the prediction argument has been known to fail for a long time, since Smart and Feigl's seminal work at the latest.

I would also suggest that Nagel's (1961) arguments in *The Structure of Science* for bridge laws make a similar point, apparently in favor of reductivism. However, it is somewhat unclear whether Nagel's bridge laws are theoretical identifications or whether they are Broad-style laws of emergence. All Nagel really tells us about reduction and bridge laws is this: A theoretical term in what Nagel calls a "secondary science" reduces to theoretical terms in the "primary science" of physics if and only if the truth of claims in the secondary science can be derived from the bridge laws and the claims of the primary science. Part of the difficulty here is that Nagel subscribes to the idea that the existence of "properties" or "essences' is theory-relative. So, for him, there may be no answer as to whether, for example, the bridge law that the mean kinetic energy of a sample of gas (expressed in units of energy) is equal to the heat of the gas (expressed in units of heat) is an identity between the property of being or having kinetic energy and the property of being or having heat.

All of this leads me to suspect, *pace* the assertion of a number of introductory texts in the philosophy of mind,[55] that the identity theory of the 1950's was created more for the purposes of providing an alternative to emergentism than as an alternative to

[55] See especially Paul Churchland (1988)

Rylean philosophical behaviorism.[56] I think the textual evidence bears this out. Nagel has a whole chapter in *The Structure of Science*, where the work of Broad and the emergentists figures heavily.[57] Feigl (1969) also discusses emergence explicitly in *The Mental and The Physical.*[58]

But we can put these issues regarding the history of philosophy and emergentism in the middle of the 20th century aside for now and return to the topic at hand. As long as the prediction argument is invalid, the identity theory (of the 50's) gives us a better explanation of the relationships between different theoretical phenomena like heat and mean kinetic energy, because it is just as explanatorily adequate and more parsimonious than positing transphysical laws of emergence. For example, identity explains why Venus and the morning star are as intimately related as they seem to be, without positing two different things and an inscrutable relation between them. It posits one thing and a clear relation, identity. So it can do the same thing for things like heat and mean kinetic energy.

The more standard problem with the prediction argument, which I mentioned earlier in connection with the seemingly trivial identity: 1 gallon plus 1 gallon = 2

[56] The story of the history of identity theory, emergentism, and behaviorism is a little more complex than this. Interestingly, Smart (1959a) explicitly critiques Ryle for accepting emergence and the prediction argument. Moreover, Armstrong (1968) formulates his causal theory of mind by explaining how mental states are the cause of behavioral dispositions, thereby making his version of the identity theory an advance on Rylean behaviorism.

[57] This chapter is also interesting in that Nagel (1961) and I take similar stands on the value of relational properties in explaining the existence of composite objects.

[58] Foreshadowing my semantic disjunctivism, Feigl (1969) even surmises that there have to be "one-many" identities. He writes, "Macro-temperature, as thermometrically ascertained, corresponds in one-many fashion to a multitude of micro-conditions, viz., a very large set of molecular states. Strictly speaking, this correspondence holds between one state description on the macro-level with a specifiable infinite disjunction of state descriptions pertaining to the micro-level. Since, as we have also pointed out, this correspondence is empirically ascertained, there is here as little reason to speak of logical identity as in the ψ-φ case. Nevertheless, we have seen that it makes sense, and what sense it makes, to regard the relation of temperature to mean molecular kinetic energy as an example of a theoretical identity." (Section 6. C)

gallons, which was first pointed out by W.T. Stace (1939),[59] is that *you can **never** predict a priori* that a property will be instantiated. In an unfortunately overlooked paper, Stace (1939) writes:

> This distinction between heteropathic or emergent effects and homopathic or non-emergent effects is not tenable. Emergentists make the distinction twofold. They say (i) that non- emergent effects are predictable, emergent effects unpredictable. And they say (2) that emergent effects are novel, non-emergent effects non-novel. Neither of these contrasts can be maintained. First, as to predictability. The heteropathic effect, the blue liquid, it is said, could not be predicted from the two white liquids without experience. Once I have experienced the sequence I can, of course, predict its recurrence, relying on the uniformity of nature. But without experience I could never predict it. This is certainly quite true. But is it not equally true that, without experience of impacts, I could not predict the angles or velocities of the new motions of the billiard balls? Indeed I could not even predict that there would be any motions at all. When a moving billiard ball strikes a ball at rest, why should both not thereupon stand still? Or why should they not turn and go backwards upon their tracks? For the matter of that why should they not turn into watermelons, or disappear out of existence altogether? … One has to wait on experience to find out what will happen as the result of an impact just as much as to find out what will happen when one mixes two white liquids. Once we have experienced the sequence, in either case, then of course we can predict future sequences of the same kind upon the basis of the uniformity of nature. I am, of course, merely repeating here considerations which have been familiar to everyone since the time of Hume. And it is accordingly unnecessary for me to elaborate them. It is only necessary to note that what Hume said really disposes of the alleged difference between emergent and non-emergent effects in the point of predictability. Both are equally predictable after experience. Both are equally unpredictable before experience... In the case of the billiard balls we have a change of velocities and direction. Now why is a change of color a novelty, and a change of velocity and direction not a novelty? Would it be rational to say that a change of color involves novelty, but that a change of smell does not? And can there be any more justification for saying that a new color is novel, but that a new velocity and direction are not? I cannot see the slightest justification for such a distinction in any of these cases. If I am right, this second distinction between emergent and non-emergent effects collapses as certainly as did the first. Why is it, then, that we all tend to think that there is some distinction? For I think it does seem to most of us that the case of the liquids is somehow unlike the case of the billiard balls. This is, I believe, a sheer illusion, and I account for it as follows. I suggest that what are called homopathic or non-emergent effects are simply those which are more common in our experience, more familiar, less surprising, less striking. The so-called heteropathic, or emergent, effects are those which are comparatively rare, so that when they do occur they seem more striking and unexpected, and so we get the impression that we are in the presence of some kind of novelty which is absent in the other cases. (pgs. 305-308)

In simple terms, Stace is charging British emergentism with selectively applying Humean skepticism about causation and prediction to a narrow range of cases in an attempt to differentiate emergent properties from mere resultants. But if Hume's skepticism is deployed consistently, *we can never predict* what property will occur in intra-level causation or in composition purely *a priori*.

[59] As near as I can tell, Stace's (1939) article has been largely overlooked, which is both strange and disappointing given Stace's prestige in philosophy and the quality of his argument.

Stace (1939) is right to suggest that the only reason we think we can predict some changes *a priori* is that some things are so familiar to us that we think we always knew them and that we never needed to learn them *a posteriori,* and so therefore we must know them innately or *a priori.* For example, a paradigm and oft-cited case of a resultant property is the motion of a solid, stationary object struck by another object of the same mass. It may seem that we just know *a priori* that the struck object will move, but really we know that because we have seen it happen many times. We learn these things at such an early age and are so familiar with them that we assume we never learned them. But we did learn them.

If you object and suggest someone *could* -by means of luck or genius- derive the laws of thermodynamics without experience of physical objects moving around, this doesn't refute Stace's point. By the same reasoning, someone could -by means of luck or genius- just as well deduce that such and such a higher-level quality will be instantiated if such and such lower-level qualities are present in a composite object. Actually, nobody can predict *a priori* the instantiation of so-called higher-level properties, and nobody can predict *a priori* the conservation of momentum. It follows that there can be no distinction between emergents and resultants of the sort Broad and Morgan tried to articulate. Thus, showing that the instantiation of a particular property could not have been predicted *a priori* doesn't show anything at all, since we can never make this kind of *a priori* prediction.

Again, Stace (1939) recognizes that we have a strong intuition that certain effects can be predicted *a priori*. It certainly seems that we can predict *a priori* that if we propel a solid object towards a stationary object with the same mass, then the effect will be that

the momentum from the first object will be transferred to the stationary object. (Think billiard balls.) But this intuition, Hume maintained with celebrated urgency, is just the result of habit and custom. It is only because we have a long history of experiences where we observe the motions of solid objects that we come to expect that the second object will move when struck by the first. If we had never observed such a collision and the aftermath of such a collision, we could not predict what the aftermath would be. Now, it may be that Hume's theory of causation is wrong. The fact remains that Broad did not supply a criterion for distinguishing properties of wholes that are predictable from part-properties, and properties of wholes that are not predictable. The intuition that this distinction is self-evident can be explained by habituation even if we reject Hume's rejection of real causation. So the prediction argument has a big hole where its foundation ought to be.

It certainly seems that we can predict what properties a swarm of gas atoms will have if we know everything about the properties of the individual atoms and the relational properties that tend to form between the individual atoms, e.g. ionic bonds. But we must give Humean psychology its due. If we had never observed gas atoms or atoms of any kind interacting with each other, so that we had never observed what relations they tended to form with one another, we wouldn't be able to predict what would happen when we put the atoms together. So, it appears that we just do, as a matter of fact, learn what relational properties exist in nature or what sorts of bonds individuals tend to form through observation. Only the most extreme rationalist would deny this. Or, at the very least, without any empirical information, we will be blind as to what relational properties

exist. But this seems true for all properties or even all things; we need empirical information to know anything about the world.

Therefore, when Broad or any of the emergentists suggest that, for example, a biological organism possesses properties whose existence couldn't be predicted *a priori* even if we knew everything about the atoms composing the cell, the proper response is to say, "So what?" After all, we can't predict *a priori* what relational properties any composite object will have, even one as seemingly microphysical as a swarm of gas atoms. All this shows is that in order to know what properties exist we need some empirical data to work with, and that is close to a truism.

What I wish to emphasize here is that, while Stace's unduly-neglected argument is quite decisive, the fundamental reason the prediction argument fails is the more general and more fundamental point that identities can be *a posteriori*, and in science typically are. It is the absolutely ground-floor error of forgetting this, or ignoring it, that vitiates Broad's argument.

V: The Intuition Behind the Prediction Argument

It might be objected that I am being unfair to the prediction argument and that perhaps it is possible to immunize the argument from my attacks. Indeed, despite the problems we have discussed, there seems to be some intuitive force behind the prediction argument, i.e. an intuition that there are higher-level properties, which can't be predicted from a lower-level perspective. It cannot be denied that there is a strong intuition at play here. However, as I will argue below, this intuition is misleading.

The intuition I am talking about manifests itself in a series of intuition pumping thought experiments and metaphors often invoked by the emergentists. The most famous of these thought experiments is Broad's discussion of the "mathematical arch-angel." The arch-angel, like Laplace's demon, is infinitely efficient at calculating. But we are told the arch-angel perceives and understands only the microphysical mechanical properties of the base-level. Broad then asks whether the angel can use his infinite calculating powers to create an understanding of higher-level properties, like the macro-physical property of being brown. Broad says that the angel cannot do this, despite his infinite powers of ratiocination, because, although the angel can see that a particular swarm of objects reflects light in a certain way, he cannot see that it is brown. And the angel cannot see that it is brown because he is looking at the world from a different perspective than we do. Thus, Broad's arch-angel example pumps an intuition, which seems to tell us that there are in principle unpredictable higher-level properties.

But this intuition about what isn't predictable is misleading. The prediction argument must prove, for example, that the predicate "is a protein" can't refer to the same properties as, say, "is five thousand carbon, oxygen, and hydrogen molecules bonded together by such and such covalent and peptide bonds." And there does seem some intuition behind saying that there must be some difference. But why do we find it so natural that these two predicates, i.e. the short biological predicate and the longer, complex microphysical predicate, don't pick out the same properties? Why does everyone suspect that the two predicates have different reference? The force of the intuition pumped by the mathematical archangel thought experiment suggests that we should answer this question as follows. The reason for believing that is just that

biological things *seem* very different to our eyes than microphysical things. That is, from the perspective of biology, there *seem* to be biological properties, which are different from chemical properties. And the further we go up the levels, the less the putative higher-level things seem like microphysical things. For example, horses seem even less like swarms of atoms than protein molecules do.

Alexander's belief in emergent qualities is ultimately the result of this misleading intuition at work, too. Alexander believes in emergent qualities because he thinks the mere fact that A and B seem to have, or be, different qualities to him implies that they actually do have different qualities. Alexander's arguments really boil down to this idea: there seem to be higher-level qualities, so therefore there are higher-level qualities.

It should have been obvious that this is a non sequitur, but in their zeal to come up with an explanatorily adequate ontology -to explain the difference in how things seem- the emergentist metaphysicians went overboard by positing more entities than necessary. The fact that there seem to be two different qualities present in a situation does not require that there are. The fact that the wall seems covered by a green slime does not require that there be a green slime, or anything of any sort, on the wall. *The fact that gold doesn't seem like a substance whose atoms have 79 protons may make it seem that "being gold" and "having 79 protons" are different properties. But all we can be sure follows from the seems-statement is that these two predicates have different senses, not different references.*

It is of the first importance to notice that we sometimes represent the same thing to ourselves in multiple different ways. This is virtually a philosophical cliché, but its significance is nonetheless often overlooked, particularly in connection with the

prediction argument. For example, we can represent heat as either that which makes me feel hot or as mean kinetic energy of a swarm of atoms. Ultimately, that which makes me feel hot just is mean kinetic energy in the swarm of atoms surrounding my body. But I represent this "that which" to myself in at least two different ways: as the cause of the feeling of heat and as the average of mean kinetic energy in a swarm of molecules.

It is undeniable that we represent the same things in multiple ways in this way, and if we didn't do it, we would have to accept that any two things that seemed different to us on a given occasion are different. But that would be absurd, given that the things that normally seem to be identical, say Superman and Clark Kent, might seem different to another person on at least one occasion. Therefore, if we accept that a difference in how things seem, which may be no more than a difference in how we conceive of a thing, necessitates an actual difference in the things conceived, we would quickly found ourselves burdened with an untenable superabundance of things in flux. That is, there would be a way things are for every way things are conceived, and the world would change as our conceptions of it changed.

Turning to putative emergent properties, I grant that it sometimes seems that "is a horse" has to pick out a different property than "has such and such atoms with such and such properties arranged horse-wise." For one thing, it would seem, intuitively, that being a horse is one unified property, while at the atomic level, there is a buzzing swarm of very different things with very different properties. But again, the fact that we have two very ways of conceiving of what a horse is doesn't mean that there is any difference between being a horse and the cluster of base-level properties of the atoms that make up the horse.

There is an obvious parallel between what I am saying about predicates and the Fregean concept of sense. In Fregean terms, I am suggesting that “is a horse” and “is a set of atoms with such and such properties arranged horse wise” are different in sense but not in what they refer to or pick out. The fact that we conceive of horses differently than we conceive of swarms of atoms does not imply that the predicate “is a horse” can’t pick out the same properties as a complex microphysical predicate. More generally, it may well be that the different branches of science differ from each other in how they go about conceiving of the world and the predicates and terms employed in each branch of science ought to have different senses, even though they all refer to the same microphysical properties. Perhaps Alexander’s mistake, then, was taking a difference in sense for a difference in real qualities. The former are part of how we conceive the world, the latter part of how the world is. And the mere fact that we conceive of the world in two different ways does not imply that the there are two different ways things in the world are.

At this point someone could object that the mere fact that we have a useful concept of being a horse, or the mere fact that the predicate “is a horse” is projectable, is evidence that there is a property of being a horse which is picked out by the predicate “is a horse.” And it might also be suggested that the fact that the two concepts is a horse and is a swarm of atoms are useful in different ways for different reasons -and the predicates “is a horse” and “is a swarm of atoms” are not projectable to the same degree- is evidence that the two predicates pick out different properties. These are serious objections and I discuss them at length in the fifth chapter.

Chapter 4: Multiple Realization and Disjunction

I. Overview

I begin this chapter by introducing and explaining the metaphysics of contemporary nonreductive materialism, the view originally advocated by Fodor (1975) and Putnam (1974). I argue that their view ought to be seen as a version of the levels doctrine, centered on the now well worn notion of multiply realizable properties. I then argue that Fodor and Putnam aren't able to give us a good reason to posit higher-level multiply realizable properties because we can explain away the appearance of multiply realizable properties with some simple, tenable claims about semantics of predicates. Just as the argument for emergent properties of composite objects dissolved when we realized that one predicate can pick out or refer plurally to multiple properties of the parts of a single object, the argument from multiply realizable properties dissolves when we realize that *a single predicate can pick out or refer plurally to different properties –or different n-tuples of properties- in different objects.*

I then explain how Kim's (1989 and 1992) now famous causal exclusion argument fails to knock down the levels doctrine, because it assumes a strong reading of what he calls Alexander's dictum, when a weaker reading is all that is needed to stave off epiphenomenalism.

II. The Metaphysics of Nonreductive Materialism

The view which I call, following Moser and Trout (1995), "nonreductive materialism" bears a number of interesting points of contact with British emergentism. The doctrines are so similar, in fact, that it is odd and a little disappointing that neither

Fodor nor Putnam thought to explain how their views relate to the once canonical work of Morgan, Broad, and Alexander. One obvious similarity between these different views is that according to both British emergentism and nonreductive materialism, the predicates of physics pick out base-level physical properties -e.g. the mass of atoms, the charge of electrons, etc.- and the predicates of the other sciences, the "special sciences," also pick out properties. To use some famous examples from the literature, both views supposedly imply that being in pain is a psychological property, being a dollar is an economic property, and being a square peg is a macro-level property. In other words, nonreductive materialism exhibits at least one defining feature of the levels doctrine: a relative abundance of properties.

As an aside, I strongly suspect the reason that Fodor and Putnam do not discuss emergentism is that once Nagel wrote *The Structure of Science* and Feigl and Smart had put forward identity theory, British Emergentism was considered dead in the water for the vast majority of philosophers. Moreover, emergence is rather obviously a robustly metaphysical notion. It is exactly the sort of idea that the anti-metaphysical movements of the early and mid-20th century were designed to do away with. So, perhaps Fodor and Putnam thought that they needed to contrast their view with identity theory and with Nagel's concept of "bridge laws" relating different scientific predicates, which is exactly what they do, and simply ignore emergentism as hopelessly passé.

Further evidence that nonreductive materialism is a form of the levels doctrine is that Fodor and Putnam are more than willing to use the language of levels. As we noted at the beginning of the first chapter, Fodor (1975) claims, "the world runs in parallel at many levels of description." He also claims that "there are special sciences not because

of the nature of our epistemic relation to the world, but because of the way the world is put together. In the Introduction to their anthology *Contemporary Materialism*, Moser and Trout (1995) suggest that Fodor has given us an argument for "shared higher-level property[s]… that makes [objects] instances of the same kind." And Putnam talks of levels frequently too, claiming that we need to posit, not only the properties of "the ultimate constituents" but also the properties of "higher-level structure."

I suppose someone could object to my linking Fodor and Putnam to British emergentism by claiming that all this levels talk is just loose talk and nonreductive materialism can be formulated without mention of higher-level properties.[60] I find it hard to believe that Fodor and Putnam could be so careless. But even if the levels talk is intended as a loose metaphor, the view that Fodor and Putnam are committed to simply has to be a version of the levels doctrine. They can't avoid this. You can't explain the metaphysics of nonreductive materialism without invoking a metaphysical relation which is, as I discuss in the next section, strikingly similar to emergence: the realization relation. Indeed, multiple realization is the *sine qua non* of nonreductive materialism. And, as we will see below, the existence of multiply realizable properties would necessitate the existence of levels. (A related objection is that nonreductive materialism is not the view that there are higher-level metaphysically real properties, but just the view that higher-level types can't be identified with microphysical types. That is to say, some see a failure to reduce the higher-level special sciences as only implying on that we have different ways of explaining the world which aren't translatable. In the next section of

[60] There's a way in which this objection is correct of course. Obviously, we don't need the word "levels" or even "higher-level property" when giving an explanation of nonreductive materialism. A term can always be avoided. If we wanted to, we could explain British emergentism without the word "level" too. That hardly gets at the substance of the view.

this chapter, I will briefly argue that while I am highly sympathetic to such a view, the view Fodor and Putnam are arguing for has to posit ontologically higher-level properties if it is going to do the work they want it to do.)

This brings us to the now celebrated concept of multiply realized properties ("MR" from here forward).[61] The MR doctrine is the claim that different objects can share a common property even if they don't share some common microphysical, base-level property. For example, suppose a small cloud of nitrogen in a bottle has the property of having such and such a temperature, and a large pool of water has the same property. Well, it certainly seems that the pool of water and the cloud don't share any microphysical similarities: the pool is made up of H20 molecules bound together in a liquid state, while the gas is composed of a different kind of atom altogether in a gaseous state. Therefore, it would seem like there is a sameness in properties between the water and the gas but no microphysical sameness.

The property of temperature may be a contentious example. It might seem odd to claim that temperature is MR and thus irreducible. Indeed, some might suggest that temperature, or at least mean kinetic energy, is a paradigmatic example of a microphysical feature. The problem with temperature is that it is not entirely clear if there is one unified microstructural property (presumably a relational property between atoms) that is picked out by the predicate "has a temperature of 200 degrees Kelvin." One might say having such and such a mean kinetic energy is multiply realizable by different clouds

[61] Fodor and Putnam themselves don't use the term "multiple realization." Nonetheless the concept is clear in both of their seminal papers on the subject. Fodor argues that a kind term in the special science can't be identified with a particular kind term in physics; at best, the special science term could be correlated with a heterogeneous disjunction of physical kind terms. He thereby avoids the term "realization," but nonetheless makes use of the concept. Putnam does use the term "realization" rather heavily, and can perhaps be given credit for popularizing it. He states that "Machines forced us to distinguish between an abstract structure and its concrete realization… the same structure could be realized in a bewildering variety of different ways; that the important properties were not physical-chemical." (p.129)

of atoms. I am not sure how to answer this question. I think the answer would require a robust account of exactly what kinetic energy is. If kinetic energy just is the average motion of molecules with certain mass, then I think we would have to say that having such and such kinetic energy is a good example of something multiply realizable by different clouds of atoms, because averages are particularly MR. On the other hand, if kinetic energy is something separate from the motion of the atoms, which then determines the motion of the atoms, then I might be inclined to say that there is a uniform realizer for all the different objects that have a certain well-defined temperature, i.e. different clouds of gasses could have the same level of energy which then determines the motions of the molecules. But this isn't relevant to the matter at hand.

The concept of MR properties yields what I call the multiple realization argument for the levels doctrine. It is really very simple, so simple and obvious, in fact, that it can seem foolish to deny it. It goes like this. We have examples of MR properties, and MR properties are, *ex hypothesi*, properties which aren't identical to any particular microphysical property. Therefore, we ought to conclude that microphysical properties aren't reducible to the microphysical or the material.

Like a number of other arguments we have discussed for levels, the argument from MR properties is abductive. It works by pointing to examples of properties and situations, and saying there is really no way of explaining these situations without denying that these properties are identical to or reducible to microphysical properties.

And putative examples of putative MR properties abound in philosophy.[62] For example, it is often said that pain is multiply realizable in a variety of different brain

[62] . As Moser and Trout (1995) point out, "compositional materialists" (p.8) believe every property or virtually every property is multiply realizable.

structures and can possibly be realized in circuit boards. Being a cell is multiply realizable by a number of different combinations of proteins and amino acids. And being a protein is multiply realizable by a variety of different chemical compounds, and chemical compounds are multiply realizable in a variety of different ionic forms. Thus, multiple realizability seems to track the hierarchy of levels that we are so familiar with from our discussion of emergence; e.g. psychology is multiply realized by biology, which is multiply realized by chemistry, which is multiply realized by physics. Interestingly, the clearest and strongest of examples of MR properties are often functional or computational in nature. For example, it is often claimed that a brain has the property of being able to compute 2+2, and so too does a small calculator. But the brain's computational abilities are, as nonreductive materialists like to say, realized by different mechanisms than the calculator's.[63]

Before I suggest what I find to be problematic in the argument from MR properties, it is important to examine exactly what the argument is supposed to prove. In other words, it is important to give a more robust account of the ontology of nonreductive materialism, which the argument from MR properties is supposed to support. As stated, the ontology of nonreductive materialism centers on the claim that there are higher-level properties which are realized by, but are not identical to, heterogeneous sets of lower-level properties. Thus, in order to explain the metaphysics of nonreductive materialism, we need to explain the metaphysics of the realization relation.

It is hard to say a whole lot more than that when giving an explanation of the metaphysics of realization. Perhaps this is surprising. Here's what we do know. For one

[63] In chapter 5, I argue that there is a reason why functional predicates seem like such good examples of MR properties. Functional predicates are often very useful in describing the world –and will remain projectable- even when they don't pick out the same property in every object they describe.

thing, while emergence is usually assumed to be 1-1 (one emergent property, one emergence base) and realizability is 1-many, it is difficult to discern any ontological differences between realization per se and emergence. In chapter two, we noted that emergence is an asymmetric, synchronic determination relation between properties. However, multiple realization, or just realization *simpliciter*, is also intended to be a kind of determination relation. The whole idea behind the concept of (multiply) realized properties is that realized properties are determined to exist by lower-level properties; e.g. if a sample of gas is in one of a number of different microphysical properties (or clusters of properties), this determines that the gas sample has such and such a temperature.

And obviously, the (multiple) realization relation is an asymmetric relation, just like emergence. When the right lower-level properties are present, this determines that a certain higher-level property will exist, but the presence of a higher-level property doesn't determine that a specific realizer property or cluster of realizer properties will be present in the same object. The determination goes from the lower-level property to the higher-level, but not vice versa, i.e. it is a one way determination relation. Moreover, realization must also be a synchronic relation, too. For example, we don't need to wait a second for a black, disc shaped object to turn into a hockey puck. It "happens" instantly. We don't have to wait for a change in voltage in a computer to become the addition of 2+2. That is, the higher-level property exists only at and at all times that its realizer property exists.

So, realization is a one-way, synchronic determination relation, and is thus very like emergence. How then is realization metaphysically different from emergence? How are emergent properties supposed to be different from realized properties? I take it that

there is supposed to be a difference in the metaphysical commitments of the two views, but the exact difference is very hard to tease out. The fact that it is difficult to tease out the differences between realization and emergence as ontological relations between higher and lower properties raises some suspicion about exactly how we are supposed to conceive of both. This is not a reason to reject either view. But it is a bit strange, given how standard it is to speak of both realization and emergence, that it is not exactly clear what we are discussing when we speak of either.

However, we can sketch some big picture, perhaps vague, differences between nonreductive materialism and emergence. Or, at least, we can sketch some ways in which nonreductive materialism is supposed to be different from its predecessor. As we saw in chapter two, traditional emergentism is supposed to occupy a conceptual middle ground between substance dualism on the one hand and eliminative mechanism on the other. That is to say, British emergentism clearly and explicitly asserts that mental, biological, and other higher-level properties are entirely real and distinct from collections of base-level properties. It is a metaphysically bold view in that way.[64] By comparison, nonreductive materialism is intended as a less bold doctrine. It is a less whole hearted, more hesitant assertion that there are higher-level properties. Perhaps this is ultimately a difference in how the two views are stated, i.e. in rhetoric but not substance. Nonetheless, it is helpful to think that nonreductive materialism is supposed to occupy a conceptual

[64] The emergentists were fairly clear that the fact that higher-level properties are anchored by the lower level properties does not imply that emergent properties are in any way identical to lower level properties But then again, there are times when Alexander sounds like he is advocating for the identity theory. And, as I explained in chapter 3, depending on how we conceive relational properties, Morgan's view might collapse into a uni-level view where higher-level relations are identical to lower level relations. So perhaps the British emergentists were guilty of oscillating back and forth between first identifying higher-level properties with lower level properties and then saying they were distinct, too. It even may be that, for some reason or another, the levels doctrine just tends to fall into this trap.

middle ground between British emergentism and eliminative mechanism, i.e. that nonreductive materialism is a sort of retreat from the strong claims of emergentism.

A side point, but a significant one, is that more recent versions of the levels doctrine, like Antony's (as we will see in the next chapter) are retreats even from Fodor and Putnam's views. And, I would maintain that they are just as, if not more, obscure as a result. For example, as I argued in the first chapter, Melnyk's (2003) "realization physicalism" is a very weak version of the levels doctrine. In fact, it is so weakly reductive that it posits higher-level properties in such a subtle way that it appears not to do so at all. All of these retreats make the levels doctrine an ever shifting target and, I think, a view that increasingly flirts with obscurantism. Indeed, even after you read volumes of material on multiply realizable properties and nonreductive materialism, it is hard to understand the answer to this central question: *what is it that isn't reduced in nonreductive materialism?*

Again, I am not arguing that pointing out that there is a seemingly unavoidable vagueness in the ontology of nonreductive materialism is a decisive objection against it. Nor do I think that pointing out that nonreductive materialism is supposed to be weaker than emergentism, even if it is not clear exactly how it is weaker, is decisive either. This is not a minor concession, because we could try to object to nonreductive materialism along these lines. Loar (1993) attempts to do something like this when he argues against supervenience based forms of nonreductive materialism. He suggests that multiply realized properties which supervene on base-level properties would have to have a sort of ill-defined "second string" status, which creates a problem for confirming when an object has such a property. I don't think these sorts of objections will ever be very successful

against the levels doctrine, though I am sympathetic to Loar's (1993) argument. As I emphasize in my discussion of Kim in the fourth section of this chapter, it really isn't so easy to prove that levels don't exist with an *a priori* argument showing that the levels doctrine is internally self-contradictory. It is easy, perhaps too easy, for proponents of levels to bite bullets or slightly alter the view to immunize it against this kind of attacks. Indeed, the levels doctrine is, as most metaphysical views are, rather malleable and can be –and, in fact, has been- reformulated to defend it against these kinds of objections. Thus, the abductive argument for the conclusion that we don't need levels is a wiser strategy to refute the levels doctrine in its various forms.

III. Properties, Types, Kinds, and Predicates

Someone could object that nonreductive materialism, particularly as it is formulated by Fodor and Putnam, does not assert that there are higher-level *properties* per se. Instead, it could be said that it is a thesis about higher-level *natural kinds.* This objection is right in at least one way. Fodor does, at least on some occasions, formulate his thesis in terms of higher-level natural kinds. But this objection is a problem for my portrayal of nonreductive materialism as a form of the levels doctrine that results from a lack of clarity in the metaphysics of nonreductive materialism.

Let's explore this objection a bit. I have tried to show how nonreductive materialism is built on the realization relation, which is a relation between real, existent properties. But maybe that is slightly wrong. Maybe nonreductive materialism asserts that there are higher-level natural kinds, not higher-level properties, per se. In other words, maybe there is some important difference between a property and a natural kind which I

am missing, which will bolster nonreductive materialism and immunize it against criticism.

Unfortunately, the views of the nonreductive materialists about these issues are hard to pin down. For example, Fodor (1975) oscillates back and forth between talk of "natural kinds," "properties," and "predicates," which makes his metaphysical presuppositions difficult to decipher. Here's one way of stating the problem. Suppose someone suggests that some kinds or types in the special sciences aren't reducible but all tokens are. This raises a question about the metaphysical status of a "type." Does "type" mean the same as "property" in the metaphysical sense we have been using it, i.e. as meaning the same as "feature" or "aspect." Or is "type" used in a more nominalistic way, to refer to a group of objects collected together and labeled with a descriptive term? If it is the latter, then a failure to reduce types or to identify them via one to one bridge laws with base-level types is perfectly compatible with the claim that there is nothing real, nothing which exists, other than base-level properties. That is, if there aren't real types, which exist over and above tokens, then nonreductive materialism might just be an epistemic or semantic thesis. But that is not enough for nonreductive materialism, as we saw above.

How, then, are we supposed to conceive of a "type" as an ontological entity that exists over and above its tokens? I think of property tokens as individual features of the world. More precisely, I can understand how individual features of objects are property tokens, and if you want to include objects in ontology, then individual objects are object tokens. Here is a hand. That is a token of a hand; it is an object token. Here is a particular color that my hand possesses. That specific color, that feature of my hand which we call

its color, is a property token. Thus, there doesn't seem to be any problem in conceiving of token properties or token objects at all. We can even point to the tokens: the hand and the color of the hand. So, the idea that types are real things that exist over and above real property tokens seems to imply that there is a very real thing, say, the type of being colored, that exists in addition to all of the specific instances or tokens of color, e.g. the color of my hand. At first glance, then, the belief that there are real, existent types of properties seems to collapse to belief in universals *ante res*, i.e. a robust view of Platonic universals. Pointing this out is, by no means, a strong reason to reject Fodor's arguments, but, once again, it does suggest that the position he is arguing for is not as clear as it could be. Surely Fodor doesn't think that types are universals *ante res*. But then, what are they, according to Fodor?

At any rate, here's a fairly clear statement of Fodor's (1975) thesis and some of the motivation behind it:

> "I am suggesting, roughly, that there are special sciences not because of the nature of our epistemic relation to the world, but because of the way the world is put together: not all natural kinds (not all the classes of things and events about which there are important, counterfactual supporting generalizations to make) are, or correspond to, physical natural kinds. A way of stating the classical reductionist view is that things which belong to different physical kinds *ipso facto* can have no projectible descriptions in common: that if X and y differ in those descriptions by virtue of which they fall under the proper laws of physics, they must differ in those descriptions by virtue of which they fall under any laws at all… Physics develops the taxonomy of its subject matter which best suits its purposes: the formulation of exceptionless laws which are basic in the several senses discussed above. But this is not the only taxonomy which may be required if the purposes of science in general are to be served: e.g., if we are to state such true counterfactual supporting generalizations as there are to state. So, there are special sciences, with their specialized taxonomies, in the business of stating some of these generalizations. If science is to be unified, then all such taxonomies must apply to *the same things.* [Fodor's emphasis] If physics is to be basic science, then each of these things had better be a physical thing. But it is not further required that the taxonomies which the special sciences employ must themselves reduce to the taxonomy of physics. It is not required, and it is probably not true." (63-64)

This is a remarkably impressive and sophisticated bit of philosophy, and there is a strong intuitive appeal to what Fodor is claiming. But the sophistication here has allowed Fodor to run together the notions of a "kind," a "type," and even a "set," along the lines I have

just suggested. This blurring of taxonomical kinds and ontological features is only exacerbated by Fodor's undefined use of the phrase "natural kind." Indeed, Fodor apparently uses "natural kind" to mean both "property" and "type" at the same time. For example, he writes that the "taxonomies" of the higher-level don't line up with the taxonomies deployed in physics. Thus, this seems to be a thesis about words or concepts, since taxonomies are human constructions. But earlier in this same passage, Fodor also seems to suggest that the taxonomical kinds deployed in the special sciences correspond to "ways the world is."

It helps to have working definitions of terms like "type" and "kind." I will take types, kinds, classifications, or taxonomical kinds to be collections of objects, any collection of objects. That is, we can think of a taxonomical kind as a category, a grouping of things: e.g. group of pigeons or a group of different visual representations of pigeons.

We need to be able to form these taxonomical kinds, these categories, in order to use and understand descriptive language. Once we have a category of things (or sense data) we then attach a lexical item like a name or a predicate to signify the type. I do not intend this remark to be at all controversial about semantics. I am just saying that a kind is a group of objects picked out by a word. I am staying at the naïve level of Locke; we could add more semantic detail if we needed to.

Of course, the semantic view of levels presented earlier says a little more about kinds than just this. The semantic view of levels -or rather semantic disjunctivism to be more precise- is premised on the claim that some taxonomical kinds needn't share any specific feature in common. For example, we have the ridiculously disjunctive

descriptive predicate "is a raven, a writing desk, a galactic cluster, or a pain." There is collection or set of objects described by this predicate, i.e. a taxonomical kind. Clearly this kind is as ontologically heterogeneous as it gets. Ravens and writing desks don't share some uniform property in virtue of which they rightly belong in the set. If they did, they wouldn't share that property with galaxies.

A property, by contrast, is a feature of objects.[65] A property is "a way the world is" or a way something in the world is modified. For example, some objects have the property of being a tau neutrino. Obviously, it is an open question as to how many properties there are and what sorts of properties there are. Some of these questions, as we have seen, are pertinent to the topic at hand, e.g. the sparse theory of properties and the existence of relational properties, other questions aren't, e.g. the dispute between universals and tropes.

One distinction between kinds and properties that falls out of these definitions is that taxonomical kinds are created by us. We group together objects into categories and we then attach terms to those categories. By contrast, properties are not created; they are mind independent. This distinction is important because if nonreductive materialism is a thesis about kinds or taxonomical categories, then it doesn't imply the existence of higher-level features of nature. Indeed, if nonreductive materialism is a thesis about taxonomies, then it would seem to accord rather well with the semantic view of levels.

I suppose it is possible that Fodor does not think that there is a simple, clear cut difference between the real features objects and the taxonomical kinds we create. It is

[65] As we saw in the first chapter, we don't need a sophisticated notion of properties here; we don't need to settle questions in the universals versus tropes. We just need to say that properties are features of objects. We are immediately aware of some features in introspecting our own experience. We're aware of pain as a qualitative feature of our experience. We can only assume, despite skeptical worries about the existence of the external world, that the physical world has at least some features, too.

hard to know. In the passage I have quoted, he says we have special sciences and the taxonomy of the special sciences because of "the way the world is." That certainly sounds as if he thinks there are features of the world that are not identical to microphysical features or even mereological sums of microphysical features. But then, at the end of this passage, he seems to take it back. He writes that the "same things" are taxonomized by the special sciences and by physical kinds.

Fodor's failure to distinguish kinds and types from real properties is, I argue, a crucial oversight and is, as we will see below, the source of the ontological obscurity of all of the different forms of nonreductive materialism. The uni-level theorist can agree with Fodor or with nonreductive materialists in general, if their thesis is just that it is difficult or even impossible for us to take the very messy and heterogeneous categories of economics or paleontology and then find some wildly disjunctive microphysical predicate with exactly the same extension.

In that trivial sense, we obviously can't reduce higher-level predicates to lower-level predicates. But, the failure of this kind of reduction isn't really at all surprising or interesting and it doesn't have the ontological implications Fodor suggests it does. It shows only that the strict type-type bridge laws that Nagel (1961) argued were the basis for reduction were never necessary for reduction.

What is necessary for reduction is that we have some way available to us for explaining how particular instances of a putative higher-level property are actually identical to microphysical properties or cluster of microphysical properties. If we had such a method, we could then proceed with a method of examining instances of higher-level properties, one by one, showing each to be identical to some collection of lower-

level properties or to some individual lower-level property. There is no requirement that we have some method for showing how all instances of a putative higher-level property are identical to some one of a number of microphysical property or cluster of properties with some one-to-one bridge law. Instead, we can make do with a sort of reductive schema or model, which would show us how a paradigm example of some individual putative higher-level properties is identical to a cluster of base-level properties, and which would suggest an avenue for how we might go about dealing with cases other than the paradigm. Once we had that schema, we could proceed case by case, showing how each instance of some putative higher level property is identical to something at the base level.

So Fodor's claim that the fact that we don't have type-type bridge laws between higher and lower-level taxonomical kinds has interesting and important metaphysical consequences is false, and it becomes easier to see how it is false when we are clearer about the meaning of terms like "predicate" and "property."[66]

Again, no matter how nonreductive materialism is formulated, it has to be conceived of as a metaphysical thesis and a form of levels doctrine. It *has* to be this way, because if nonreductive materialism doesn't assert the existence of higher-level properties, then it loses its biggest selling feature: the fact that it offers a possible solution to the mind-body problem. If there really are higher-level properties which are realized by, but distinct from, lower-level properties, then perhaps mental properties are one such example, i.e. perhaps mental properties are realized by but distinct from neurological properties. But, if realization should be conceived as an epistemological or linguistic relationship between predicates or kinds, with no metaphysical implications about the

66

existence of realized properties, as I suggest, then nonreductive materialism cannot offer a solution to the mind-body problem. The mind-body problem is ultimately a metaphysical question about what relationship mental properties or mental states have to neurological, or even physical properties. It is immune to any linguistic cure. In other words, if nonreductive materialism can explain the very the things it is ultimately designed to explain, then it has to be robustly metaphysical.

As an illustrative aside, I would like to point out that this confusion over higher-level properties and higher-level predicates isn't found only in Fodor's nonreductive materialism. A good deal of the literature on supervenience, nonreductivism, and levels is somewhat unclear on predicates and properties in roughly the same way. As a result it is not always clear whether various forms of supervenience-based formulations of physicalism and materialism are covert versions of the levels doctrine or double covert versions of the sort of semantic view of levels laid out in the first chapter.

For example, it is sometimes hard to tell whether Davidson's (1970) now seminal view in "Mental Events" is committed to the existence of higher-level properties. Certainly, Davidson's view is normally construed as being in favor of higher-level, autonomous, irreducible mental events. However, I would argue that Davidson is arguing only for a linguistic or epistemological thesis that is not inconsistent with the semantic view of levels.[67] There is some textual evidence for this. For example, Davidson (1970) claims that he is arguing for "a version of the identity theory" (p.108) and that the absence of strict "psychophysical laws," i.e. one to one bridge laws, does not imply that the identity theory is false. He also claims that "Anomolous monism shows an

[67] John Heil rightly pointed me in this direction.

ontological bias only in that it allows the possibility that not all events are mental, while insisting that all events are physical." (p.112)

On the other hand, he also claims that "such a bland monism, unbuttressed by correlating laws or conceptual economies, does not seem to merit the term "reductionism""(p.112) Indeed, he claims that the mental is autonomous and irreducible, and at best supervenient on the physical. This would seem to put Davidson in the same camp with Fodor and Putnam.

A brief excursion into what Davidson actually says, though, suggests he is better seen as an early advocate of what I have called the semantic view of levels. Davidson (1970) wants to argue that the following three statements are not, despite appearances, mutually inconsistent. 1.) Mental events are caused by physical events and vice versa, at least on some occasions. 2.) Causation must be governed by laws. 3.) There are no strict psychophysical laws, "on the basis of which mental events can be predicted and explained."

He attempts to smooth over the apparent contradiction in these three statements by admitting that mental events can't always be fully explained in physical terms with strict psychophysical laws. Rather, the latter supervene on or "are in some sense dependent on" the former. (p.112) (Note the vagueness in Davidson's "in some sense dependent on.") That is, there are no *strict* psychophysical laws which we can point to in order to fully explain the mental in terms of the physical; there is just a non-explanatory dependence relation, i.e. supervenience. However, Davidson argues that the lack of strict psychophysical laws doesn't imply that it is false that "all events are physical" (p.112) To explain this, he writes

> It should now be evident how anomalous monism reconciles the three original principles. Causality and identity are relations between individualized events no matter how described. *But laws are linguistic*; and so events instantiate laws, and hence can be explained or predicted in the light of laws, *only as those events are described in one way or another.* The principle of causal interaction deals with events in extension and is therefore *blind to the mental-physical* dichotomy. *The principle of anomalism of the mental concerns events described as mental, for events are mental only as described.* The principle of the nomological character of causality must be read carefully: it says that when events are related as cause and effect, *they have descriptions that instantiate a law.* It does not say that every true singular statement of causality instantiates a law. (p.112-113) [italics are mine]

In the simplest possible terms, it seems that Davidson is suggesting here that as long as we construe the failure of one to one reduction as a linguistic problem, i.e. as a failure to find physical predicates that can be correlated with mental predicates, we don't need to give up on monism, physicalism, or materialism. As he puts it, the problem is just that "mental and physical predicates are not made for one another."

Unfortunately, Davidson goes farther than this point, suggesting that because mental predicates display intentionality while physical predicates do not, we cannot translate mental descriptions to physical transcriptions. He then adds, without much explanation, that he wants his view to be akin to Quine's "indeterminacy of translation."(p.118) The point though is clear enough. Davidson is pessimistic that we will ever find a set of physical predicates that pick out the same properties as mental predicates. However, I can find no real reason for Davidson's pessimism in what he says other than the standard worries about identifying intentional states with physical properties.

Moreover, a gradualistic approach, along the lines that Martin (2008) takes, can resolve these worries and can dispel pessimism about reducing intentional states to physical states. If you can explain how a simple Coke machine selects for certain inputs, e.g. American quarters but not Canadian quarters, in terms of physical properties of the Coke machine's electronics, and you can explain the more complex selectivity of, say,

chess playing computers, in terms of more complex physical circuits, then you can also explain the intentionality of desires, emotions, and even beliefs in similar terms. But this discussion of intentionality will take us too far afield into specific questions about the mind, when we are focused on levels more generally.

While Davidson's view is clear enough, though it has been misinterpreted as a version of the levels doctrine, other views are obscure versions of the levels doctrine, whose evidence relies on conflating predicates and properties. One such view is William Wimsatt's (1994).[68]

The core of Wimsatt's view is what he calls "robustness." In and of itself, robustness is nothing unusual or controversial. According to Wimsatt, when there is more than one useful way of describing or modeling some entity or event, then there is robustness. So, for example, psychiatric illnesses are robust in that there are psychological and physical descriptions (or models) of them. We simply can't deny that the sort of robustness Wimsatt is describing does occur. However, we ought to remind ourselves that it is only obvious that theories, models, and descriptions are robust. It remains, as we have seen repeatedly in the past few chapters, an open question as to whether things themselves are robust. To say otherwise, without more evidence, would be a question begging argument in favor of the levels doctrine.

A little more controversially, Wimsatt then suggests that robustness is a "criterion for what is real." (p.1) It seems that his idea is that if there are a bunch of useful descriptions or models describing a certain thing, this is evidence that the phenomenon so described is real. (We will discuss a similar, though clearer, criteria for "what is real" in

[68] Another example, which Wimsatt points out he is largely in agreement with, is John Dupre's (1993) view.

connection with the claim that projectable predicates must pick out unique properties in the final section of chapter five.) He (1994) writes,

> Before I say what there is in this complex world, I should give my criteria for regarding something as real or trustworthy. Particularly among those of a foundationalist persuasion, it is common to start by providing some criterion, be it indubitability, incorrigibility, or other means of picking out things or assumptions whose veracity is not open to question. One then says that those things are real (true, indubitable, or whatever) if it is either one of these primitive things or if it is derivable from them via a valid series of inferences. Only things admitted in one of these two ways are allowed. I share the foundationalist's concern with securing reliability for our conceptual structures. But I don't think that there are any criteria which both give indubitability or render error impossible, and permit any interesting inferences from that starting point. Thus, I would rather give a criterion which offers relative reliability, one that you're better off using than not, indeed better off using it than any other, and which seems to have a number of the right properties to build upon. Rather than opting for a global or metaphysical realism (an aim which bedevils most of the analyses of "scientific realists"), I want criteria for what is real which are decidedly local—which are the kinds of criteria used by working scientists in deciding whether results are "real" or artifactual, trustworthy or untrustworthy, "objective" or "subjective" (in contexts where the latter is legitimately criticized--which is not everywhere). When this criterion is used, eliminative reductionism is seen as generally unsound, and entities at a variety of levels--as well as the levels themselves--can be recognized for the real objects they are, and traditional foundationalism and ontic fundamentalism are in trouble. They will survive, if at all, as a local kind of problem-solving technique of significant but limited usefulness… Following Levins (1966), I call this criterion robustness. (See Wimsatt, 1981a, for an analysis and review of the concept and methodology, 1980a, 1980b, for relevant case studies, Campbell, 1966, whose concept of "triangulation" captures many of the same ideas, and whose classic work with Fiske (1959) on the "multi-trait-multi-method matrix" brought this methodology to the social sciences). Things are robust if they are accessible (detectable, measureable, derivable, defineable, produceable, or the like) in a variety of independent ways. A related but narrower criterion (experimental manipulability via different means) has since been suggested by Hacking (1983), who draws a close link with experiment, and limits his discussions to the realism of entities. But robustness plays a similar role also in the judgement of properties, relations, and even propositions, as well as for the larger structures--levels and perspectives--described below (see Wimsatt, 1981a, and also 1974, 1976a). Furthermore, the independent means of access are not limited to experimental manipulations but can range all the way from non-interventive observation or measurement to mathematical or logical derivation, with many stops in between. Experimental manipulation is just a special case. We feel more confident of objects, properties, relationships, etc. which we can detect, derive, measure, or observe in a variety of independent ways because the chance that we could be simultaneously wrong in each of these ways declines with the number of independent checks we have.[2] We can only make the probability of failure decline--though it can get very small, it does not go to zero. This criterion does not give certainty. Nothing does. There are no magic bullets in science--or anywhere else, for that matter. But if that's so, then certainty is not so important as generations of philosophers have supposed.

In the simplest terms, Wimsatt's argument seems to be that if we accept robustness as a criterion for what is real, then eliminative reductionism must be "unsound." But he never really explains why we should accept this implication.

The only thing I can think of is that Wimsatt has blurred the distinction between predicates and properties such that he does not see his position as question begging. And indeed, if there is no such distinction, then perhaps Wimsatt's argument is sound, though also trivial. But if we are right in what we have said so far, then there is a difference between our descriptive predicates and what they describe. Thus, Wimsatt can only help himself to the premise that predicates deployed in theoretical descriptions are robust. But instead, he sometimes tries to appeal to a metaphysical conception of robustness as a premise in his argument form higher level properties. He describes it in the premise of his argument as a relationship between properties. But if his premise is that properties are robust, then his premise itself asserts that there are levels. And his argument for the conclusion "there are many higher-levels of properties corresponding to the hierarchy of descriptions we deploy in the sciences" would have to be question begging.

Perhaps, the references to the pragmatic usefulness of robustness as a criterion for reality are supposed to yield an argument that isn't question begging. That is, perhaps Wimsatt wants to argue that it is useful for scientists to believe in an ontology of levels, and therefore the ontology has evidence in its favor. But Wimsatt never establishes that the ontological doctrine that there are higher-level properties has pragmatic value for the sciences. How could he? At best, he establishes that it is useful for scientists to deploy higher-level descriptions and higher-level predicates. Indeed, no one denies that it is useful to have, say, a biological and a psychological model of mental illnesses, but whu should we take this to imply that the two different descriptions don't describe the same thing?

Perhaps this is unfair. It may be that I simply don't understand Wimsatt's argument. However, there is some textual evidence to suggest he has conflated predicates and properties. For example, in his explanation of what a level is, he writes,

> Ontologically, one could take the primary working matter of the world to be causal relationships, which are connected to one another in a variety of ways and together make up patterns of causal networks. (I won't address problems with causality in this essay. Those who favor "Humean scepticism" will also find lots else to object to here, and can stop reading now unless they want to see how far you can get without it!) These networks should be viewed as a sort of bulk causal matter--an undifferentiated tissue of causal structures--in effect the biochemical pathways of the world, whose topology, under some global constraints, yields interesting forms. Under some conditions, these networks are organized into larger patterns which comprise levels of organization, and under somewhat different conditions they yield the kinds of systematic slices across which I have called perspectives. Under some conditions, they are so richly connected that neither perspectives nor levels seem to capture their organization, and for this condition, I have coined the term "causal thickets". (p.5)

Here Wimsatt distinguishes between the "bulk" causal relationships and how they are organized. He then wants to say that levels just are the organization of the causal relationships. But this is odd given his stated realism about levels. Isn't it possible that all that exists are the causal relationships, and how we organize them is a matter of how we describe them? If so, then it is not clear why Wimsatt would remain a realist about levels. He never rules any of this out, and in not ruling it out he puts his ontology in a sort of interesting but unclear middle ground between the levels doctrine and the semantic view of levels.

IV. The Failure of the Causal Exclusion Argument

It is standard in philosophy to think that the best hope anyone has of disproving the existence of irreducibly higher-level MR properties is found in Kim's (1989 and 1992) causal exclusion argument. However, as I argue below, Kim's causal exclusion argument is insufficient for knocking down the levels argument. And the failure of Kim's

a priori criticisms suggests the sort of *a posteriori*, abductive approach I have chosen is more likely to be fruitful.

A brief review of the causal exclusion argument is in order. The argument is designed to show that there can be no such thing as an irreducibly higher-level MR property. It begins with a truism. Either higher-level properties come prepackaged with causal powers or they don't. If higher-level properties don't come pre-packaged with causal powers then they are mere epiphenomena. And we cannot claim that mere epiphenomenal properties exist, because if something doesn't have any causal powers, it wouldn't be able to affect us, i.e. we couldn't detect it, and so we would never have reason to say that a truly epiphenomenal property existed. This idea that properties have to come prepackaged with causal powers is what Kim calls "Alexander's dictum," alluding to the very same Samuel Alexander we discussed in connection with emergence.

Now, on the other hand, if higher-level properties do come prepackaged with causal powers, i.e. if they do make a contribution to what the object possessing them can do, then either those causal powers exist in addition to the causal powers associated with the base-level properties, or they are identical to the causal powers at the lower-level. If higher-level properties endow objects with novel, higher-level causal powers, then this, Kim argues, entails the causal overdetermination of the physical. (I discuss this crux at greater length below.) According to Kim, what happens at the physical level is fully determined by prior events at the physical level. For example, if a molecule is accelerated at a rate of 1 meter per second per second, that event must have been determined by some purely microphysical cause, e.g. a collision with some other particle. There is no room for anything to else to be involved in causing the acceleration of the molecule because the

microphysical collision is sufficient for causing the acceleration, and it is necessary that the acceleration have some microphysical cause: e.g. a collision, gravitational pull, etc. It is hard to deny that the physical is causally closed in this way. The problem that plagued Descartes' causal interactive dualism was that it didn't seem reasonable to claim that mental substances could make an impact on physical bodies. The physical world is closed to the impacts of Cartesian souls. Higher-level properties, according to Kim, have a similar problem.

Now, the trickiest part of the causal exclusion argument is proving that the existence of higher-level causal powers would imply a violation of the causal closure of the physical. (And this is where the argument fails.) Let's start with what we know. We know that if there are higher-level causal powers, then these powers must make a difference at the higher-level. That is to say, there must be intra-level causation at the higher-levels, if there are higher-level powers, e.g. pains have the power at the psychological level to make us want an Advil. (Certainly, it wouldn't make much sense to say that pains have causal powers but they don't have the ability to affect changes at the mental level, i.e. that pains couldn't change our desires or emotions.) Moreover, we know that realization is a sort of "upward causation."[69] If the right lower-level, realizer property or properties are present, e.g. a complex set of neurological properties, then a particular realized property is now determined to be present.

[69] In chapter two, we discussed whether there is a difference between emergence and ordinary efficient causation. We determined that the only possible difference is that ordinary intra-level causation is diachronous, while emergence is instantaneous. However, we also realized that some intra-level causation is instantaneous, too, and we decided it was difficult to distinguish intra-level causation from emergence. Nonetheless, we allowed that there *could* be such a distinction. All of the same applies to realization and intra-level causation.

But Kim argues that is all you need to prove that higher-level powers entail violations of causal closure. How so? Suppose something has the psychological level realized property of being in pain, because it possesses neurological properties A and B. Well, those neurological properties also have to come with intra-level causal powers, i.e. the neurological properties have to make an impact at the neurological level. Suppose, upon closer investigation that we discover that neurological properties A and B can cause neurological properties C and D, which are, let's say, realizers for the psychological property of wanting to take an Advil. If A and B didn't bring about the instantiation of some neurological property that instantiated wanting an Advil, and wanting an Advil is caused by being in pain, then wanting an Advil can occur without any neurological cause. But that can only be true if some mysterious form of dualism is true, i.e. if there can be changes in the mind with no neurological cause. And no one in with a dog in this fight over levels wants that.

There are two stories we can now tell about what caused the psychological property of wanting to take an Advil to be instantiated. On the first story, the fact that I am in pain is explained by the fact that I, or my brain, possess neurological properties A and B, and those neurological properties caused neurological properties C and D, which in turn caused me to want an Advil via upward causation But, on this story, pain is a property, and so pain must be associated with a set of causal powers. Therefore, on this story, pain must also have caused the wanting of an Advil. But that means the psychological property of wanting an Advil is causally overdetermined. It is caused at the psychological level by the property of being in pain and its realizer is caused at the neurological level, which in turn causes the higher-level property to be realized as well.

Moreover, the neurological property that realizes the wanting of an Advil is over determined for the same reasons: the realized property is caused to occur at the higher-level, which in turn requires the presence of some realizer property. This is downward causation at work, and this downward overdetermination cannot occur, because it implies the causal overdetermination of base-level events and properties. And that constitutes a violation of the causal closure of the physical, which, we have assumed, does not happen. So this story can't be the truth.

The other story we can tell is the one Kim gives. He suggests that the causal powers of a particular realized property are identical to the causal powers of the property or properties that realize it. So, for example, the causal powers that come prepackaged with my pain just are the causal powers of the neurological realizers of my pain. The advantage of this story is that it avoids causal overdetermination. On this story, the causal work done by the neurological realizer properties A and B *is the same causal work* done by pain. For example, if all of the brains that realize being in pain share a common neurological level causal power, let's say, the power to bring about some common neurological property X, then all those brains share a common property. After all, if properties are identified and individuated by their causal profiles, *a la* Alexander's dictum, then sharing a common causal power implies sharing a common, uniform realizer property. On the other hand, if the neurological properties have different causal profiles, we can conclude that they are different properties. And unless we want to say that pain has a higher-level power, which will threaten causal over determination and a violation of the causal closure of the physical, then we have to admit that there is no such thing as the

property of pain, but rather there are many different properties, each associated with a different set of causal powers, and they are all called “pains.”

One response to Kim’s causal exclusion argument, then, is to just bite the bullet and say, “Okay, so there is causal overdetermination, and violations of the causal closure of the physical. So what?” Of course, Kim will respond that violations of the causal closure of the physical are just as mysterious, magical, and unbelievable as the claim that Cartesian souls can affect matter. Again, that second bullet doesn’t seem like a bullet that should be bitten.

However, we should note that although the kind of causal overdetermination which Kim thinks is implied if higher-level MR properties have causal powers does imply that higher-level changes can bring about a base-level, microphysical change, it doesn’t imply that higher-level changes can occur without a concomitant lower-level change. Kim’s critics can agree that it would be very strange indeed if a higher-level property like the psychological property of wanting a donut could bring about a change in the physical world without effecting a base-level change, say a change in my brain-chemistry, which caused my hand to move, which caused the donut to accelerate towards my mouth. I can’t just want a donut and have the telekinetic power of my wanting affect a change in the position and motion of the molecules constituting the donut. That kind of magical, telekinetic change in the molecules making up donut would be a truly gross violation of the causal closure of the physical. But, it is less clear how gross a violation of the causal closure of the physical if the donut is moved towards my mouth by my wants, and the physical changes in my brain and the nerves connected to my arm, which instantiate the property of wanting a donut. What this shows is that if someone wants to

defuse the causal exclusion argument, they have to differentiate problematic violations of the causal closure of the physical, like Cartesian souls telekinetically affecting donut molecules, from non-problematic ones, like souls telekinetically affecting arm molecules, and this doesn't seem easy to do, at least not in some principled question begging way. (That is, perhaps someone can bite the first bullet, i.e. accept causal overdetermination, without biting the second, i.e. a violation of the causal closure of the physical.)

A better, though related, objection is to say that Kim has misconstrued the relationship between causal powers, laws, and properties. I think the best way for Kim's critics to do this is to differentiate between a strong and a weak reading of Alexander's dictum (AD) as follows:

Strong AD: Every property has its own unique causal profile, i.e. set of causal powers. That is, no two different properties share a common causal profile. This is a one causal power-one property view.

Weak AD: Every property has a causal profile. Perhaps some different properties have the same causal profile. This is a one causal power – many properties view.

Notice that weak AD is enough to stave off epiphenomenal properties. Weak AD states that no property can be instantiated if it doesn't have some power to effect changes in the world or changes in us. The problem with truly epiphenomenal properties is that if, *ex hypothesi*, they can affect no changes at all, then they can affect no changes in us or in our observations, and we could never have reason to believe in them. Weak AD bars such unobservable epiphenomenal entities, because it asserts that all properties are instantiated with causal powers.

Interestingly, it is fairly clear Alexander himself accepted only weak AD. Thus, it is somewhat ironic that Kim dubs the stronger reading of the relationship between powers and properties "Alexander's dictum." In the third chapter, we examined Alexander's ontology of emergent qualities and we determined that Alexander's ontology consists of a hierarchy of different qualitative properties, with no new causal powers. For example, on Alexander's view my brain possesses microphysical qualities, chemical qualities, biological qualities, and mental qualities. But the "motions" or the changes in the qualities possessed my brain are the same, i.e. there is one set of motions or events in my brain, resulting in changes at each of the levels. In this way, Alexander's ontology, and the levels doctrine more generally, are quite explicitly and quite heavily indebted to Spinoza's causal parallelism.[70]

Indeed, Spinoza would not have accepted strong AD, but he still could've and would've accepted Weak AD. For Spinoza, the events that occur in God or Nature occur in at least two different attributes at once: the physical and the mental. Spinoza writes:

> A circle existing in nature and the idea of the existing circle, which is also in God, are one and the same thing, which is explained through different attributes. Therefore, whether we conceive nature under the attribute of Extension, or under the attribute of Thought, or under any other attribute, we shall find one and the same order, or one and the same connection of causes, i.e., that the same things follow one another. (IIp7s)

So, if Spinoza's metaphysics is at least possible –never mind well grounded empirically- then there is no reason to think that we need strong AD. That is to say, if we can conceive of mental and physical attributes that are qualitatively distinct, but identical in terms of their causal powers, then the strong reading of AD might very well be too strong.

[70] As I mentioned in chapter 3, it is regrettable that recent work, see for example McGlaughlin (1992) on Alexander and Broad fail to mention Spinoza given the important similarities between Spinoza's ontology and the levels doctrine and explicit discussions of Spinoza by Alexander and Broad.

Moreover, this is roughly the tack Antony takes in pushing back against the causal exclusion argument. Antony writes (2003):

> "… mental properties can be causally efficacious: because a mental property is instantiated via the instantiation of a physical property, the mental property ''inherits'' all the causal powers associated with the realizing physical property. (p.1)

She goes on to explain that the powers associated with higher-level powers are not mere "bogus" powers, i.e. there are powers that are associated with higher-level properties that are not identical to the powers associated with any particular physical realizer property. She thinks the powers associated with, say, being pain are not identical to the powers associated with any of the particular realizers of pain, e.g. with the physico-chemical powers associated with my C-fiber firings or your C-fiber firings. Of course, she does think that any given instance of pain will "inherit" its causal powers from its realizer(s). (This is how she deflects the causal overdetermination problem.) Granted, this may seem self-contradictory to someone who is tempted to agree with Kim or anyone who accepts Strong AD. But Antony attempts to smooth away this apparent contradiction with an appeal to Shoemaker's (2007) distinction between a "core realizer" and a total realization. The core realization of pain, in my case, is my C-fibers. In a Martian, it will be the filling of cavities in his feet with fluid. Of course, a set of C-fibers firing in a Petri dish will not instantiate pain. Pain is only instantiated by C-fiber firings that are properly hooked up to an organism. The C-fiber firings, when hooked up, are a total realizer of pain, while the C-fibers themselves are a core realizer. She (2003) then explains that,

> It's true enough that for a property to be truly ''multiply'' realizable, its realizers must differ from each other substantially in respect of total causal powers. But it doesn't follow from this that the realizers can't share a significant set of causal powers. In fact, we can see from the case of physical properties that this must be so. SILVER is a paradigmatically physical property, Kim would agree. And yet, instances of SILVER differ widely with respect to their causal powers. Silver coins, in sufficient numbers, have the power to free a candy bar from a vending machine; my silver ring lacks this power, but compensates by being able to encircle my finger. But heterogeneity of causal powers among instances hardly shows that SILVER isn't a perfectly respectable property with its own distinctive set of causal powers. What are they? To find out, we

> could take the intersection of the causal powers of all things silver. Now consider the question whether BEING SILVER adds anything to the causal powers of some instance of silver. BEING SILVER is determined by the atomic structure of the molecules that compose the thing in question. But atomic structure is an abstract property. So there will be a multitude of particular molecular arrangements alike only in the abstract respect of sharing atomic structure. In any given instance, given the particular molecular assembly that constitutes that instance, the specific causal powers of that instance are fully determined. What does or could BEING SILVER add? The question is simply a vexed one. The causal powers associated with any properties above the level of fundamental physics all supervene on the properties of the fundamental particles that compose their instances. Higher-order properties are not distinguished in this way from any others above the fundamental level. (p. 18-19)

So the powers we associate with being silver cannot be identified with the powers of any particular microphysical property. Nonetheless, these powers are always inherited from one of the core realizers of the property of being silver.

Again, Antony's account of powers will not sit well with anyone who is prone to accept Kim's position. But unfortunately for Kim and his allies, the burden of proof is on them to show that Antony's position on powers is somehow untenable or internally inconsistent. That is a heavy burden because it requires arguing for a rather strong thesis about the relationship between powers and properties.

Because Kim doesn't give us any good reason to favor strong AD over weak AD, Kim's causal exclusion argument doesn't provide us with a reason to not accept Alexander's emergent qualities nor does it give us a reason to deny favor Antony's version of the levels doctrine. What this means is that the causal exclusion argument doesn't rule out, on its own anyway, the possibility of a multi-level ontology consistent with the causal closure of the physical. Kim's critics only need to say that Kim's argument leaves some conceptual ground to assert the existence of higher-level properties that share powers with base-level properties.

There is one avenue still open for Kim or his allies. Kim could argue that accepting weak AD but not strong AD makes a mystery out of how we go about

distinguishing any two different properties that share a set of causal powers. For example, if the property of being a cell is a different property from being a certain collection of proteins and amino acids, even though the causal powers of a cell are identical to the causal powers of the amino acids and proteins, then it is not clear *how* we go about identifying the two different properties. In other words, perhaps weak AD suffers from an analog to the problem afflicting epiphenomenalism. Epiphenomenal properties, by definition, don't have any causal powers, so that they can't tell us that they exist, i.e. we would have no criteria for determining when such a property is present. By analogy, how does the sort of higher-level property Antony is arguing for tell us that it exists? If it does so by exercising the causal powers associated with its base-level realizer properties, then how does the higher-level property tell us that it is in any way distinct from those base-level realizer properties? In other words, what criteria do we use to determine when two properties or three properties are inherent in an object when that object only has one set of causal powers?

Kim's critics will respond to this by saying that we just do distinguish higher-level properties from lower-level properties, presumably by their qualities. Kim will say that we obviously can't distinguish properties without noting differences in causal powers, and that anyone who denies this is risking epiphenomenalism. And Kim will argue that this is more evidence that strong AD, and not just weak AD, is true. This appears to be another case where one man's *modus ponens* is another man's *modus tollens*. Moreover, there doesn't seem to be a clear, principled way of figuring out who is right.

I have some sympathy for Kim's position and for the idea that strong AD must be true. But his advocacy of strong AD hits a massive problem when we begin discussing mental qualities, i.e. qualia, and the physical properties that supposedly realize them. Through introspection, I can know, for example, that my headache is very painful. And I can distinguish being in pain from pleasure and other experiences. Notice, I don't need to know anything about the causal profile associated with the property of being in pain or the neurological properties that supposedly realize my pain. Indeed, even if I know everything about the causal powers of the neurons realizing my being in pain and the causal powers associated with the property of being in pain, I can conceptually distinguish the property of being in pain from any cluster of neurological properties. This alone suggests that we don't always identify a property by its causal powers. At the very least, we don't determine when a certain property is present in our experiences by looking at its causal profile. The experience and the quality itself are simply present to us in introspection.

I suspect some believers in the levels doctrine believe all higher-level properties are just like mental properties in this way. That is to say, introspection presents mental properties as being different from physical or neurological properties, and nature presents higher-level properties in general as different from the lower-level realizer properties from which they emerge. Indeed, this is very likely what Alexander (1918) is trying to express when he states that we must accept that there are higher-level qualities with "natural piety," i.e. Alexander thinks that nature just shows us that there is some difference between, say, biological and chemical properties even if we are not sure how to explain the difference.

But even if this claim seems untenable with respect to biological properties or chemical properties, it is certainly tenable with respect to mental properties. And that is where Kim runs into trouble. Mental properties appear to have something in them which we know about independent of what we know about the causal profile associated with pain or its realizers.

Perhaps this is why Kim (2005), in *Physicalism or Something Near Enough,* gives some ground to nonreductivism and allows that mental properties exist over and above their physical realizers. (That is, he ultimately allows that qualia are epiphenomenal.) In the simplest terms, Kim says that we don't identify the qualitative feel of pain with its causal profile, so therefore it can't be like any other property, it can't make any sort of causal contribution whatsoever, and so it must therefore be a sort of weird epiphenomenal property. Of course, admitting epiphenomenalism is a massive concession to make just to knock down the levels doctrine. I would point out that the abductive argument against levels of properties we have offered here does not require us to posit epiphenomenal properties.

It is here that the hidden weakness of Kim's causal exclusion argument shines through. Kim tries to show that there can be no MR properties *a priori* given a certain conception of the relationship between properties and causal powers, which is speculative, and which is very controversial when applied to mental properties. Indeed, this relationship between properties and powers is a deep and difficult matter, especially when it comes to mental properties. Kim's critics can push back and argue that mental properties do exist and we do identify them independently of their causal profile. If we follow Kim and his critics in this debate, the question of whether the levels doctrine is

true will turn on some very abstruse questions about causal powers, qualitative properties, and the very tricky subject of mental causation.

Summing all this up, it looks like Kim's causal exclusion argument only works if we accept some strong metaphysical claims, i.e. that strong AD is true. The abductive strategy we are deploying does not make any such claim. The particular metaphysical claims Kim makes create problems for him down the road, i.e. he ends up having a problem with qualia and strong AD, because qualia aren't distinguished by their causal powers. The abductive strategy we are deploying has no such problems because it does not make strong metaphysical claims. Indeed, the abductive strategy we are deploying makes no metaphysical assumptions that the levels doctrine does not. The only assumptions that are needed to show that higher-level properties are not necessary are the simple, plausible semantic assumptions we have made, e.g. semantic conjunctivism and semantic disjunctivism.

V. Semantic Disjunctivism Versus Multiple Realization

In the first section of this chapter we discussed multiple realization and presented the argument from MR properties in broad terms. But just to review, the argument goes like this:

P1: There are MR properties

P2: A multiply realizable property cannot be identical to or reduced to any microphysical property.

C: There are properties that aren't identical to any microphysical property.

The argument is obviously valid. Moreover, I think the second premise, or at least something similar, must be true for a variety of reasons. One such reason is just that multiple realization is a one-many relation and identity is not, and so realization cannot be the same thing as identity. If P can bear the "is realized by" relation to Q and R, and Q is not identical to R, then "is realized by" cannot be identity, because identity is transitive.

That leaves only one possible avenue we could take that would allow us to deny the conclusion: perhaps P1 is false. Thus, to undermine the argument from multiple realizability we must show that the first premise is itself unwarranted. At first glance that can seem like a rather odd strategy. MR properties *seem* to exist. That is to say, we have perfectly good predicates like "is an axe," "is a protein," and "computes 2+2." Indeed, someone might scoff at any attack on premise 1, saying "Surely these predicates, we can call them MR predicates, pick out properties. After all, MR predicates describe the world successfully, just as successfully as microphysical predicates, and so it would be ridiculous to say that MR predicates don't pick out properties."

I have telegraphed my response to this objection so often in previous chapters that I suspect readers will know what I am about to say. My response is to note that *there is a difference between "There are no MR properties" and "MR predicates don't pick out properties*." Indeed, I don't deny that MR predicates describe the world correctly nor do I deny that MR predicates pick out properties. In fact, on my view, MR predicates pick out lots of properties: different properties in different objects.

In the first chapter we discussed semantic disjunctivism, which is just the idea that a single predicate can pick out different properties in different objects. Semantic

disjunctivism should not be confused with semantic conjunctivism, which holds that a single predicate can pick out a cluster of properties in the objects it describes. Of course, semantic disjunctivism and semantic conjunctivism can be combined: a single predicate can pick out different clusters of properties in different objects. That complicates matters a bit, but affects none of what I am going to argue for below.)

Let's consider an example of a putative MR property, say, the property of being an axe. A lot of very different objects are axes. Some axes are made of wood with steel heads. Others are made of plastic with carbon-graphite heads. Some axes are small. Some are so large they can barely be lifted. Some, but not all axes are double-bladed. Indeed, there are a whole host of axes made of any number of materials in a plethora of different shapes. As swarms of atoms, the objects we call "axes" don't really have anything in common. Or, even if they do have something in common, it is not reasonable to think there is some relevant microphysical property or cluster of microphysical properties which we can say is identical to the property of being an axe. Any individual microphysical feature that all axes have in common, like having a mass above 0 grams, would be a feature that things other than axes would share, and so that property couldn't be picked out by the predicate "is an axe." Again, there is an argument or at least an intuition here; there is some *prima facie* case for the MR property of being an axe.

But that *prima facie* case doesn't hold up. I agree that we do have an *MR predicate* "is an axe." That is to say, there is a group of things we call "axes," i.e. "is an axe" is a way of classifying or "taxonomizing" objects, and there is no singular microphysical predicate that describes all and only the objects that are in the kind "axe." But it does not follow that there is a unique, singular property of being an axe, which all

and only axes have in common. The class of axes need not be ontologically homogeneous in any relevant way at all. Instead, the predicate “is an axe” may apply to a set of ontologically heterogenous objects. Indeed, it’s quite plausible, at least initially, that the class of things described by the predicate “axe” brings together a set of objects that don’t share some one common feature in virtue of which they belong in the class. Rather, perhaps the predicate “is an axe” describes objects in virtue of their possession of some one of a number of different properties from a disjunctive set.

Again, at least at first glance, there appears to be no reason we can’t say the same of every other example of a supposedly MR property, including the supposed best and strongest examples of MR: functional properties like being able to chop wood or being able to compute 2+2. What this means is that we can give an account for the fact that there are MR predicates that describe the world without positing MR properties. We can do so as long as the semantic disjunctivism I explained and argued for is plausible. (In the next chapter, we will consider the reasons given by nonreductive materialists for why the “disjunctive strategy” I am suggesting here will not work.)

Chapter 5: Higher-level Laws and Projectability

I. Overview

Unfortunately, the mere fact that semantic disjunctivism yields a tenable alternative to MR properties, as I explained in the fourth chapter, is not sufficient to dismantle the levels doctrine or the argument from MR properties. The argument from MR properties is buttressed by three separate arguments involving laws and projectable predicates, two of which I have alluded to briefly in previous chapters: the projectability argument and the argument from *ceteris paribus* (hereafter CP) laws. The third of these arguments, which I call "the argument from pragmatically valuable laws," is so intimately related to the argument from CP laws that it can be seen as a part of that argument or as independent. All three arguments are related in that they are all attempts to show that higher-level MR predicates must pick out properties that are not picked out by disjunctions of lower-level predicates. And all three arguments try to find evidence for the existence of MR properties in the nature of laws and inductive confirmation. If any of these arguments -or all of them in concert- is sound, then we have reason to accept MR properties and to reject semantic disjunctivism. And, as we saw in the previous chapter, MR properties are almost certainly higher-level properties.

In this chapter, I attempt to reconstruct and explain each of these arguments. And for each of these three arguments, I attempt to offer an alternative, uni-level explanation of the various facts about laws and confirmation that are supposed to point towards higher-level properties.

II. Law-Statements, Causal Relations, and Causal Powers

Before I begin explaining and critiquing the three arguments for levels, I would like to make a few preparatory remarks about laws, causal connections, and causal powers. First, I think it is critically important to distinguish between sentences that state laws and the laws themselves, i.e. that which is described by the law-statement. For example, we have the sentence "Bad money drives out good money." This is a sentence which states or describes Gresham's law. The law itself determines what happens in economies. Gresham's law determines that money that has value less than it is face value, will drive from circulation money that has value equal to or greater than its face value. By contrast, *the sentence* describing the law, i.e. the "law-statement," does not govern the behavior of anything. After all, a string of words on a page cannot impact an economy.

Laws govern the behavior of objects. Law-statements just describe or *track* those laws. But this raises another question: What are laws? What is their ontological status? There are at least two competing answers here. The first answer we could give is that laws are not entirely unlike Platonic universals, i.e. laws have a kind of existence independent of the things whose behavior they govern. Perhaps a less loaded and tendentious way of stating this point is to say that laws are causal relations between objects, and the ontological status of these laws can be explained in a number of different ways.

A second and entirely different answer is that there are no laws existing apart from the objects they possess. In other words, there are no causal relations, per se; rather individual objects possess *causal powers* in the same way they possess categorical, qualitative properties like redness or being cubical. On the causal powers view of laws, then, law-statements describe or pick out the causal powers of objects. That is to say, the

causal powers which inhere in objects are *those which are responsible* for the ebb and flow of causality. On this view, the causal powers determine what happens, so we don't need causal relations.[71]

Just to be perfectly clear, all of the arguments I will advance and all of the arguments that may be offered for levels are compatible with both views of the ontological status of laws. For example, we can think of Fodor (1975) and Putnam (1974) as providing arguments for higher-level causal powers or higher-level "autonomous" causal powers. Take your pick. Conversely, the claim I want to make can be stated by saying we can explain everything we know about laws and causation while positing only *base-level causal powers* of base-level entities or base-level causal powers, e.g. either the causal power of atoms to impart momentum to other atoms or base-level causal relations between atoms, like the conserving of momentum in interactions. So anywhere in the following chapter where I write "causal power" or "higher-level causal power," I could just as easily have written "higher-level causal relation," and vice versa.

Distinguishing questions about law-statements from questions about the ontological status of what those statements describe allows us to distinguish what we *know* to be true about laws, i.e. what we can use as evidence, and what needs to be determined on the basis of that evidence. That is to say, these foregoing considerations allow us to be clear what the data are when we are considering whether those data point towards the existence of ontologically existent higher-level laws.

[71] Furthermore, it is also quite possible that the existence of causal powers is fully compatible with the existence of *causal relations between events*, e.g. between the event of striking a match and the event of the match igniting. On this view the causal relation between two events reduces to some operation of the causal powers of the object(s) that are undergoing the event, in this case the match.

So what do we know? Well, for one thing we know that there is *a hierarchy of law-statements.* We know that there are sentences expressing true generalizations in the various special sciences. For example, we can read psychological law-statements in psychology texts, chemical law-statements in chemistry texts. Moreover, I think we can safely assume, putting Humean skepticism aside, that we know that some true generalizations are about either causal relations or causal powers, i.e. some sentences that express true universal generalizations track or describe causal laws of some sort. For example, the psychological law-statement "If someone has borderline personality disorder (BPD) they will experience splitting (which is oscillating back and forth between feeling extreme love and hate for those close to them.)" But we can also describe what is going on within BPD patients with neurological terms. For example, we can say this patient is bipolar because he has too much dopamine in the amygdale and is experiencing splitting because a particular region of his temporal lobe is activated. Moreover, we have a lower-level law that says "Too much dopamine in the amygdale causes activity in a particular region of the temporal lobe."

There is one more thing we know about the hierarchy of law-statements. We have learned empirically that there is a kind of determination between the higher and lower-level law-statements. For example, the truth of the lower-level law-statement "Low dopamine in the amygdale causes activity in a particular region of the temporal lobe" determines the truth, for this particular patient anyway, of "His BPD caused his splitting." Now, we have to be careful not to characterize this determination relation between law-statements too strongly as an explanatory relation or as a one-to-one bridge law. Doing so would imply that *we already know* that we can explain higher-level laws statements in

terms of lower-level law statements. And, as we will see, this would beg the question against the arguments from higher-level laws we will be discussing.

Indeed, the nonreductivist's position is that even though higher-level laws are, in some sense, determined by, or as Putnam (1974) says "deducible from," lower-level laws, higher-level laws still can't be fully *explained* by lower-level laws. In other words, it seems like something would be lost in any attempt to explain higher-level laws in terms of lower-level laws. Thus, nonreductive materialists don't and can't deny that the fact that this person's BPD caused him to split *because* something in his brain realizing his BPD caused something else in his brain which realized his splitting. If they did deny that there was a lower-level explanation for the behavior of a particular object, they would lose their last tie to materialism. Fodor (1975) expresses and maintains this tie to materialism by accepting, as he puts it, "the generality of physics in at least the sense that any event which falls within the discourse of a special science also falls within the universe of discourse of physics." (p. 54)

The point about the hierarchy of law-statements is that it is absolutely undeniable that the predicates deployed in many law-statements in the special sciences are often MR *predicates*. (In my terminology, MR predicates are disjunctive predicates or, to use Bechtel's terminology, "coarse grained" predicates.) Accordingly, it is impossible to line up higher-level law-statements in 1-1 correspondence with lower-level law-statements. At best, we can say that a higher-level law-statement is correlated with a messy disjunction of lower-level law-statements or a single law-statement that deploys disjunctions of lower-level predicates. For example, if "has BPD" and "is splitting" are multiply realizable *predicates* –though not necessarily MR properties- then a lower-level

law correlated to the law-statement "BPD causes splitting" would have to be something like "If a patient has low dopamine in the amygdale, or neurological condition X, or neurological condition Y, then that patient will have activity in the temporal lobe or neurological condition Z."

Interestingly, we sometimes tend to think of this picture of a hierarchy of law-statements as something that is necessarily true and known *a priori*. The picture is so familiar to us that it is tempting to think we could not have pictured the world differently. But this temptation should be resisted. It is a contingent fact that scientific law-statements come to us in a hierarchy and it is something we learn empirically. Clearly, other possible worlds could have different sciences. Indeed, our division of sciences into subspecialties is and has been subject to change throughout history. We have learned empirically that the regularities described by higher-level statements are determined by what happens at the lower-level. And this isn't necessary either. We could have discovered a set of law-statements with very different relations between them.

I want to encourage philosophers to look at these facts about law-statements as a kind of datum, an empirical datum, which stands in need of interpretation. We can and should ask: What does the hierarchy of law statements tell us about nature and about metaphysics? And the question before us is just this: What is the best explanation for our having this hierarchy of law-statements? The upshot of looking at the stratification of law-statements as a kind of empirical datum to be explained is that it turns the question of the existence of higher-level laws or higher-level properties more generally into a more *a posteriori* question than it is sometimes considered.

Of course, nonreductive materialists think there is something in this picture of a hierarchy of law-statements that points strongly towards the existence of higher-level causal powers, or "autonomous" higher-level causal relations. And they offer a number of arguments to that effect. But it is important to note that the burden of proof is on them to show that this hierarchy of law-statements is evidence for ontologically real, higher and lower-level laws.

Before we get to the specific evidence cited in favor of the levels doctrine as it concerns laws and law-statements, I will lay out a fairly simple, hopefully plausible, uni-level explanation for the mere fact that *law-statements* are stratified into the hierarchy of law-statements we have been discussing. To that end, I want to suggest that a law-statement that relates two MR predicates may not tell us about one, common causal relation between every one of the objects in the two microphysically heterogeneous kinds described by the predicates. In other words, I suggest that it is possible that a single law statement with MR predicates *tracks* or describes multiple causal relations existing between the different objects in the kinds. For example, on the levels doctrine view of laws, there is a single causal relation linking every object in the kind "depressed" and every object in the kind "helped by antidepressants," *and* there is a motley of neurological level causal relations linking the diverse neurological realizers of depression with the many different neuro-chemical realizers antidepressants. There is one causal relation which exists over and above the many lower-level causal relations. By contrast, on the uni-level view, the higher-level law statement "depressives are helped by antidepressants" picks out multiple causal powers or tracks multiple causal relations. The uni-level view is more parsimonious because it doesn't posit one higher-level law to rule

them all. On the present view, a good higher-level law-statement is just a "multi-track law-statement."

Of course, the mere fact that there is an alternative view is not the end of the debate over higher-level laws. As I suggested earlier, Fodor, Putnam and Antony have three more rounds in their clip. So it remains for us to show that the uni-level view is just as capable of explaining everything we know about law-statements and laws as the levels doctrine.

III. The *Ceteris Paribus* Laws Argument

Let us consider the first salvo fired in defense of higher-level laws and MR properties. In "Philosophy and our Mental Life," Putnam (1974) argues that higher-level laws cannot be reduced to lower-level laws as follows:

> Very often we are told that if something is made of matter, its behavior must have a physical explanation. And the argument is that if it is made of matter (and we make a lot of assumptions), then there should be a deduction of its behavior from it is material structure. *What makes you call this deduction an explanation?* [Emphasis in the original] (p.126)

A little later Putnam explains how he intends this distinction between deduction and explanation:

> People have argued that I am wrong to say that the microstructural deduction is not an explanation. I think that in terms of the purposes *for which we use the notion of explanation*, it is not an explanation. If you want to, let us say that deduction *is* an explanation, it is just a terrible explanation… Explanation is superior [to deduction] not just subjectively, but *methodologically*, in terms of facilitating the aims of scientific inquiry, if it brings out relevant laws. An explanation is superior if it is more general. (p. 126)

But why does Putnam (1974) think that the "deduction" of higher-level properties from lower-level properties fails to be general enough to be explanatory? Well, obviously explanations involve laws, e.g. the explanation we give of why a sample of gas in a container exerts such and such pressure on the walls of its container points at Boyle's

law. It is ultimately the law that does the explanatory work. So, Putnam is saying that a mere deduction of a higher-level property from lower-level properties is not explanatory because, presumably, it does not point to a law in the same way. That is to say, mere deductions of higher-level properties from lower-level properties don't point to laws. But why should we believe this? Why can't a mere deduction of a higher-level property from a lower property point to a law?

The answer lies in the fact that higher-level properties seem to be involved in higher-level laws, while the disjunctive predicates which pick out their microphysically heterogeneous level realizers do not. To use Putnam's primary example, there are higher-level "generalities," like the macroscopic generality "2 inch square pegs will fit through 2 inch square holes," whereas at the lower-level there doesn't seem to be any general law relating all the different realizers of square pegs and boards.

So, Putnam's argument here is really that higher-level law statements are more *general* than the corresponding disjunctive lower-level laws they are "derived from," and the *generality* of higher-level laws implies that higher-level laws must be distinct from lower-level laws. Or, to use Putnam's words, higher-level laws must be "autonomous." After making this attempt at proving the autonomy of higher-level laws, Putnam then concludes -and here one may agree- that the existence of autonomous higher-level laws implies the existence of "higher-level structural features" related by those laws. But I agree only in the sense that the implication is vacuous.

To save the uni-level view, then, the inference from the generality of higher-level laws to autonomous higher-level laws must be shown to be invalid. One might try to do this by denying the premise of the argument, but that doesn't seem very plausible,

because there does seem to be something right about the idea that higher-level laws (or law-statements as we will see below) are more general than their lower-level counterparts.

Before we turn to criticisms of Putnam's view, we need to explore the concept of generality, a concept which Putnam does not fully explain despite its centrality to his position. In fact, there are two interpretations of what Putnam might mean by "general" that fit with what Putnam actually says. These two interpretations yield two different but related arguments for levels.

The first way of interpreting what Putnam might mean when he says that a law is general, or more general than a rival, is that it is a *ceteris paribus* (hereafter "CP") law, i.e. a law describing the behavior of objects *all other things being equal*[72]. The concept of a CP law is a familiar one, and it is hard to deny that there are such laws since both common sense and scientific practice rely on them. For example, the law we looked at above, that depressed people are helped by antidepressant medication, has exceptions, but to say it isn't a law would leave psychiatry in a difficult state. Some people who are depressed are not benefitted by antidepressants. Others are helped by some antidepressants but not others. But despite these exceptions, despite the fact that the law

[72] In his earlier work, Fodor (1974) also gives a version of the CP-laws argument along similar lines: "To put this discussion less technically: we could, if we liked, require the taxonomies of the special sciences to correspond to the taxonomy of physics by insisting upon distinctions between the natural kinds postulated by the former wherever they turn out to correspond to distinct natural kinds in the latter. This would make the laws of the special sciences exceptionless if the laws of basic science are. But it would also loose us precisely the generalizations which we want the special sciences to express. (If economics were to posit as many kinds of monetary systems as there are kinds of physical realizations of monetary systems, then the generalizations of economics would be exceptionless. But, presumably, only vacuously so, since there would be no generalizations left to state. Gresham's law, for example, would have to be formulated as a vast, open disjunction about what happens in monetary system or monetary system under conditions which would themselves defy uniform characterization. We would not be able to say what happens in monetary systems tout court since, by hypothesis, 'is a monetary system' corresponds to a natural kind predicate of physics.)" (p.112)

is not entirely universal, it certainly seems that we still have a law -or at least something suspiciously law-like- here. For example, Boyle's law ($pV = k$) is CP. If the particles of a gas are at an extremely high temperature, or if the gas molecules are particularly heavy organic gas molecules, Boyle's law will not describe the behavior that gas. Nonetheless, Boyle's law is certainly law-like.

Before we go further into the notion of general laws as CP, let us recall the distinction between law-statements, i.e. sentences that describe regularities, and the regularities themselves that the law-statements describe, i.e. the existing causal powers or causal relations that determine the behavior of objects. The alternative explanation of the CP character of laws that I want to put forward rests on the idea that it is not so much a feature of the causal relations between objects, a real feature of the world, as it is *a feature of the law-statements* that pick out those causal relations. Somewhat more precisely, I want to suggest that it is possible that the CP character of a law-statement is a consequence of how multi-track law-statements refer (in a loose sense of "refer") to the variety of different causal relations between different objects tracked by the statement. The CP character of CP laws does not reflect anything about the supposed higher-order properties denoted by the predicates of CP laws.

Returning to our example, we have a sentence "People who are depressed will be benefitted by taking antidepressants." Notice that this law-statement only applies in some cases and not others. There are some objects described by the predicate "depressed person" which are not causally impacted by any of the objects described by the predicate "is an antidepressant." And some of the objects described by the predicate "is an antidepressant" have no effect on some objects described by the predicate "depressed

person" For example, imagine one person is depressed but drinks a bottle of whiskey a day and as a result no antidepressant will help him. Or imagine another person who has gene-X, whic causes a rare neurological irregularity, which in turn makes some kinds of antidepressants -say the serotonin reuptake inhibitors- ineffective in his case.

What this shows is that, if we want to make the law-statement more precise and exceptionless, we would need to add a series of clauses to the law-statement describing the exceptions. For example, we might add a series of exceptions to our psychological statement, rewriting it like this, "A person with depression will be benefitted by antidepressants, except if they drink alcohol regularly or have gene-X." But even this law with these explicit exception-clauses is not guaranteed to be exceptionless. Imagine someone takes an antidepressant medication that is near its expiration date and the chemical has slightly broken down. (This thing called an "antidepressant" won't cause this thing called a "person" to be less depressed.) Or suppose a depressive is eating poorly, their diet leads to malnutrition, the malnutrition leads to an extreme shortage of the neurotransmitter chemicals associated with feeling well, and the person taking the antidepressant doesn't feel better.

So, in order to clear the exceptions from this psychological law-statement we would need to list all of the possible exceptions. But to do that we would end up with, at best, a massively exception-laden law-statement describing all the possible brain states that are the so called realizers of depression and all the possible chemical interactions that various anti-depressant chemicals could have with those brains states. And that bulky law-statement would be mostly useless in teaching and practicing psychiatry. It would require too much effort and time to teach to psychiatric students if you had to continually

restate such a long, exception laden law-statement. So, for the sake of simplicity and brevity, we need a shorter law-statement that doesn't explicitly list all the exceptions. Of course, we will have to make clear to the people to whom these law-statements are communicated that it has a variety of unstated exceptions. To do that we can add a sort of implicit clause to the law-statement: "There are a lot of exceptions to this law, not all of them can be listed" or "This law doesn't apply universally, only *all other things being* equal."

Let's look at a different example. Suppose we have two taxonomical kinds A and B. Let's say that everything in kind A is described by the predicate "is arsenic" and everything in kind B is described by the predicate "is a human." We have a law statement, "Arsenic kills humans." Fortunately, for those with enemies, the law-statement has exceptions, i.e. it is a CP law-statement. For example, a small amount of arsenic might kill a small person but not a large person, or an elderly person but not a young adult. So, there is a micro-chemical causal relation between some objects described by the predicate "is arsenic" and some objects described by the predicate "is human," but that particular causal relation doesn't exist between all of the objects described by the predicates. And this may be why some samples of arsenic won't kill some people. Moreover, on this view, there will be different micro-chemical causal relations between different objects called persons and different objects called "arsenic." The law-statement "Arsenic kills humans" will track all of these causal relations. So, one way of explaining the CP character of "Arsenic is fatal" is by noting that a particular person who has taken a sample of arsenic may not die because none of the several causal relations that are tracked by "Arsenic kills humans" hold between this entity and this sample.

This interpretation of CP-ness as a consequence of the multi-track character of law-statements predicts that the more heterogeneous the kinds picked out by the predicates deployed in a law-statement are, the more exceptions the law-statement is likely to have. For example, the kind described by the predicate "is a poison" is presumably more heterogeneous than the kind described by "is arsenic," and the kind described by "is an animal" is more heterogeneous than the kind described by "is ahuman being." (I do not claim that measuring and comparing heterogeneity with precision is an easy matter. But I think these examples are fairly plausible and intuitive.) And obviously, the law-statement "Poison kills animals" is even more CP than "Arsenic kills humans." The proposed explanation for this is that the former law-statement tracks a larger and more diverse set of causal relations, and thus there is a greater likelihood that a particular animal, on any given occasion, might be immune to whatever particular poison might be on tap. For example, if we give person X poisonous spider venom, it will kill him, but spiders will be immune to it. By contrast, some persons might be immune to small amounts of alcohol, which might kill a small spider. That is, there is a specific causal relation between spider venom and persons that will make the generalized law-statement "Poisons kill animals" a correct description in some cases and not others.

IV. The Argument from the Pragmatic Value of Higher-level Laws

This leads us headlong into the second interpretation of what Putnam might mean by a "general" law and also into the second of the three arguments for higher-level laws. Suppose Putnam objects to the idea that higher-level law-statements are multi-track by saying that this view can't account for *why* higher-level, CP laws are more pragmatically

useful than their disjunctive, lower-level, exceptionless counterparts. And Putnam could object, along similar lines, that the view proposed doesn't explain why we would find it useful to bundle sets of causal relations together and describe the relations in that set with one unified law-statement "X's cause Y's"

Let us look at this objection a little more carefully. When Putnam says that higher-level laws are more general than their lower-level counterparts, he might mean something like the following. We look at the world and we notice regularities between phenomena of two different taxonomical kinds. Returning to a psychiatric example, psychiatrists and psychologists notice that if you have observed someone correctly described as being in the category described by "clinically depressed," then you should expect to observe that this person will also belong to the kind described by "is helped by antidepressants." I think Putnam wants to suggest that if we looked at the brains of people with depression "from the perspective of microphysics," we wouldn't notice these same regularities. Moreover, he suggests that "from the perspective of microphysics," it wouldn't be useful to think of some of the various different swarms of atoms and chemicals constituting our brains as being similar. That is, grouping swarms of atoms together into the kind described by the term "depressed" would be totally arbitrary and accidental fro the microphysical perspective. I think the best way to think of Putnam's argument is to think of it as a challenge: Why should we have a single, true, and pragmatically valuable sentence that tracks all these different base-level causal relations, when all these different microphysical causal relations are very different?

Putnam's inference from perceiving similarities from a higher-level perspective to their being such similarities is suspect. Nonetheless, unless the uni-level view can provide

a plausible explanation for the generality of laws in this sense, i.e. in the sense that higher-level laws draw out salient, pragmatically valuable, "general" similarities between objects by ignoring their particular differences, the uni-level account of laws would lack the explanatory power of the levels doctrine. In that sense Putnam has hit upon an important and powerful argument.

However, I think the needed explanation is available without positing higher level properties. In rough outline, the alternative explanation works like this. First, as C.B. Martin (2008) observed, very different causal powers often have the power to bring about the same effect in different situations. Indeed, it seems to me, Martin has to be right about how we individuate causal powers. His observation about causal powers can then be used to explain why it would be useful to form metaphysically heterogeneous taxonomical categories. The pragmatic value of higher-level law, multi-track law-statements follows naturally from the pragmatic value of the heterogeneous kinds picked out by the MR predicates deployed in higher-level law-statement.

Let's start with Martin's idea about causal powers. Philosophers often talk about causal powers in a loose way. For example, it is commonly said that some objects possess the power to cause us pain or the causal power to make us have an experience of seeing yellow. There is nothing wrong with this loose talk *per se*, but, as Martin points out, we can't properly individuate causal powers with predicates that just describe what a causal power is for the power to do.

According to Martin (2008), if we want to properly individuate causal powers, we need to identify, not just what the power is a power to bring about but also how the power is, to use Martin's term, "selective." Describing the selectiveness of a causal power

means describing the conditions under which the causal power successfully results in manifesting what it is for. For example, a sharp steel knife has the causal power to cut a tomato *when it is pressed with enough force*. It is unclear how refined our dispositional predicates would have to be to properly individuate causal powers. In fact, if Martin is right about the failure of the conditional analysis of dispositions, i.e. if a disposition's selectiveness outstrips what can be described conditionally, then our dispositional predicates might never be precise enough to ensure that every object described with that predicate shares a common causal power. Nonetheless, even if we can never come up with perfectly individuated dispositional terms, we can still say that the more "individuative" a dispositional predicate is, the closer it is to an asymptotic mean.

On Martin's view, there is no such thing as a power to cut, or a power to X, *simpliciter.*[73] Martin's view can be argued for via a simple *reductio.* Suppose Martin is wrong and there is a causal power that is identified and properly individuated from other causal powers by the simple phrase "the power to harm." Well, if so, what objects would possess this power? Certainly knives, guns, fists, and a whole host of other objects would all possess this power since they can, in some situation or another, bring about harm. But really, every object *could*, in some circumstance or another, result in harm. For example, a teddy bear thrown at sufficient velocity could result in harm. Air can cause harm when heated or accelerated. Empty space is deadly if you need to breathe. Turning to other examples, every object *could* appear red in the right circumstances, e.g. in the right light,

[73] I suspect Martin's (2008) idea about dispositions is in direct conflict with the view, associated primarily with Shoemaker (2007) and Clapp (2001), that there is an "overlap" of causal powers between, for example, all objects that are colored. Shoemaker says you can identify a series of objects that are yellow, red, blue, etc. The overlap of all these causal profiles is the dispositional property that distinguishes being colored. On Martin's account of causal powers, different red, yellow, and green objects all share a variety of very different dispositions to reflect light in different ways in different contexts. There is no "overlap."

to the right observer, perhaps most especially to creatures with very different sensory organs. Pigs could fly, if gravity wasn't pulling down on them so forcefully. So too could everything else fly in some circumstances. But it is absurd to say that all objects have the disposition to fly, to be harmful, and to appear red. Indeed, if we individuate powers only by what they result in, then every object has every causal power. Every object has the power to do X, where X is anything the object could do in some imaginable circumstances. On this view the only powers that all objects wouldn't have would be the power to do the logically impossible. Every object would therefore have every real causal power and causal powers wouldn't serve to distinguish different objects from each other.

It might be objected that this is a slight overstatement. Perhaps it is merely the case that if dispositions or causal powers are identified by what they are powers to do, then most every object has most every power or many powers, since most every object can do almost anything in some circumstance or another. Even so, this weaker statement, i.e. that every object has most every power, is absurd. It would imply among other things that knives possess both the disposition to cut and the disposition not to cut. (I would argue that a particular knife has one disposition to cut in certain circumstances and not others.) And the absurdity of the weaker claim gives us reason to reject the idea that dispositions are distinguished only by what they are dispositions to do. To prove the stronger claim that distinguishing dispositions only by what they are dispositions to do implies that all objects possess all dispositions would require that it is true that there are some *possible* conceivable circumstances where a particular object does anything any other object can do. That is to say, to prove the stronger claim, we would have to agree that there is some situation or other where, for example, a bird's song could create a black

hole or bring a person back to life. These circumstances needn't be likely to occur, but they would have to be possible. And certainly some otherwise absurdly unlikely things are at least possible, which means that distinguishing dispositions by what they are dispositions to do would imply that everything can do, at least most everything.

Notice that even though there is no such thing as a power to do X *simpliciter*, we still find it useful to create a category of things picked out by the simple, unindividuated, predicate "is harmful" or "has the power to cause harm" even when all the things that fall into the kind picked out by that predicate have very different causal powers to bring about harm in very different contexts. For example, we teach children that hot stoves and sharp knives fall into the kind described by "is harmful" but teddy-bears and broccoli do not. This is a pragmatically valuable rule for parents to follow, because in most ordinary circumstances hot stoves will cause harm. Of course, teddy bears and broccoli could cause harm in the right situations, so it might seem arbitrary to exclude them from the category. In some circumstances, we might also find it useful to form narrower kinds, picked out by more refined predicates like "is something that causes burns when touched" "is something that causes illness if you eat enough of it." But of course, these minimally refined terms will pick out heterogeneous kinds, too, just like "is harmful." For example, a sharp knife does not cause harm just when it is touched, it has to be touched with a certain amount of force. Does it fall into our minimally refined category or not? It doesn't take a great many examples to show that if we wanted to form perfectly individuated kind terms, they would be so refined that they would be unwieldy and nearly impossible to teach, i.e. we would have one kind term "is a thing that causes 2nd degrees burns in children but not adults when touched for more than 10 seconds." So, to make life

easier on ourselves, we notice that in the ordinary world there is a large set of objects that tend to result in a particular outcome in common situations, e.g. a wide range of things that cause harm in a number of very different but common situations. This, it may plausibly be conjectured, accounts for the use of the predicate "is harmful."

Suppose everything just stated about the pragmatic value of the descriptive predicate "is harmful" applies equally well to all the kinds described by sensation language, too, e.g. "is something that looks red" or "Is something that tastes sweet." Surely the same analysis might then also apply to functional predicates and the nonreductivists favorite examples of functional properties. If Martin is right, it is quite plausible that we create a taxonomical category "is red" and place in it a variety of different objects that cause red experiences in us in a variety of different ordinary contexts. But if so, then the objects that cause sensations of redness in us need not share any common causal power to cause us to have that red experience. When we try to create properly individuated kind terms, e.g. "looks red to everyone who has rods and cones of such and such sort when lit exactly so" we realize that many of the objects in the taxonomical category of "red looking" actually have very different causal powers. The same applies to functional predicates as well. Not all axes share a common, simple power to chop wood. Nor do all computers and brains possess the simple power to add 2+2. Different computing machines and brains have the power to yield the answer 2+2 under different circumstances, when different inputs are applied. Quine[74] (1969) puts this point eloquently when he writes,

> "… color is cosmically secondary. Even slight differences in sensory mechanisms from species to species, Smart remarks, can make overwhelming differences in the grouping of things by color. Color is king in our innate quality space, but undistinguished in cosmic circles." (p. 238)

[74] Quine credits Smart (1963) for the same point.

So far, I have only suggested why we might find it pragmatically useful to form heterogeneous taxonomical kinds. I haven't explained why we might find it useful to form law-statements that track apparently arbitrary collections of base-level causal relations. Well, in the simplest terms, the pragmatic value of higher-level law-statements *simply follows from* the pragmatic value of the heterogeneous taxonomical kinds these law-statements relate.

To see how the pragmatic value of higher-level, multi-track law-statements is a consequence of the pragmatic value of the heterogeneous taxonomical kinds the law-statements relate, let's look at a fanciful example. Suppose there are primitive people living on an alien planet. They often encounter creatures -some with fangs, others that can fly, some small, some large, some mechanical robots, other biological organisms- that kill and eat people. The primitive people don't have the time or the ability to zoologically analyze the creatures as being the same or different in terms of their genes or their phenotypes. They simply call all such creatures that tend to kill people under normal circumstances "monsters." Furthermore, let's imagine that the primitive people also notice a number of plants in their environment have a distinctive, strong, pungent smell. Suppose they call all such plants, "pungents"

It is almost inevitable that the predicates, "is pungent" and "is a monster," describe heterogeneous kinds of objects. That is to say, we may suppose that all creatures in the kind described by "is a monster" all tend to kill humans under normal circumstances. But, as we have seen, the mere fact that monsters tend to cause people to die under a variety of different ordinary circumstances is not sufficient to show that all things in the taxonomical kind "monsters" share a common, unique causal power. We

may also imagine that " pungent" things, by very different chemicals undergoing very different reactions, act on the olfactory receptors in the noses of the primitive people, causing their neurons to fire, eventually causing a sensation of a pungent smell. Actually, let's imagine that some of our imaginary pungent plants do not create a pungent smell by releasing a chemical, but by emitting a weak electrical field, which affects the olfactory receptors of our primitive people directly.

Returning to multi-track law statements, let's suppose that over time the primitive people begin to observe that carrying something pungent tends to result in monsters running away. As a result, they form the causal law-statement "Pungents repel monsters." Clearly this is a pragmatically valuable law-statement. It helps our primitive people to avoid being eaten.

The law-statement "Pungents repel monsters" is multi track because different pungents have very different causal powers to repel monsters in very different situations. Perhaps the pungent plants that emit a chemical repel monsters because the chemical burns their skin. Perhaps the pungent plants that emit an electrical signal repel monsters because the electric signal affects their brains and causes them to feel intense fear. Maybe other monsters –those with noses- just don't like the smell.

There is a sense in which it is arbitrary to lump all these different causal powers of pungents together into a category. These causal powers aren't really the same and they certainly wouldn't look the same "from the perspective of microphysics." But that doesn't really matter for practical purposes to the primitive people. They've already formed one taxonomical category based on sense experience, i.e. the kind described by "is pungent." They then begin to notice that the presence of something from the category

pungent object tends to make the presence of a monster less likely. Granted, it may be an accident that very different pungent plants tend to result in similar outcomes in ordinary circumstances. But once the heterogeneous kind terms are formed, and the primitive people begin to categorize objects according to those terms, they are going to notice regularities between the occurrences of objects in the kinds, even if those regularities obtain because of very different causal relations between different objects in the heterogeneous kinds. And even if the regularities have exceptions, and even if it is impossible to state the exceptions systematically.

Ultimately, one can interpret the use of MR, coarse grained predicates to describe heterogeneous sets of objects as showing that these coarse grained predicates are useful. They are useful because the objects they describe tend to behave similarly under a variety of different, commonly occurring conditions. The objects in the heterogeneous kinds will have a variety of different causal relations with each other. But since we are looking at the world through the prism of the heterogeneous categories we already have, we will see these causal relations as general, probabilistic, CP regularities between most objects of one kind A and most objects of another kind B. Thus, we have a model for CP laws that does not require any laws beyond the microphysical ones. Other things being equal, considerations of economy favor this account.

V. The Projectability Argument

We come now to the projectability argument, which is the last, best hope of nonreductive materialism. In the simplest terms, the projectability argument finds evidence for the existence of higher-level properties in the fact, or the seeming fact, that

there is a difference in the "projectability" of higher-level predicates and the projectability of the disjunctive predicates that pick out their so-called microphysical realizers. The phenomenon of projectability is familiar and so the following brief definition will do. A predicate "is an X" is projectable if and only if the observations described in sentences like "This X is Y" confirm the more general law-statement "All X's are Y's"

The simplest version of the projectability argument is just this:

P1: Some higher-level MR predicates are projectable even when the lower-level disjunctive predicates that pick out the corresponding lower-level properties are not projectable.

P2: As Louise Antony (2003) says, the fact that a predicate is projectable is evidence that "reality is playing along"[75] and that the predicate picks out a property.

C: There are higher-level properties which are distinct from lower-level properties. (Note that Antony's (2003) version of the argument involves restating P1 to allow the fact that disjunctive predicates do pick out properties: they pick out disjunctive properties which exist over and above any individual disjunct property. We will examine this in the next section.)

The argument is valid as stated. However, I think both of the premises, as stated, are false, although statements which are only subtly different from them are true. A little more precisely, I want to argue that the evidence we have about which predicates are projectable and which aren't projectable can be explained without conceding P1 and P2.

[75] As Antony (2003) admits, Goodman himself, who introduced the notion of projectibility, would not have agreed with such a realist position. However, it has become common in philosophy to think that if a predicate is projectable, then it must pick out a property or a "natural kind." In light of Goodman's extreme nominalism, it is not even clear that he would understand the uni-level vs. multi-level problem.

Before I begin to criticize this argument, I should explain and motivate P2, since P2 links the epistemic issue of projectability to metaphysical questions about properties. The idea is just this. The fact that a predicate is projectable shows that the predicate is successful in describing the world. For example, one difference between the predicate "is an emerald" and "is one of the things on my desk" is that the former is projectable and the latter isn't. How do we explain the difference? Well, there are true non-accidental law-statements, i.e. universal generalizations that support counterfactuals, which can be discovered by observing that individual objects are described by predicates like "is an emerald" and "is green" and then generalizing to a claim like "All emeralds are green." The same cannot be done with the predicate "is one of the things on my desk." "Everything which is on my desk is green" may be accidentally true, but it is not a counterfactual-bearing law; if you put something on my desk, and then observe it, you do not have reason to suspect it will be green.[76] What this suggests is that there is something worthwhile, i.e. something right and veridical, about grouping objects together into the category "is an emerald," which in turn suggests that "is an emerald" is at least a better candidate to pick out a property than "is something on my desk."(I will come back to P2 later.)

Before we begin discussing this argument, note that a common strategy in the debate over higher-level properties is to find some test for the reality of properties in general –e.g. that real properties are distinguished by their causal profile or that real

[76] This does suggest that the category of things described by the predicate "is an emerald" share salient and important *similarities* in how they behave and in how they are. But note that I say *similarities* and not the singular "similarity." I say this because it very well *may be* that projectable predicates pick out different properties on different occasions. I don't want P2 to rule out my view before we even consider the argument. That said, I don't want rule out the levels doctrine just by pointing out that P2 is compatible with my view, either. The meat of the argument is in P1. If the nonreductivist can establish that there is something different in how lower and higher-level predicates pick out properties, then that is evidence for the levels doctrine and a problem for the uni-level view.

properties are picked out by projectable predicates- and to then leverage that test into an argument either for or against higher-level properties. This is the battle that Kim has with Fodor and Antony. There are problems with this strategy on both sides. First off, since the different sides deploy different tests, i.e. Kim has a causal powers test and Antony and Fodor have a projectability test, they will tend to argue past each other. And second, it may be that neither of these tests is as clear and decisive as may be assumed. I will take a somewhat different tack.

The main objection I would bring against P1 is that disjunctive predicates may be projectable. This sort of objection is another way to highlight the difference between the present attack on nonreductivism and Kim's (1992). Kim argues that disjunctive predicates aren't projectable *and* that higher-level MR predicates aren't projectable either. Having done that, he suggests that the predicate "is jade" is identifiable with the predicate "is jadeite or is nephrite" and so "is jade" is no more projectable than the disjunctive predicate "is jadeite or nephrite." Accordingly, Kim must conclude and does conclude that there are no laws about jade as such. Rather, Kim argues, there are only "local laws" about the realizers of jade, i.e. jadeite and nephrite. Kim then argues by analogy –as Fodor usefully points out (1997)- that if pain and other higher-level predicates really are MR predicates, then they're just like "is jade," i.e. that there can be no laws about these putative higher-level, MR properties either. But that implies, according to Kim, that there are no MR properties with causal powers distinct from the causal powers of their lower-level realizers, which in turn implies –assuming that properties are identified by their causal powers- that there are no MR properties. Finally, Kim suggests that if a predicate is an MR predicate, it isn't involved in laws, per se;

rather there are different laws involving each of the different realizers. To reduce the laws deploying MR predicates, therefore, we need to break them apart and do a "local reduction."[77]

The position defended here is in line with the reductivist aims of Kim's argument, but Kim's causal exclusion argument and arguments like it can't defeat the levels doctrine. Kim tries to argue that higher-level MR predicates are no more projectable than lower-level, disjunctive predicates and that therefore premise P1 of the simple projectability argument above is false. However, as Antony (2003) argues and Fodor (1997) comes close to arguing, Kim's strong ban on projectable disjunctive predicates is not warranted, and therefore Kim's critics have the best of him.

In short, as a result of the fact that pain must be carved up into different subtypes, Kim is forced to conclude that there isn't some one causal power existing in all minds described as "being in pain" in virtue of which they are all in pain. Nonetheless, there is a strong intuition pressing us to conclude that many of the creatures belonging to the different subtypes of being in pain are likely experiencing the same qualitative feeling when they feel pain, despite the fact that their brains don't possess some common causal power. Kim is therefore forced to accept that qualia don't endow the brain with a set of causal powers. But this in turn makes qualia epiphenomenal. The argument presented here does not suffer this same problem. Saying that pain is projectable, as is the disjunction of its base-level realizers, does not lead to epiphenomenalism.

Let me give an outline of how I will proceed. After we have rejected Kim's strong ban on projectable predicates, I will examine Antony (2003) and Fodor's (1997) more subtle restatements of P1, i.e. their more subtle versions of the argument from

projectability, which don't state anything quite so strong as "no disjunctive predicates us projectable," but nonetheless point to a difference in the projectability of higher and lower-level predicates.

Kim's argument against the projectibility of disjunctive predicates is straightforward. He writes:

> "Does…"Jade is green" pass the projectability test? Here we have a problem. For we can imagine this: on re-examining the records of past observations, we find, to our dismay, that all of the positive instances of (L), that is, all the millions of observed instances of green jade, turn out to have been samples of jadeite and none of nephrite! If this should happen, we clearly would not, and should not continue to think of (L) as well-confirmed." (p.520)

This argument is clear enough. Suppose you keep observing samples of something you know to be jade. You keep discovering that all of them are green. Or, put a little differently, you keep making observations described by the sentence "This sample of jade is green." Now, suppose that by mere misfortune, every sample of jade you've been observing is jadeite and none were nephrite. If the particular observation described in each sentence "This sample of jadeite is green" confirms that "Jade is green," then it would seem that by observing samples of jadeite, you can learn about the color of samples of an entirely different substance, i.e. nephrite. And that is just as mysterious as thinking that observing that coffee is bitter can somehow confirm that butter is bitter, because "this coffee is bitter" also confirms "coffee or butter is bitter."

We are now sailing into the painfully complex problem of inductive confirmation. Obviously, I don't have a well worked out theory of confirmation to offer, which makes it more difficult to argue conclusively about what (if anything) confirms hypotheses using disjunctive predicates. (However, the converse is also true; the argument from projectability is suspect until we have a good handle on what counts as projectable.) But as this is where the discussion of the levels doctrine has gone, we must follow.

Even in the absence of a positive theory of confirmation, there is reason to reject Kim's general ban on all projectible disjunctive predicates. In fact, Kim appears to have conflated general Humean worries about inductive confirmation with more specific questions about the projectability of disjunctive predicates. Fodor (1997) comes close to arguing this by saying that Kim has mistaken worries about disjunctive predicates with worries about "sample bias." The general worry about inductive confirmation –what Goodman (1983) calls the old riddle of induction- is this: no matter how many times I observe particular apples falling towards the earth when dropped, I can never be certain that all apples will fall when dropped. *Maybe* there is something wrong with the apples I was observing or maybe gravity operates differently now than it will in the future. *Maybe* other apples will be different. Well, the same is true with jade too. *Maybe* all of the jade we have observed was all jadeite and none of it was nephrite. If so, we haven't confirmed that all "jadeite or nephrite" is green or that all nephrite is green.

But those italicized "maybes" are big maybes. The fact remains that we do know that many of our observations of jade have been observations of jadeite and many have been observations of nephrite. Now, we *could* be wrong about this; maybe some evil demon-geologist has confused us about how common nephrite is. If so, then what we thought was true on the basis of inductive confirmation will turn out to be false[78] But the mere fact that a sentence *may* turn out to be false doesn't imply that the sentence is not well confirmed. After all, confirmation is not deductive implication and so no matter how

[78] Ironically, despite Kim and Fodor's claims to the contrary, not all jade is green. Both jadeite and nephrite can be white, yellow, *or* brown, too. And there are some rarer cases of other colors of jade as well. That is, the law regarding the color of jade would be something like this: "If something is jade then it is *either* green, white, brown, or yellow. Actually, pure nephrite and pure jadeite are both cloudy white. The color in samples of jade is due to traces of other minerals. Thus, any law regarding the color of jade or even a law regarding jadeite would have to contain a disjunctive predicate.

much confirmation we have, there is always a chance that what we have confirmed will turn out to be false.

Kim could easily respond that there is a distinctive weakness about disjunctive predicates. He might say if we observed samples of jadeite and samples of nephrite and found both to be green, we really have two confirmed laws, not one. In other words, Kim can say that "This sample of jadeite is green" confirms "All jadeite is green," and "This sample of nephrite is green" confirms "All nephrite is green." That is, Kim can say it is not really that the disjunctive predicate is projectable; rather, each disjunct predicate "is jadeite" and "is nephrite" is projectable, and the statement "If something is jadeite or nephrite then it is green" is just a consequence of the two more basic laws "All nephrite is green" and "All jadeite is green."

Well, Kim may well be right that unless we have some reason to believe that our observations of jade are distributed so as to ensure that some were observations of jade were observations of jadeite and others were observations of nephrite, we shouldn't think we have confirmation for "If something is jadeite or nephrite, then it is green." But this does not mean no disjunctive predicates are projectable. It means something weaker: that disjunctive predicates are only projectable *under certain circumstances*. That is, only a ban on disjunctive projectable predicates more limited than P1 is warranted.

The limited ban could be stated as follows. Instances of a disjunctive predicate "is an A or a B" do not confirm a general law-statement "Everything that is A or B is X" *unless* we have some reason to believe that our observations of things that are "A or B, [but we don't know which]," are likely to be A's reasonably often and B's reasonably

often. Call this a disjunct-distribution law. I claim that disjunctive predicates are projectable if and only if they are backed up by a disjunct-distribution law.

The example of jade affords us a good picture of what a disjunct-distribution law would look like. We know from geology that both jadeite and nephrite are reasonably common, geologically speaking; so we know miners are going to dig up samples of both from the earth. Moreover, we know that miners, metallurgists, and jewelers are likely to treat jadeite and nephrite similarly since they tend to look the same and have similar hardness, weight, and cleavage. These two things we know about jade –which are confirmed inductively- insure that any given sample of jade has reasonably similar probabilities of being jadeite or nephrite. We can think of these things we know about jade as a disjunct-distribution law, because they ensure that observing many samples of "jadeite and nephrite" is likely to result in observing samples of each.

I am not prepared to say how disjunct-distribution laws figure formally in a larger theory of confirmation, but the following seems clear. It is already widely admitted that the bearing of evidence on a hypothesis depends on the background. So, in order to explain the difference between successful and unsuccessful inductive confirmation of hypotheses involving disjunctions we must take into account what is going on in the background, i.e. what we have already inductively confirmed elsewhere. For example, as long as we have background knowledge that jadeite and nephrite are roughly equally distributed throughout the world, then a series of observational sentences "This is jadeite or nephrite and it is green" would serve to confirm "All things that are jadeite or nephrite are green," even if in every single case –or in the vast majority of cases- the person who was observing the jade being green didn't know whether the jade was jadeite or nephrite.

Note that a disjunct-distribution law needn't be as strong as "the occurrence of A's and B's are equally probable." If we have made millions of observations of things which are "A or B or C, but we don't know which," then our background claim only needs to make it probable that a reasonable fraction of the things were A, a reasonable fraction were B, and a reasonable fraction were C. The disjunct-distribution law needs only to insure that we would observe instances of every disjunct A, B, and C eventually if we continually made observations of things being "A or B or C." Moreover, there is no such guarantee with arbitrary disjunctions like the one introduced by Armstrong (alluding to Lewis Carol) e.g. "is a raven or a writing desk."

Suppose you learn that Mr. X has been observing objects and found 10,000 that belong to the category "is a raven or a writing desk," And he has found all to be black. That is all we know. In this case, there is no guarantee that Mr. X hasn't observed 10,000 ravens and no writing desks. That is, there is nothing we know about Mr. X or about ravens and writing desks that would allow us to conclude that he isn't just playing some weird game or that he is an ornithologist selectively observing ravens instead of writing desks, or that by chance only ravens have come his way. Goodmanian make believe disjunctions, then, are those where there is no disjunct-distribution law operating in the background to make the disjunctive predicate projectable.

Part of the problem with denying that *all* disjunctive predicates are projectable can be seen in the following. Suppose you want to buy a bracelet for your wife and a friend says "I can sell you a jade bracelet." Furthermore, suppose that neither you nor your friend can distinguish jadeite from nephrite, though you are familiar with the fact that jade can be either jadeite or nephrite. That is, you know only that the bracelet is

jadeite or nephrite. Suppose you go from jewelry store to jewelry store encountering objects that are described as being "either jadeite or nephrite, but we don't know which." And imagine that every such object out of thousands observed was green. Wouldn't it be reasonable to conclude that your friend's "jadeite or nephrite" bracelet will be green, despite the mere outside chance that all the bracelets you observed were jadeite and not nephrite? Surely, as long as there is a disjunctive-distribution law operating somewhere in the background, then you have confirmed that your friend's bracelet will be green, even though you don't know if it is jadeite and not nephrite or nephrite and not jadeite.

So, some disjunctions of lower-level predicates are projectable and others aren't. But "is jadeite or nephrite" and "is jade" are equally or comparably projectable when we realize that there is a disjunct-distribution law operating in the background. Thus, there is no difference in the projectability of the MR predicate "is jade" and the disjunctive predicate "is jadeite or nephrite" that suggests that higher-level MR picks out properties other than being jadeite and being nephrite. I conclude that P1 is too strong, even though it is true that some disjunctive predicates, i.e. those that aren't aided by a disjunct-distribution law, aren't projectable.

It might be objected that this criticism of P1 only works for the case of jade, since it has not been shown that *every disjunctive predicate* that picks out the realizers of an MR predicate will be aided by a disjunct-distribution law. That is, I need to show that it is plausible that every projectable MR predicate has an associated lower-level disjunctive predicate which is also projectable because it is aided by a disjunct-distribution law.

We obviously can't consider each example of a putative MR property. So the best response to this objection is to take the strongest, best example of an MR predicate that is

supposed to be projectable although its lower-level disjunctive correlate is not projectable, and show that even in this case, there is a disjunct distribution law which guarantees that the lower-level disjunctive predicate is just as projectable as the higher-level predicate. A suitable example is the predicate “is in pain.” It is a strong example for the nonreductivists because it is hard to deny that there are general laws about pains that are confirmed by observing pains, and “being in pain” certainly seems to denote a property of suffering sentient organisms, and the disjunctive set of realizers of pain is a massive, “open” disjunction of different realizers.

Before I argue that even the disjunctive set of the realizers of pain is projectable, it is useful to examine what Fodor says about this case. Indeed, Fodor’s (1997) reformulation of the projectability argument, specifically P1 is found in the following passage:

> “Nonetheless, I think the intuition that open disjunctions are at best bad candidates for laws is basically sound. Here's why: Open laws suggest missed generalizations. To offer a law of the form RI v R2 v...--> Q is to invite the charge that one has failed correctly to identify the property in virtue of which the antecedent necessitates the consequent. Or, to say the same thing the other way around: Someone who offers such a law undertakes a burden to provide positive reason that there isn't a higher-level but nondisjunctive property of things that are RI v R2...in virtue of which they bring it about that Q. But we still haven't got to the bottom. No doubt, if there is a higher-level property that subsumes all the disjuncts of an open disjunction, then we will want to state our laws in terms of it. But why is that? Not, surely, because we are prejudiced against disjunctions 'as such'? Rather, I think, it's because formulas that express closed laws are stronger than the corresponding open ones and, ceteris paribus, we want to accept the strongest generalizations that our evidence confirms. Accepting the strongest generalizations that one' s evidence confirms is what induction is about. I'm suggesting that the intuition that open disjunctions invite closed laws isn't different in kind from the intuition that open lists invite universal generalizations. Swans are white constrains more worlds (hence supports stronger counterfactuals) than x is white if (x is swan A, V x is swan b, V x is swan c, ...). So too, and for the same reason, pain causes avoidance constrains more worlds than x causes avoidance if (x is neural state a, V x is neural state b.... V x is silicon state f.. , V x is Martian state g, V...). The cost of the former generalization is reifying not just the properties of being-swan-a and being-swan-b and being-swan-c.... etc. but also the more abstract, higher-level property of being-a-swan. The cost of the latter generalization is reifying not just the properties of being-in-neural-state-a, and being-in-neural-state-b.. .and being-in-silicon-state-f...and being-in-Martian-state-g, etc. but also the more abstract, (Special Sciences / 159) higher-level property of being-in-pain. Pretty clearly, standard inductive practice is prepared to hypostatize in both cases. And the success of standard inductive practice suggests pretty clearly that it is right to do so.” (1997, 158-159)

So, Fodor says that "open" disjunctions of lower-level predicates are" bad candidates" for forming laws and closed higher-level predicates are good candidates for laws. This, then, is his restatement of P1 of the projectability argument: sometimes there are higher-level MR predicates –which are single "closed" predicates- where the disjunction of predicates picking out all of the realizer properties of that MR predicate would have to be an "open" disjunction, i.e. "Is an A or a B or a C…"

So, what exactly does Fodor mean when he says there are "open" disjunctive predicates, and why are they bad candidates for projection? Well, one kind of open disjunction is a disjunction that contains so many disjuncts that we would be unhappy writing them all down. (Three disjuncts seem to be the limit before most authors will introduce ellipses.) For example, it would take a lot of time and effort to list all the realizers of "is a mammal." We can start by writing "is a cat, is a dog, is a whale," but before we get far, we just write "…" to let the reader know we could have finished the list, but we have decided not to. Let's call these kinds of open disjunctions, "large disjunctions"

Another kind of open disjunction is that in which we try to write all *the actual or possible* realizers of "is in pain." The fact is, we just don't know enough about the neurological basis (or bases) of pain to do so. Even though we might be able to identify what realizes the sensation of pain in the brains of human patients Mr. X and Mr. Y, we can't be sure there are not very different realizations in Martians or mad humans (to use Lewis's examples). So, we start listing the microphysical realizers of pain by saying "is in neurological state x, or y, or…" Here the "…" indicates a lack of knowledge about

whether and how many other realizers there are. Accordingly, I call this kind of open disjunction an "epistemically open" disjunction.

Fodor says there is a difference in the projectability of closed, MR predicates and the open disjunctions that pick out their various realizers. This allows Fodor to claim a disanalogy between jade -which he thinks is reducible to being jadeite or nephrite- and what he thinks aren't reducible MR properties like being in pain or being an economic transaction. The disanalogy is that the set of putative realizers of being pain is an open disjunction whereas the set of realizers of jade is a closed disjunction. If the open disjunction isn't projectable, but the associated higher-level predicate MR is projectable, then that MR predicate must pick out a higher-level, MR property. That, then, is Fodor's version of the projectability argument.

In response to Fodor's version of the projectability argument, consider large disjunctions. The set of all *actual* realizers of pain, i.e. all the realizers of pain we come across in the real world, is a very large disjunction: "is Mr. X's neurological state, or is Mr. Y's neurological state, or is such and such neurological state in this octopus, or …" But it is not an epistemically open set. We do not yet know exactly what goes on in each neurological state that realizes pain. (Of course, a philosopher knows far less than a neurologist.) Nonetheless, we can say with reasonable confidence that the actual realizers of pain are things that we have observed. It might be objected that it is possible that Martians have pains realized, as Lewis imagines, by fluid filled cavities in their feet. True, that is *possible*. But the issue here is the disjunction of realizers of pain we know to be actual.

As a result, it is reasonably clear that there is a disjunct-distribution law for the large disjunctive set of all actual, observed realizers of pain, just as there is a geological distribution of jadeite and nephrite which makes each more or less equally likely to be found. Moreover, we know that jadeite and nephrite look superficially similar, and have the same hardness and feel. This implies that if we go around observing things that look to be either jadeite or nephrite, and make enough observations, we are guaranteed to observe both instances of jade and instances of nephrite. The same is true of the actual realizers of pain, e.g. the realizers of pain in humans X, Y and X, the realizers of pain in abnormal humans A and B, the realizers of pain in giraffes… Animals and humans are distributed in the way jade is. And the realizers of pain are going to look (feel) the same, superficially in the way that jadeite and nephrite look similar. That is, different persons and animals are going to display the same behaviors in the way that jadeite and nephrite display the same color and hardness. (The same, I think, would be true of we could directly observe how pains appear, internally, to humans, giraffes, etc.) Thus, if you observe sufficiently many objects from the disjunctive list of actual realizers of pain, you are guaranteed to observe every disjunct.

We can therefore see that the large set of actual realizers of pain is just as projectable as the closed predicate “is in pain.” Thus as long as we stick to actual, observed instances of realizers of pain, Fodor’s version of the projectability argument doesn’t go through.

I suspect Fodor would be willing to agree that large disjunctions aren’t what he means by open disjunctions. After all, I can imagine someone thinking that a disjunction of 4 disjuncts A,B,C, and D is just barely too large to state. Should we conclude that this

disjunction is less projectable than the two-disjunct disjunction, “jadeite or nephrite?” I think not. So how is 5 disjuncts any different? Or 10? Or 1000? It seems arbitrary to draw a line as long as there is a back-up disjunct distribution law. So, I think Fodor really has epistemically open disjunctions in mind.

This leaves one last question: Are there higher-level projectable predicates with epistemically open disjunctions of lower-level predicates which aren’t projectable? Here I think the answer is clearly “yes,” and it is these epistemically open disjunctions that Fodor wants to use in his version of the projectability argument.

It is pretty clear that epistemically open disjunctive predicates aren’t projectable. The reason for this should be obvious. Suppose all actual jade is either jadeite or nephrite, but jade could’ve been realized by any number of heretofore unknown alien materials, magicite, absurdite, confusite, deceivite, and more. Suppose we have confirmed that all jadeite and nephrite are green. Have we confirmed that all magicite is green? Surely not. Maybe magicite is green like jade on some occasions, but turns a different color, magically, on other occasions. That is, if we are dealing with a disjunctive set that is epistemically open, i.e. if there are things in that set that we don’t know about, then we can’t confirm any universal generalizations about that set. This problem might be tolerable if there were one or two things in the set that we didn’t know about, but if there are untold, unknown numbers of unknown things in the set, any time we try to confirm a universal generalization about the set, there is no guarantee that some of the unknown things in the set won’t violate the generalization.

As a result, it is clear that a disjunctive predicate picking out the epistemically open set of all *possible* realizers of pain cannot be projectable. Of course, I think Fodor

will claim otherwise. He will claim that we can't confirm that all realizers of pain are going to result in certain behaviors like avoidance or saying ouch, but we can confirm that all creatures that are in pain, even those crazy Martian's whose pains are realized in their feet, will engage in common pain behaviors, or rather behaviors common to their species. Thus, instances of closed MR predicates can confirm general laws about pain that cannot be confirmed by instances of the associated epistemically open sets of realizers.

Part of the problem with many of the examples cited in the projectability argument, especially Fodor and Putnam's preferred examples of functional properties, is that there are analytic truths about some of these predicates, and Putnam and Fodor may be conflating laws that must be confirmed empirically with mere *analytic* truths. For example, it is true that all mousetraps catch mice. Does that mean the sentence "This is a mousetrap and it catches mice" counts as confirming "All mousetraps catch mice?" The answer must be "No." The probability of an analytic statement being true is already 1, so the statement cannot be confirmed by evidence, since "confirming" means increasing the likelihood that the confirmed statement is true.

By comparison, suppose Fodor were to claim that we have confirmed, as a higher-level generalization, that all creatures who are in pain will exhibit pain behavior. And, of course, all the instances of confirmation we have of that law are observations of humans and other animals, not Martians or other possible realizers of pain. Fodor thinks that we would be justified in believing that, if we were to encounter Martians, and if were to discover that one of them was in pain, that we should expect that Martian to engage in pain behavior. But what sort of pain behavior? Certainly, we should not expect the

Martian to behave as we do by wincing or crying or saying "I am in pain" in English. If Fodor Fodor's account of confirmation suggests that we should expect the Martian to behave as we do, because we have confirmed a general law relating pain and behavior for all possible realizers of pain, then so much the worse for Fodor's account of confirmation and laws. Fodor could respond by saying that we shouldn't expect the Martian to behave as we do or as animals do in response to pain; rather we should expect the Martian species to have their own set of reactions to their pain. But the law-statement "All creatures will respond to their pain with their own pain caused behavior" is analytic. It is just a truism that any creature in any psychological or physical state will react in someway to being in that state. Indeed, suppose the ordinary Martian reaction to pain is to do nothing. That would be Martian pain behavior and it would be predicted –if we can call this predicting- by the analytic claim "All of the possible realizers of pain will exhibit their own sort of pain behavior."

What this shows us is that analytic truisms, like "Axes chop" or "Computers can compute answers," which deploy higher-level predicates, can appear to be more confirmed than any generalization made deploying lower-level predicates. But this is an illusion. Analytic statements don't say anything that can be confirmed. And if we interpret them as saying something that needs to be confirmed, as in the case of expecting specific sorts of Martian pain behavior above, they will turn out to be no more and no less confirmed than a similar statement deploying lower-level predicates describing the actual observed realizer properties.

So, there are two responses to the objection that we can confirm generalizations about pain that can't be confirmed about the realizers of pain. First, someone making this

objection may be thinking of “is in pain” as a functional predicate, i.e. a predicate which describes all objects that are disposed to engage in pain behavior. If so, the so called laws that the objector is pointing to, which he thinks can be inductively confirmed, are really analytic truths about pain, e.g. “All things that are in pain, functionally defined as a state that result in pain behavior, exhibit pain behavior.” That analytic truism is not and cannot be confirmed by anything. Its confirmation value is already 1. Or, if the predicate “is in pain” is thought of as picking out the qualitative property that we are familiar with internally, i.e. the quale of being in pain, then it is false to say that our observations of actual pains confirm something about all possible creatures that could experience pain. That is, “I am in pain and am avoiding its source” does not serve to confirm “All possible creatures who are feeling the same qualitative pain we feel will engage in the sort of pain behavior we engage in.” Again, it is entirely possible that Martians, whose pain is realized by fluid filled sacks in their feet will not engage in the same sort of pain behavior that we will. We are as ignorant of how Martians will react to their pain as we are of how they would realize pain.

Antony and Levine (1997) have a way of responding to all of this. They think there is at least one generalization, which is not just an analytic truism, which is only projectable at the higher-level. They (1997) write,

> We think that such regularities do exist. To see what we have in mind, consider the second-order property of dormativity. Dormativity is a property a sub-stance possesses if and only if it possesses some first-order physical property that induces sleep. (We can leave aside niceties about dosages and clinical populations.) There is, therefore, an analytic regularity in which the property of dormativity participates: "Dormative substances cause sleep." But arguably, the fact that this regularity holds does not certify the existence of distinct causal powers associated with dormativity, any more than the fact that C's cause E's certifies the existence of a distinctive C-power. There are indeed some empirical regularities that hold of many dormative substances: many of them cause dry mouth, for example. But this regularity is realization-specific. It holds for some but not all of the realizers of dormativity: antihistamines and tricyclics, for example, but not barbiturates. However, consider this generalization: "*dormative substances, if ingested before driving, cause traffic accidents."* This generalization is true of anything that induces sleep, regardless of the mode of realization. It is an empirical, realization-independent regularity. Note in

> this connection that while this is a ceteris paribus law, there is no reason to expect that the exceptions will be systematic with respect to the realizations; the conditions that must be equal may all have to do extrinsic contingencies (like the density of traffic at the time the driver falls asleep) that are themselves realization-independent. We contend that the availability of this sort of generalization certifies "dormativity" as a genuine property, associated with a distinctive set of causal powers. [italics mine] (pp. 92-93)

But this example doesn't work. This isn't a clear case where the generalization is true "regardless of the mode of generalization." For example, lullabies are instantiated by sound waves. This is distinctively relevant to whether they cause accidents. Suppose I hear a lullaby while I am driving. I can choose to ignore the lullaby. At the lower, realizer level, my ignoring the lullaby is explained by the fact that my brain has a way of reacting to sound waves differently than it does to "tricyclic" chemicals. Antony and Levine might respond by saying a lullaby isn't a substance. The proper retort to this response is a thought experiment. Suppose scientists come up with a substance called Lullazac, which is dormative. The great thing about Lullazac is that it doesn't cause someone to be sleepy unless they want it to, i.e. people can will themselves to "ignore" the effects of Lullazac as they can do with lullabies. Of course, Lullazac would have to be realized by a very different chemical than ordinary sleeping pills, since we know that the effects of these chemicals are involuntary. If Antony and Levine object that Lullazac is an imaginary substance, or that there is something wrong with my example, then I can point them to the fact that warm milk is dormative and relaxing, but it does not cause accidents because the chemical effects of milk on the brain are soporific, but they are soporific in such a way that the brain can quickly snap to attention if danger -in particular the possibility of a car accident- looms.

Indeed, with the case of dormative substances, what we project for the realizers is exactly what we should project for the realized properties and vice versa. For example, a dormative substance is likely to cause car accidents if and only if it is realized by a

chemical –or something else like a song, as in the case of lullabies- that affects the brain in such a way as to cause the brain to go into one of the states described by the predicate "is asleep" *and* the realizer effects the brain in such a way that the brain does not have the power to counteract the soporiphic effect of the "dormative" substance.

VI. Antony, Disjunctive Properties, and Type Theory

In a more recent paper, Antony (2003) provides another way of arguing from projectable predicates to higher-level properties. She discusses the same arguments from Fodor (1974 and 1997) and Putnam (1975) and then writes:

> But I think this argument goes wrong. Fodor seems to be worried that if we cannot distinguish functional (and more broadly, higher-order) properties from the disjunction of their realizers, then the reductionist will have won. But what form of ''reductionism'' can claim this victory? Suppose, for the moment, that we bite the bullet and simply identify presumably higher order mental properties with their lower-order disjunctive associates. If the latter count as ''physical'' properties, then we would be acceding to the weak reductionist thesis that mental properties are physical properties. But so what? We can still maintain that mental properties are distinct from any of the disjunct properties… (p. 7)

Her argumentative strategy here involves Clapp's (2001) notion of a disjunctive property, i.e. the property of being A or B, and it involves it in a very clever way. For Clapp (2001), the property of being colored, for example, is just the disjunctive property of being red or blue or brown, and so on. Notice that a disjunctive property is supposed to be different from any one of its disjuncts in the same way that a conjunctive property is supposed to be different from its conjuncts.[79] So, if Antony can prove that a disjunctive predicate is projectable in a way that none of its disjuncts are, then, it would appear she can prove that there are disjunctive properties, picked out by disjunctive projectable predicates that are not identical to any particular base-level property. These disjunctive

[79] One of Clapp's (2001) arguments for disjunctive properties is to point out the hypocrisy present in, for example, Armstrong's acceptance of conjunctive, compound properties –e.g. being red and round- but not disjunctive properties, e.g. being red or blue. This argument will not work against any strongly reductive view, such as the one presented here, which does not recognize compound, conjunctive properties.

properties would then, it seems, have to exist over and above their individual disjuncts in the way that emergentists thought emergent properties of composite objects had to exist over and above the properties of the parts of said composite. At any rate, Antony has cited a datum that appears to point towards higher-level properties that needs to be explained without positing properties other than the properties picked out by individual microphysical predicates.

The weak point in her argument is found halfway through, where she introduces an odd caveat. She (2003) states, "What I really think is that there is no such thing as a ''disjunctive property''-rather, there are only disjunctive predicates."(p.9) This is an interesting claim, because it allows her to avoid any worries about the metaphysics of disjunctive properties. She further explains her position by suggesting that if a disjunctive predicate composed of individual microphysical predicates as disjuncts is projectable, this shows that it picks out a property present in all the objects described by the predicate, and this property isn't the same as -and thus exists separately from- any of the properties picked out by the disjuncts in the disjunctive predicate. That is to say, Antony wants to argue that higher-level properties are singular properties which are picked out by disjunctive predicates and singular higher-level predicates really just are disjunctive predicates of this sort.

I want to argue that Antony's position rests on an unwarranted and mistaken view about disjunctive predicates and the properties they pick out. But explaining the assumption and how it is unwarranted requires that we examine some basic concepts from Russell's (1903 and 1908) type theory in a very informal way.

One of Russell's concerns was that there are expressions like "the set of all sets that are not members of themselves" that give rise to paradoxes. Suppose we imagine that there is such a set which we will call R. Then R cannot be a member of itself, by its very definition. But if R is not a member of itself, then it must, again by its very definition, belong in R, which would mean R does contain itself, which violates the definition of "R." Russell rightly diagnoses that the problem with such expressions is that they're defined circularly, and they stand in violation of the vicious circle principle, i.e. the term to be defined is referred to in the definiens.

A little more generally, the problem with the expression "R" or the property R is that it quantifies over elements that include itself. If we were to define R formally, it would look something like this, where S is any set: $(\forall S)(S \in R$ if and only if $S \notin S)$. Notice that S is supposed to quantify over all sets including sets that contain only basic elements, sets that contain sets, and sets that contain sets of sets, including itself, i.e. the set R. Here's where some simple type theory comes in. Let's say sets that contain only urelements are type 1 entities. (We can say the urelements are type 0 entities.) And we will say the sets containing sets of type 1 are type 2 entities. (For our purposes it doesn't matter, but they will still be type 2 sets even if they include type 1 *and* type 0 entities as members.) And type 3 sets will have type 2 sets as members and so on.

Again, we needn't get into the details of type theory, but there is a simple way of expressing the vicious circle principle in terms of type theory and showing that expressions like R are illegitimate. We can say no unique property is picked out by a predicate if that predicate is defined in such a way that its definition deploys a quantifier ranging over entities where the set containing all and only those entities has a type of n

and the members of that set have a type of n or greater. To borrow another term from type theory, let's call any property which doesn't meet this criteria an "impredicative" property, and we will say that the expression that picks out such a property is defined impredicatively. One way of expressing the vicious circle principle, then, is to say all impredicative properties must be banned and all impredicative definitions are illegitimate.

However, a ban on all impredicative properties and definitions is too strong. That is, some impredicative properties cause trouble and must be banned, while others are harmless. For example, there are expressions like "Napoleon has the properties of all great generals" which are seemingly quite harmless and generate no paradoxes whatsoever. Banning the existence of such a property as having all the properties of great generals seems too strong.

One solution is to weaken the ban on impredicative properties. Part of Russell's solution (1903 and 1908) to this problem is to introduce a criterion for determining which impredicative expressions should be banned and which shouldn't, (Once again, we are leaving aside worries about ramified type theory and a number of important details of type theory more generally.) He calls this criterion, the "Axiom of Reducibility" and in the simplest informal terms, it states that impredicative expressions are legitimate and meaningful if and only if they are extensionally equivalent, i.e. "reducible to," properties which aren't defined impredicatively.

For example, we could list the properties of great generals. Napoleon is a great general and he possesses the properties of being daring, cunning, short, and French; whereas Rommel possesses the properties of being cunning, being daring, being German,

and so on. Let's say that these properties are type 1 properties and let us say that P ranges over all the type 1 properties.

Well, what should we make of the predicate "has all the qualities of great generals?" It would appear to pick out a property, too. Indeed, it would appear to pick out a property belonging to generals like Napoleon. As such, it should be a type 1 entity. Let's call that property G. Here is how we would have to define G: *(∀P)[(∀y)(y is a great general ⊃ Py) ⊃ Px]*. But if that is the case, then G is an impredicative property because it is defined, impredicatively, in terms of type 1 properties. But that is the exact sort of thing that is supposed to be ruled out by the ban on impredicative expressions used to ban paradox-creating expressions like the "set of all sets that don't contain themselves as members." Here we can see how the ban on impredicative properties is problematically strong.

The solution to this problem, as I mentioned above, is to say that the property G is extensionally equivalent to or "reducible to" some type 1 property(s) which can be defined without the use of a quantifier which ranges over type 1 entities. Indeed, Russell's Axiom of Reducibility (AR) solves the problem along these lines. In simple informal terms, AR states that an impredicative property, i.e. a property X of type n which is defined with quantifiers ranging over entities of type n, should not be banned if and only if X is extensionally equivalent to some property of type n which is defined *without* using quantifiers which range over type n entities.

And the property of having all the properties that great generals have certainly is reducible -i.e. extensionally equivalent to- such a type 1 property. Let's say all great generals share the property of being daring and cunning. Notice, that it is quite clear that

the expression “has all the qualities (properties) of great generals” really just is a shorthand way of picking out the type 1 properties of being daring and cunning. Now, if there were a property of being daring and cunning, which was at all distinct from these type 1 properties, then that property would have to be impredicative. It is absolutely critical to note that AR would ban such properties, i.e. any property which was distinct from a type 1 property. Indeed, AR implies that impredicative properties are only to be accepted if they turn out to be reducible to or identical to type 1 properties.

Notice that the expression “has all of the qualities of great generals” does not refer to specific qualities, e.g. daring and cunning. One reason we might use such an expression is if we are in epistemic ignorance about exactly which qualities are the qualities of great generals. Or perhaps we don’t want to list all these qualities, e.g. perhaps there are 313 such qualities, so we need a shorthand way of referring to all 313 qualities. In either case, it seems quite plausible to say that “all of the qualities of great generals” actually picks out type 1 properties, just as “is daring and cunning” picks out type 1 properties. The only difference is that the former is defined impredicatively.

The point I wish to make here is this: if AR is true, it would be a mistake to think that impredicatively expressions imply the need to posit impredicative properties which are distinct from type 1 properties that can be defined predicatively. But this is the mistake that Antony makes. She thinks that there is a higher-level property picked out by a disjunctive predicate which is distinct from the properties picked out by any of the disjuncts.

Suppose we have a series of axes in front of us. Each axe has a collection of base-level properties. On my account each axe will have a collection of microphysical

properties, but for the sake of simplicity let's think of these base-level properties just as simple properties like hardness, sharpness, is one meter long, is one foot long, is made of iron, is made of steel, is double bladed, and is single bladed. Let's call these base-level properties, type 1 properties.

Antony believes that there is a property of having some one of a number of these properties in virtue of which an object is an axe, and this property is distinct from the type 1 properties. As a result, it is a higher-level property of the object that possesses the type 1, more base-level properties. And her view can seem correct because the putative higher-level property she is discussing is defined in terms of type 1 properties. That is, something has the property of being in axe in virtue of having some one of a number of other properties. For example, she would say that there are type 1 properties an object possesses like sharpness and being made of steel, and an object has a separate type 1 property of being an axe. Again, the fact that we define the property of being an axe in virtue of the type 1 properties does seem to suggest that the property of being an axe is distinct from any of those type 1 properties.

Perhaps this is extending AR a little bit, but AR implies that any type 1 property defined in terms of other type 1 properties must be reducible to those other type 1 properties. That is, AR implies that Antony's properties either don't exist or they are, *pace* Antony, identical to properties picked out by base-level predicates employed as disjuncts in the disjunctive predicates she is arguing about.

One possible reply Antony could try to give is that the higher-level properties she is arguing for are not intended to be type 1 entities at all; they are intended to be type 2 or higher. That is, she could say that the higher-level properties are type 2 properties in that

they are properties of properties. For example, she could say being an axe is a property belonging to more base-level properties like being made of steel.

But it is pretty clear that this objection won't get off the ground. It may be that there are type 2 properties of properties. Perhaps a pain possesses the property of being achy. And perhaps the achiness possesses the property of being irritating. Personally, I suspect that there are no such properties; rather there is a type 0 object, in this case an experience, which possesses the type 1 properties of being painful, achy, and irritating. But even if there are type two properties of properties, it is quite clear that being an axe is not one of them. Being an axe is not a property belonging to the property of sharpness or the property of being made of steel, any more than being this laptop computer is a property of being black. Like I say, it is possible that there are properties of properties. But certainly properties like being an axe or being a computer are type 1 properties of objects.

Obviously, we can't prove that we should accept AR. It is an axiom, after all. But certainly, AR, even the metaphysical version of AR, is a plausible assumption. Indeed, AR may be an essential axiom in the study of logic. So, Antony cannot argue in reply to what we have here discussed that the axiom of reducibility is false without some compelling reason. She can only argue that the axiom of reducibility should not be generalized in the way suggested here.. But it is not clear what reason she could cite for this conclusion.

In fact, there are very compelling reasons to generalize the axiom of reducibility along the lines needed to deflate Antony's argument. The axiom of reducibility rests on a simple and intuitive idea about the semantics of predicates and referring expressions

more generally. That idea is just this. We can come up with short hand ways of describing sets of type 1 properties. I can say, “has the properties of all the axes on the table” or “has the properties of one of the axes on the table.” But this shorthand expression is intended to pick out the type 1 properties of the axes. Again, there are a variety of reasons we might come up with such a short hand impredicative expression. Perhaps we don’t want to list all of the properties of all of the axes. So, for the sake of brevity, we say, impredicatively, “has all of the qualities of the axes on the table.” But, again, when we say that, we are intending to use the shorthand phrase to pick out a cluster of individual type 1 properties of axes. We do not intend to say there is a property of having the properties of the axes on the table.

Another, more relevant case, occurs when we find ourselves in epistemic ignorance about what the type 1 properties there are but we want to talk about the properties nonetheless. For example, we might want to talk about one of the properties, even though we don’t remember exactly which one. In that case, we would use a phrase like “has one of the properties of the axes on the table (I am just not sure which)” But here again, then the predicate is meant to pick out a type 1 property. That is to say, it is intended to pick out the same property as some predicate that picks out a type 1 property.

These examples show how we use legitimate impredicative expressions, and that legitimate impredicative expressions are merely indirect or shorthand ways of referring to or picking out type 1 properties. So, an examination of the meaning and use of impredicative expressions, including disjunctive expressions like “has one of the properties of the objects on the table,” seems to suggest that we should generalize the axiom of reducibility for all such expressions.

And so it appears that the sort of disjunctive predicates Antony mentions are just a particular form of impredicative expressions. (I would argue that disjunctive predicates are impredicative expression where we are in epistemic ignorance of the type 1 properties we wish to pick out. But that is less important.) And impredicative expressions are – according to the axiom of reducibility and what we know about how we use impredicative expressions- merely indirect ways of picking out type 1 properties, then it certainly seems like we should conclude that disjunctive predicates do not pick out properties other than the type properties their disjuncts pick out. Thus, Antony's case for the existence of higher-level properties which are picked out by disjunctive predicates falls apart.

At the very least, AR gives us a plausible way of explaining away the premises in Antony's argument for higher-level properties. As long as AR is true, and AR is at least plausible, then there are no higher-level properties corresponding to projectable disjunctive predicates which are distinct from the properties referred to by the individual, base-level disjunct predicates. Rather, if AR is true, Antony's higher-level properties would have to be reducible to the base-level.

VII. Conclusion

It may be that in the future some new piece of evidence will surface that will support the existence of higher-level properties. Or it may be that there is some piece of evidence that we have overlooked. But as it stands now, all of the evidence for the existence of higher level properties that we have catalogued and discussed is insufficient to show that we should not accept the semantic view of levels. The semantic view of

levels is capable of explaining why we have projectable, pragmatically useful higher-level predicates and *ceteris paribus* higher-level law-statements. It can explain the appearance of multiply realizable properties. And it can explain why composite objects appear to have properties that emerge from the properties of their parts. Moreover, we have seen that the intuitions that sometimes seem to weigh in favor of the levels doctrine are not always so straightforward, and others are misleading.

I am sure that many will find that what I have argued here is not decisive. I respect that just as I respect the subtlety and sophistication of the various versions of the levels doctrine. However, I do think that we have learned, at the very least, that the semantic view of levels deserves consideration as an alternative view. I hope it will be considered as such and that other philosophers might try their hand at formulating their own detailed versions of the semantic view. This is at least an interesting avenue for further philosophical research.

Works Cited

Alexander, S. (1918). *Space, Time, and Deity: The Gifford Lectures at Glasgow*. London: Macmillan and Co. Available online at: www.giffordlectures.org/Author.asp?AuthorID=5

Antony, L. (2003).Who's Afraid of Disjunctive Properties. *Nous-Supplement: Philosophical Issues. Vol. 13*, 1-21.

Antony, L. (1999).Making Room for The Mental: Comments on Kim' Making Sense of Emergence. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition. 95*, 37-44.

Antony, L. (1999). Multiple Realizability, Projectibility, and The Reality of Mental Properties. *Philosophical Topics* 26:1-24.

Antony, L., & Levine, J. (1997). Reduction with Autonomy. *Nous-Supplement: Philosophical Perspectives,. 11*, 83-105.

Armstrong D. (1968) *A Materialist Theory of the Mind.* New York: Humanities Press

Armstrong, D. (1989). Universals: An Opinionated Introduction. (chapter 5), In Jaegwon Kim and Ernest Sosa (eds.), *Metaphysics:* An *Anthogy*. Blackwell Publishing

Baker, L. R. (2007). *The metaphysics of everyday life: An essay in practical realism*. Cambridge: Cambridge Univ Press

Bechtel, W. & Mundale J (1999). Multiple Realizability Revisited: Linking Cognitive and Neural States. *Philosophy of Science*, 66: 175-207.

Boolos, G. (1985). Nominalist Platonism. *Philosophical Review, 94*, 327-344.

Boyd, R. (1989) How to be a Moral Realist. In Paul Moser & J.D. Trout (eds.), *Contemporary Materialism: A Reader* (pp.297-356). Routledge: NewYork

Broad, C.D. (1925). *The Mind and its Place in Nature.* Routlegede and Kegan Paul. Available Online at: http://www.ditext.com/broad/mpn/mpn.html

Churchland, P. M. (1981). Eliminative Materialism and the Propositional Attitudes, *Journal of Philosophy* 78: 67-90.

Churchland, P.M. (1988). *Matter and Consciousness: A Contemporary Introduction to the Philosophy of Mind.* MIT Press

Churchland, P.S. (1986). *Neurophilosophy: Toward a Unified Science of the Mind/Brain*. Cambridge, MA: MIT Press.

Clapp, L. (2001). Disjunctive properties: Multiple realizations. *Journal of Philosophy, 98*(3), 111-136.

Dawkins R. (2000) *Unweaving the Rainbow: Science, Delusion, and the Appetite for Wonder.* Mariner Books

Davidson, D. (1970) "Mental Events." In Paul Moser and J.D. Trout (eds.), *Contemporary Materialism: A Reader* (pp.107-121). Routledge: New York, NY

Dupre, J. (1993) *The Disorder of Things: Metaphysical Foundations of the Disunity of Science.* Harvard University Press: Cambridge, MA

Evans, G. (1973), 'The Causal Theory of Names,' *Proceedings of the Aristotelian Society*, Supplementary Volume 47: 187–208.

Feigl, H. (1968). *The mental and the physical: The essay and a postscript*. Minneapolis: Univ. of Minneapolis Pr. Also Available at: http://www.ditext.com/feigl/mp/mp.html

Fisk, M. (1972) "Relatedness without Relations," *Noûs*, 6; 139-151.

Fodor, J. A. (1997). Special sciences: Still autonomous after all these years (volume 11: Mind, causation, and world). *Nous-Supplement: Philosophical Perspectives, 11*, 149-163.

Fodor, J. A. (1974). Special sciences (or: The disunity of science as a working hypothesis). *Synthese: An International Journal for Epistemology, Methodology and Philosophy of Science, 28*, 97-115.

Frege, G. "On sense and Nominatum." In Martinich, A.P. (ed.) *The Philosophy of Language*. Oxford University Press.

Goodman, N. (1983). *Fact, Fiction and Forecast.* Cambridge, MA: Harvard University Press.

Heil, J. (1999). Multiple realizability. *American Philosophical Quarterly, 36*(3), 189-208.

Heil, J. (2005). *From an Ontological Point of View.* New York: Oxford University Press

Hey, T. and Walters, P. (2003) *The Quantum Universe.* Cambridge University Press

Kant, I. (1999) *The Critique of Pure Reason.* Guyer P. and Allen W. (ed.) Cambridge University Press

Keats, J. (1884) *The poetical works of John Keats: reprinted from the original editions, with notes by Francis T. Palgrave.* London: Macmillan

Kim, J. (1989). The Myth of Nonreductive Physicalism. *Proceedings and Addresses of the American Philosophical Association*, 63: 31-47.

Kim, J. (1992). Multiple Realization and the Metaphysics of Reduction. *Philosophy and Phenomenological Research*, 52: 1-26

Kim, J. (1999). Making sense of emergence. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, *95*(1-2), 3-36.

Kim, J. (2005) *Physicalism or Something Near Enough.* Princeton University Press.

Kripke, S. (1980). *Naming and Necessity*. Harvard University Press

Lewes, G. H. (1875). *Problems of Life and Mind*, Volume 2. London: Kegan Paul, Trench, Turbner, and Co.

Lewis, D. (1983) In Jaegwon Kim and Ernest Sosa (eds.), *Metaphysics:* An *Anthogy*. Blackwell Publishing

Linnebo, Ø. (2003). Plural quantification exposed. *Noûs* 37:71-92. 2004. Plural quantification. *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta, Available online at: http://plato.stanford.edu/archives/win2004/entries/plural-quant/

Loar, B. “Can we confirm supervenient properties?” *Philosophical Issues* 4, Naturalism and Normativity: 74-92

Locke, J. (1979) *An Essay Concerning Human Understanding*. Nidditch, P. (ed.) Oxford University Press

Martin, C.B. (2008) *The Mind in Nature*. Oxford: University Press

McLaughlin B. (1992) "The Rise and Fall of British Emergentism," In A. Berckermann, J. Kim, and H. Flohr (eds.) *Emergence Or Reduction.* (pp.49-93) De Gruyter

Melnyk, A. (2003). *A Physicalist Manifesto: Thoroughly Modern Materialism.* Cambridge: Cambridge University Press

Merricks, T. (2001). *Objects and persons*. Oxford: Oxford University Press

Mill, J.S. (1843). *System of Logic*. London: Longmans, Green, Reader, and Dyer. (Eighth edition, 1872). (Relevant Chapter: On The Composition of Causes)

Morgan, C.L. (1923) *Emergent Evolution: The Gifford Lectures*. London, UK: Williams and Norgate. Available online at: http://www.giffordlectures.org/Browse.asp?PubID=TPEMEV&Cover=TRUE

Moser P.K, and Trout J.D. (1995) *Contemporary Materialism: A Reader.* New York: Routledge

Nagel, E. (1961). *The Structure of Science.* New York: Harcourt, Brace, and Wilson.

Pereboom, D. (2002). Robust nonreductive materialism. *Journal of Philosophy, 99*(10), 499-531

Phillip, C. (2006). *Mind and Emergence: From Quantum to Consciousness.* Oxford University Press

Putnam, H. (1975). Philosophy and Our Mental Life. In Paul Moser & J.D. Trout (eds.), *Contemporary Materialism: A Reader* (pp.122-132). Routledge: NewYork

Quine, W.V.O (1969) “Natural Kinds” In Kim, J. and Sosa, E. *Metaphysics: An Anthology* (pp. 233-242).Blackwell Publishing

Rosenberg, J. F. (2005). *Accessing kant: A relaxed introduction to the critique of pure reason*. Oxford: Clarendon Press.

Russell, Bertrand (1903) "Appendix B: The Doctrine of Types," in Russell, Bertrand, *Principles of Mathematics*, Cambridge: Cambridge University Press, 1903, 523-528.

Russell, Bertrand (1908) "Mathematical Logic as Based on the Theory of Types," *American Journal of Mathematics*, 30, 222-262. Repr. in Russell, Bertrand, *Logic and Knowledge*, London: Allen and Unwin, 1956, 59-102, and van Heijenoort, Jean, *From Frege to Gödel*, Cambridge, Mass.: Harvard University Press, 1967, 152-182

Searle, J. (1983), *Intentionality*, Cambridge: Cambridge University Press.

Schaffer, J. (2003). Is there a fundamental level? *Noûs*, *37*(3), 498-517.

Shafer-Landau, R. (2003). *Moral realism: A defence*. Oxford: Clarendon Press.

Shapiro, L. (2000). Multiple Realizations. *Journal of Philosophy*, 97: 635-654.

Shoemaker, S. (2007). *Physical Realization*. Oxford: Oxford University Press.

Smart, J.J.C. (1963). *Philosophy and Scientific Realism.* Humanities Press

Smart, J. J. C. (1959). Ryle on mechanism and psychology. *Philosophical Quarterly*, *9*, 349-355.

Smart, J. J. C. (1959). Sensations and brain processes. *Philosophical Review*, *68*, 141-156.

Stace, W.T. (1939). Novelty, Indeterminism, and Emergence. *Philosophical Review 48*, 296-310

Stoljar, D. (2009). Physicalism" In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta, Available online at: http://plato.stanford.edu/entries/mereology/

Unger, P. (1980). "The Problem of the Many." *Midwest Studies in Philosophy*. 5: 411-67

Varsi, A. (2009). "Mereology". In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta, Available online at: http://plato.stanford.edu/entries/mereology/

Wimsatt, W. (1994) The Ontology of Complex Systems: Levels of Organization, Perspectives, and Causal Thickets." *Canadian Journal of Philosophy*, supplemental vol #20, 1994. Available at http://www.institutnicod.org/Reduction/7.OntComplSys.pdf