# An Ideal Solution to Disputes about Multiply Realized Kinds

Colin Klein

**Abstract**

Multiply realizable kinds are scientifically problematic, for it appears that we should not expect discoveries about them to hold of other members of that kind. As such, it looks like MR kinds should have no place in the ontology of the special sciences. Many resist this conclusion, however, because we lack a positive account of the role that certain realization-unrestricted terms play in special science explanations. I argue that many such terms actually pick out idealizing models. Idealizing explanation has many of the features normally associated with explanation by MR kinds. As idealized models are usually mere *possibilia*, such explanations do not run afoul of the metaphysical problems that plague MR kinds.

## 1 Introduction

Under enough force, some things bend. Others just break. Materials science calls the latter sorts of things the *brittle* ones.[1] Hardened steel, glass, fired ceramic, and hard candy are all brittle. They aren't brittle for the same reason, though. Hardened steel is brittle because of the lack of atomic dislocations within crystalline grains. Glass is

---

[1] Note that unlike fragility, brittleness is not the feature of an *anthropocentric* kind. Brittle materials are simply those that fracture before they yield, or, more formally, those materials that have a stress-strain curve with no tangents ([Roylance, 1996] 211). Hardened steel and bricks are brittle but not fragile; although the force necessary to fracture these materials is far more than commonly encountered in everyday life, they nevertheless break before yielding.

not a crystal; it fractures because the molecules of a super-cooled silica liquid form easily broken bonds.

The set of brittle things is thus an example of a *Multiply Realizable* kind. The realizers of the members of this kind are heterogeneous from the perspective of a lower-level science.[2] So unlike *Realization-Restricted* kinds (like 'brittle steels', 'brittle glasses', etc.), MR kinds cannot be reductively explained by lower-level sciences. Since MR kinds are widespread in the special sciences, we should all embrace nonreductive physicalism.

Closer inspection reveals a serious problem. There is hardly anything we can *discover* about the brittle things. Materials science has told us much about the realization-restricted types. In steels, brittleness is proportional to hardness. Softer grades of steel deform under pressure; only harder grades fracture. Steel can be made less brittle through tempering, which re-converts some of the martensite formed during fast quenching into a structure with more crystalline dislocations. The addition of manganese to a steel alloy generally decreases brittleness, while a high sulfur content generally increases it.

None of this is true of glasses. Glass only fractures and never plastically deforms under stress; hence the brittleness of glass is uncorrelated with hardness.[3] Different chemicals have different effects when mixed with glass rather than steel. And so on.[4]

That's a problem. Virtually nothing true of brittle steel is true of brittle glass, and vice-versa. There is plenty to discover about realization-restricted types of brittle things. But there are few discoveries that hold of all brittle things. Because of this, discovering $\phi$ about one kind of brittle object—steels, say—doesn't give us any reason to believe that $\phi$ holds of any other kind of brittle objects. This

---

[2] I'll talk freely in what follows about 'higher' and 'lower' level kinds and sciences. This distinction has been used in many ways, some objectionable. I will use it in the straightforward compositional sense, first endorsed by Putnam and Oppenheim. An kind $A$ is of a higher level than an kind $B$ (or a collection of such kinds) just in case (1) $A$ and $B$ belong to different sciences and (2) the $A$s are at least partially composed of $B$s, but not vice-versa. Similarly, say that a science $\mathcal{A}$ is lower-level than a science $\mathcal{B}$ just in case some of the members of the kinds $A$ of $\mathcal{A}$ compose some of the members of the kinds $B$ of $\mathcal{B}$, but the converse is not the case.

[3] Molten mixtures of silica deform, but I shall follow convention and use 'glass' to refer to the super-cooled phase of these liquids; glasses are thus always brittle.

[4] These are just a few of the many differences. For an excellent introduction to Materials Science, see [Gordon, 1984]; for fracture mechanics, I have benefitted greatly from [Roylance, 1996] Chapter 7.

represents a serious challenge for defenders of MR kinds. If there are MR kinds, they must be proper scientific kinds. If they are scientific kinds, then we should be able to project generalizations about them across all instances of that kind. But there aren't any such projectable discoveries; it looks like we must therefore abandon MR kinds—and not just in metallurgy, but in all of the special sciences, and psychology in particular.[5]

Many have resisted this conclusion. No doubt this is in part due to problems in various formulations of the argument. But I suspect that the real resistance has a deeper root. After all, isn't it just obvious that the special sciences refer to things that span realization kinds? Isn't their literature chock full of references to brittleness, genes, and computers—the very paradigms of multiply realizable kinds? Certainly these facts must form a sound basis for a modus tollens of any argument against MR kinds.

Perhaps not. Working scientists talk about many things that span realization kinds, true. Consider a few more: ideal gasses, turbulence-free flow, and perfectly attentive perceivers. All of these are invoked to explain the behavior of a number of different kinds of things. They are not counterexamples to the above. None of these things have actual instances; they are *idealizations* of actual kinds.

I argue that many putative cases of MR kinds actually turn out to be idealized models. The use of idealizations thus complicates a simple view many hold about the special sciences. Philosophers of mind have often assumed that one could determine which mental states existed just by looking at what working psychologists happen to talk about. But idealized models do not exist; there are no ideal gasses or friction-free planes. If read as referring to the actual world, statements involving these models are false. Scientists know this, and idealize anyway. Therefore, not all scientific talk is ontologically committal, and careful work is required to determine the ontological commitments of theories.

It is work that is worth doing. In what follows, I set out the beginning of that project. First, I run through an argument originally due to Jaegwon Kim, suggesting that MR kinds don't meet minimal standards for inclusion in explanatory theories. I then introduce idealizations, showing that many features cited as characteristic of MR kinds are also features of idealized models. This helps explain the

---

[5] See for example [Kim, 1992, Kim, 1998, Millikan, 1999, Shapiro, 2000].

intuitions that led nonreductive physicalists to embrace MR kinds in the first place. I conclude by suggesting that many special science terms—including mental terms—are ambiguous. Sometimes they refer to actual, realization-restricted kinds while other times they refer to features of explanatory but nonactual idealized models.

## 2 The Argument against MR Kinds

It's worth devoting a bit to unpacking the argument suggested above. It is by no means a flawless argument, but it remains under-appreciated. Further, by spelling out precisely what the problem for MR kinds is supposed to be, we can make some headway towards understanding what putatively MR kinds might actually be doing.

First, arguments about MR kinds are best understood as arguments about whether they are proper *scientific* kinds. I will remain agnostic about whether there are any nonscientific MR kinds—surely whether you think so will depend on larger commitments that are not at issue here. The real issue is whether there are MR kinds that are worth including in our scientific ontology; it is in that capacity that they would do the reduction-blocking work that is distinctive of MR kinds.

I will focus on one feature that is often thought to be a necessary condition on scientific kind hood: that they are what Kim calls 'nomically projectable kinds' ([Kim, 1998] 109). That is, if we discover something about a number of members of a scientific kind, we should have reason to expect that discovery to hold of other members of the kind. Of course, not everything is projectable: members of a kind can differ. However, members of a kind should share a cluster of interesting, non-trivial properties that distinguish them from other scientific kinds. These properties are what we go looking for when we do empirical work on a kind, and what we appeal to when we appeal to kinds to explain.

Projectability of discoveries is not a sufficient condition for scientific kind-hood, but it is a plausible necessary condition. One of science's important jobs is to discover general facts about the kinds it uses. This is only possible if we can be confident that some of the discoveries we make about some members of a kind will be true of the other members that we cannot survey. Conversely, if we find out that many of our discoveries aren't in fact projectable, that should erode our confidence that we have successfully divided the world into useful kinds. Hence, if $\phi$ is a proper scientific kind, then we can expect at least some empirical discoveries about members of $\phi$ to be true of all other members of $\phi$.

This notion of a scientific kind allows us to make the notion of an MR kind more precise.[6] MR kinds are the ones whose realizers are so diverse that they don't have enough in common to form a type in a lower-level science. The brittle things form one such kind. Mental kinds are classic examples: the mental kind of 'being in pain' could be instantiated by neural events in humans, silicon processes in robots, or by really any process (it seems) that plays the appropriate causal role of pain. The realizers of an MR kind are not merely different. They are *so* different in their behavior that they are completely heterogeneous from the perspective of a lower-level science. As Fodor puts it, MR kinds *cross-classify* kinds at a lower-level [Fodor, 1975]. So if $\phi$ is an MR kind, then the set of realizers of $\phi$ must not form an empirical kind of some lower-level science.[7] Since the realizers of instances of MR kinds don't have enough in common, we can't explain the behavior of MR kinds by citing facts about the set of their realizers realizers. Thus we get the expected result that facts about MR kinds are *irreducible* to facts about the set of their lower-level realizers.

The conjunction of cross-classification and projectability makes trouble for any physicalist defender of MR kinds. Barring an unpopular commitment to emergentism, physicalists should believe that higher-level kinds don't have additional causal powers over and above the causal powers of their realizers. In other words, the properties of an instance of a high-level kind are *grounded in* the properties of the realizers of that instance. This notion of grounding captures a distinctive commitment of non-emergent physicalism: instances of a kind have the properties they do *because of* the properties of the lower-level entities that compose them. A sample of hardened steel is brittle. That brittleness is not *sui generis*, not something that emerges only

---

[6] Defining multiple realizability is a nontrivial problem, and has attracted a fair deal of attention in recent years. For a recent exchange see [Shapiro, 2000, Rosenberg, 2001, Clapp, 2001, Jaworski, 2002].

[7] A more rigorous treatment would make a kind MR relative to specific lower-level kinds, without assuming that if a kind is MR relative to one lower level kind it is MR relative to all lower kinds. However, since that is a common simplifying assumption in the literature (likely stemming from the transitivity of early models of reduction), I will also assume so here. It should not matter for my argument.

at the level of metals. Instead, it is brittle because of the causal and relational properties of the grains that compose it.[8]

Grounding implies in turn that differences between lower-level realizers should 'trickle up' and manifest as differences between instances of a higher-level kind as well. It is precisely the differences between realizers of brittle things that are responsible for differences between the brittle things themselves; the micro-differences between steels and glasses are what underlie the macro-differences between the different kinds. So consider some putative MR kind $\phi$, and suppose that it is a scientific kind. Since $\phi$ is MR, we know that we should expect the set $\rho$ of realizers of $\phi$ not to form a scientific kind. That is, for any property that we discover of a member of $\rho$, we should expect that there's another member of $\rho$ that lacks that property.

Since $\phi$ is a scientific kind, we know via projectability that there is some reasonably large set of properties shared by instances of $\phi$ and about which we could make empirical discoveries. By the grounding of high-level properties in lower-level realizers, we also know that any causal property of an instance of $\phi$ will depend on $\rho$ that instantiates it. But the set of $\rho$ things are too diverse to hang together as a kind. Grounding ensures that these differences should manifest as differences at the higher level as well. This means that instances of $\phi$ should not have enough in common to form a proper kind, generating a contradiction. Hence $\phi$ cannot be both MR and a scientific kind. Since the argument is entirely general, we may conclude by denying

---

[8] There are a number of ways to cash out 'grounded in', and I intend the argument to remain as neutral as possible between them. I am partial to Kim's "Causal Inheritance Principle" outlined in his ([Kim, 1992] 16), and the argument that follows is deeply indebted to his discussion of the CIP. However, Kim's own formulation is not without problems. For one, Kim claims that the causal powers of $M$ will be *identical* to the causal powers of $P$. This is contentious. While MR kinds may not have more powers than their realization bases, they may well have *fewer* causal powers. Such a line has been advanced by Sydney Shoemaker and Stephen Yablo as a way to make sense of the realization relation ([Shoemaker, 2001], [Yablo, 1992]). On this line, the causal powers of MR kinds form a proper subset of the causal powers of their realizers. Every MR kind picks out a collection of realizers that share certain causal powers. Given that this kind is MR, though, the realizers will differ in other of their causal powers. Hence, not *every* causal power of the realizers can be a causal power of the high-level kind. As Karen Bennett has brought to my attention, even Kim's later writings endorse a subset view (for example, in [Kim, 1998] 54). For another, not everyone in the literature thinks that a subset view is necessary for cashing out the realization relation—see for example Gillett's recent [Gillett, 2003]. Thanks to **[suppressed for review]** for pressing me on this point.

that there are *any* MR kinds in science.

No doubt there are many places to object—but the result is stronger than it may appear. First, it is not merely a problem with *justifying* generalizations about MR kinds. Fodor, for example, reads this point as a worry about about sample bias. If we confirm our discovery on only one realization-kind, we do not confirm it for all—but that's to be expected, for "*Biased* samples don't confirm *anything*" ([Fodor, 1997] 152). [9]

But that isn't the problem. It's not that we merely lack reasons to believe that $\phi$ will apply across realizations. Instead, we also have good reason to think that $\phi$ *won't* apply to other realizations. This is just because we have good reason to believe that members of $\phi$ don't have many empirically discoverable properties in common. As such, anything we find out about a set of $\phi$ things is unlikely to hold of all of the $\phi$ things. Grounding gives us every reason to expect that any $\phi$ that differs in realization from the ones we observe might also differ in its properties from $\phi$ things we have observed. As such, we should expect most generalizations about $\phi$ things to fail.

Second, variants of the argument will still run even if we weaken our constraint on kindhood in various ways. Regardless of what you think kinds are, you should think that MR kinds are *dispensable* in a well-ordered science.

This is in part because we *do* have good reasons to include realization-restricted versions of MR kinds in our scientific ontology. As projectable kinds, they bring with them obvious explanatory and methodological benefits. We can discover interesting clusters of facts about them, and use those discoveries to explain. Hence, even the most ardent supporter of MR kinds shouldn't disagree that we can include realization-restricted kinds into our scientific ontology, and benefit greatly thereby.

Once we include realization-restricted kinds, however, then MR kinds start looking superfluous. As Kim has emphasized, we can get by supposing that MR kinds (and functional kinds, and whatever) have a *conceptual* unity without referring to a single unified kind ([Kim, 1998] 110). MR kinds (like gerrymandered kinds) can be eliminated from scientific ontology without reducing the predictive or explanatory power of the theories they are supposed to figure in.

---

[9] Actually, that's a bit optimistic. Fodor himself claims that MR kinds apply to a potential infinity of realizers ([Fodor, 1997], p155). If so, we always have biased samples of MR kinds.

Whatever science can do with a putatively MR predicate, it can also do with the related concept. So I don't gain anything by supposing that 'brittle things' picks out a scientific kind: everything I can say about the brittle things I can also do if I treat 'brittle' as a simple shorthand for the set of realization-restricted brittle things. Once we have realization-restricted kinds in our ontology, it is easy to so augment a theory's vocabulary. MR kinds thus aren't any more useful than any other second-rate kinds; it looks science can get by without them, and loses nothing by eliminating them.

# 3 Intermission: The Stalemate

You might not buy the argument just given. If not, you probably reject it for the same reason most do: what seem like very obvious facts about the special sciences. Fodor, for example, responds to Kim:

> "The very *existence* of the special sciences testifies to reliable macro-level regularities that are realized by mechanisms whose physical substance is quite typically heterogeneous. Does anybody really doubt that mountains are made of all sorts of stuff? Does anybody really think that, since they are, generalizations about mountains-as-such won't continue to serve geology in good stead?"
> ([Fodor, 1997], p160)

Working scientists don't seem to care about realization. The explanations they give often use many kinds that look MR. And the special sciences still seem to work awfully well. Consider the oft-cited abstract computational models in cognitive psychology, which look like they can be implemented in any sort of material. Even the brittleness pops up to explain lots of things across all sorts of realization-restricted kinds. No matter how good the argument above looks, why don't such obvious facts just give the beginning of a modus tollens?

This debate usually ends in stalemate. Nonreductive physicalists have strong intuitions about the role of certain terms in the special sciences. Their opponents have given equally strong reasons to think that they're wrong. Without accounting for the intuitions of the nonreductive physicalists, folks like me can make little headway in this debate.

What we need is a good positive account of what putative MR terms are doing in explanations. A Kim-style account like the one suggested above—that is, an account that says that we should treat such terms as picking out *concepts* rather than kinds—accounts for the use of many such terms in causal explanations of token events.[10] However, proponents of MR kinds often stress their use in explaining *generalizations* as well as token events. By cross-cutting realization kinds, it is sometimes thought, MR kinds give good explanations of high-level generalizations. To break the stalemate, therefore, I propose that we also need a positive account of these explanations.

I claim that putatively MR terms in explanations of scientific generalizations in fact refer to elements in nonactual idealizing models.[11] Idealization is a common scientific tool; scientists often turn to it when a phenomenon is too complex to study directly. Idealized models are especially useful when one wants to explain generalizations about complex systems. They are thus a plausible candidate for the actual referent of putatively MR terms.

Although I claim that many putatively MR terms refer to idealizing models, it is beyond to the scope of this paper to properly *prove* that. To do so even for a single special science would require a careful look at the kinds it typically uses, showing that they in fact refer to either realization-restricted kinds or else elements in an idealizing model.

That is a large task, and I will here content myself with the more modest goal of making the thesis plausible. First, I will show that the special sciences *do* use idealizing models. Second, I will show that idealizing models have a number of features that are usually cited as characteristic features of MR kinds. This makes it plausible that the nonreductive physicalist has *confused* reference to a scientifically problematic category of things (the MR kinds) for reference to a scientifically legitimate category of things (the idealizing models).

Before we get to the examples, I want to motivate what follows by way of a broad observation. The stalemate we have reached is only tenable given a very specific view of science. Call this view *Literalism.* According to literalism, the primary task of scientists is the construction of accurate theories. Theories pick out real laws that connect and explain real causal processes in the world. Assuming a mild realism,

---

[10] That is, MR concepts allows for indirect reference to a specific, realization-restricted causal kind. See Kim's [Kim, 1998], especially Ch4, for more details.

[11] My original introduction to the idea of idealizing models came from the work of Peter Godfrey-Smith, to which the present work owes a deep debt. See his [Godfrey-Smith, 2004] for a similar attempt to treat folk psychology as an idealizing model. Yablo reads Quine as making a similar point, especially in ([Quine, 1960] ch 7), a point with which I am inclined to agree [Yablo, 1998].

we may read off the special sciences properties and objects that exist from explanatory theories.

Hence the stalemate. Fodor claims that since scientists get a lot of explanatory mileage out of MR kinds, MR predicates must pick out real causal properties. That is how explanation works: to explain just is to make reference to scientific types. Kim agrees that science always talks about explanatory kinds. Because there are good reasons to doubt that MR kinds fit the bill, Kim concludes that they must not have a place in science. The crucial move in both is an appeal to literalism; that's what makes possible the move from explanation to ontology.

In practice, however, we cannot determine the ontological commitments of a particular science just by looking at what terms show up in their best explanations. True, physics is plausibly committed to the existence of electrons, evolutionary biology to populations of rabbits, and engineering to steel plates—and what's more, they are so committed in part because these things show up in good explanations. On the other hand, these sciences certainly *aren't* committed to the existence of frictionless planes, infinite populations of rabbits, and perfectly thin, homogenous plates. Just ask the scientists who talk about them. But there are many quite good explanations that invoke frictionless planes and infinite populations and perfectly thin plates. That's a problem—or at least a problem for the literalism that led us to the stalemate.

When scientists talk about idealizations, they aren't talking about the actual world. They are talking about a greatly simplified version of the actual world, one that behaves quite differently than any actual things behave.[12] And this means that we can't determine the

---

[12] Many authors have taken the problems of idealization to be about how models can *represent* the world, despite the fact that they appear to deliberately falsify aspects of the very systems they discuss. For various reasons, I think that the representational view of models is unpromising. First, as Callender and Cohen point out, there is little reason to think that scientific models pose some *special* representational puzzle over and above general puzzles about any sort of representation [Callender and Cohen, 2006]. Second, on the representational view it is unclear what makes the difference between good models and poor theories. Newton failed—admirably, but failed nevertheless—when he tried to describe the dynamics of the actual world. The representational view must deny this, or else some serious pragmatic wiggling to get the right result. For if models are just slightly shabby representations of the world, then it's unclear how Newton went wrong. Finally, I show elsewhere that representational views must be either unrealistically simple or else give a an unreasonable and unmotivated semantics for scientific terms.

ontological commitments of a science from the terms that appear in its best explanations—for quite often, those terms explicitly refer to things that don't exist.

In short: the stalemate presupposes Literalism about scientific talk, and this literalism is wrong. Realizing this is the first step to dissolving the stalemate, and figuring out the role that putative MR kinds really play. With that in mind, let us turn to details.

## 4 Idealized Explanation

Consider again the humble property of brittleness. Discoveries about brittle things as such turned out to be unprojectable, and there was a good explanation from inheritance of heterogeneity as to why this should be expected.

Despite this, there is a legitimate family of explanations that use the term 'brittle' in an unrestricted sense. Fracture mechanics is the subfield of materials science devoted to the study of how things break. One of the first important breakthroughs in fracture mechanics was due to A.A. Griffith, who first discussed the notion of *critical crack length*. In any reasonably stiff material under tension, cracks below a certain length will be stable: barring further stress, they will not grow longer. Beyond the critical crack length, however, cracks become self-propagating. At this point, they will grow without increase in stress until they destroy the material completely. Griffith created a quite general model of crack propagation in brittle solids. This breakthrough came by idealizing away from the properties of real solids, studying instead the properties of the *ideally brittle solid*.

In the ideally brittle solid, cracks are always perfect notches perpendicular to the direction of stress. The material itself is assumed to be a perfectly thin, regular lattice of atoms extending to infinity in the direction of of the stress. The environment in which a material exists is assumed to be free of fluid pressure, radiation, excessive

---

Following Godfrey-Smith and Weisberg (and more distantly, Giere), I therefore treat idealizations as the non-actual referents of various representational practices in science [Godfrey-Smith, 2004, Weisberg, 2006, Giere, 1988]. That is, when scientists talk about perfectly homogeneous materials, the ideal gas, infinite populations of rabbits, and so on, they aren't trying to talk about the real world in a roundabout, confusing way. They're instead referring—directly, and well—about mere possibilia: perfectly homogenous materials, ideal gasses, and infinite populations of rabbits, respectively. Godfrey-Smith's and Weisberg's positions are explicitly pragmatic, however, while mine is explicitly not.

temperature, and so forth.[13] An ideal crack under tension is thus as pictured in figure one.

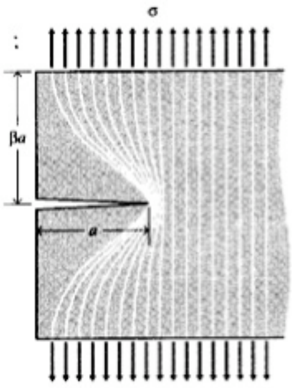

Figure 1: A crack in the ideally brittle solid

The triangular regions around the crack are assumed to be under zero tensile stress, as the crack prevents that portion of the material from carrying any load. The released strain energy from the unloaded areas is proportional to $a^2$ in the diagram, where $a$ represents the length of the crack. Additional stress is thus concentrated at the tip of the crack. This stress provides energy that can break atomic bonds and further extend the crack.

In addition to breaking energetic bonds, however, there is a further energetic requirement on crack propagation. Extending a crack requires the creation of more surface area around the crack face. Solids have a certain surface energy—think of it as like the potential energy stored in the surface tension of a liquid. Making a crack longer increases the surface area of the crack face, and thus requires some additional energy to compensate for this increased surface energy. If a crack is to propagate, therefore, the free energy released by the unloaded areas must be greater than the energy absorbed by the creation of new crack faces.

Griffith demonstrated that below a critical length, only further applied stress will tip the energy balance in the right direction; materials with shallow cracks simply don't have enough free energy to

---

[13] This is not comprehensive. See [Griffith, 1920] for Griffith's classic paper, wherein all of these conditions are laid out.

tip the thermodynamic balance in the direction of propagation. After the critical crack length, however, the energetic balance tips the other way. Strain energy released by growing cracks is greater than the energy required to make new crack faces; such a crack will propagate spontaneously, destroying the material. The critical crack length is given by the equation:

$$l_g = \frac{2\gamma E}{\pi s^2}$$

Where $l_g$ is the critical crack length, $\gamma$ the surface energy of the material, $E$ Young's modulus for the material, and $s$ the stress on the material.[14]

## 4.1 Brittleness as Idealization

The above explanation used a term 'brittle'. That term shows up in many explanations in fracture mechanics, in contexts that span a number of different realization-kinds. This does not vindicate Fodor, though. For the idealizing use of the term 'brittle' differs in important ways from the MR use of the term.

If the term brittle picked out an MR kind, remember, its referent would have a number of features. Each member of the kind would have to have a high level of similarity of causal powers. MR kinds figure in true (though usually *ceteris paribus*) laws. This means that if 'brittle' picked out an MR kind, we should expect the laws it figures in to give reasonably accurate descriptions of actual things, and to do so precisely because it's picking out a projectable scientific kind that each of these things belongs to.

That is not what happens when we talk about the ideally brittle solid.[15] For starters, we can be certain that nothing actual is 'brittle'

---

[14] The full derivation of this equation relies on Inglis' work on stress concentration. See [Griffith, 1920] for the full derivation, and ([Roylance, 1996] Ch7) for a simplified derivation.

[15] Depending on how we read 'reasonably accurate,' that is not what is going on with many scientific statements—as Nancy Cartwright famously argued in her [Cartwright, 1983]. The present position is indebted to Cartwright's argument, but differs from it in an important way. Cartwright argues that the laws of physics, while not accurate descriptions of the dynamics of actual things, do serve to pick out the causal properties of actual things ([Cartwright, 1983] 61). My position does not require even that. The Griffith model does not function by picking out causal powers of actual brittle solids; hence descriptions involving it cannot be read as descriptions of facts about actual entities.

in the way that the ideally brittle solid is brittle. Griffith's model is a severe simplification of the world. Many properties of the ideally brittle solid aren't—indeed, *couldn't*—be possessed by actual things. That is precisely the point. The world is very messy. Rather than try to work his way from the unimaginably complex details of particular materials to a true theory, Griffith began by leaving out a lot of important details.

It should not be surprising, then, that the behavior of Griffith's model often diverges severely from the behavior of the anything actual. Unlike MR kinds, we cannot use the the laws covering the Griffith model to predict or describe the behavior of most brittle materials. Relying on Griffith's work alone will give predictions that are wrong—and not wrong in the same way, but to differing degrees for different kinds of materials. For some brittle materials it will be quite close. It fares worse with ductile materials, far worse with many kinds of composites, and disastrously with fatigued metals. Serious corrections are also needed to deal with fracture in environments of high pressure, chemical agents, radiation, and so on.

MR kinds and idealizing models are very different beasts. Despite this deep difference, there are a number of features of *appeal* to idealization that are nearly identical to those usually cited as features of MR kinds.

First, the defining feature of MR kinds was supposed to be their ability to be instantiated in a diverse group of materials. This was precisely the feature that got them in trouble—spanning realization kinds meant that MR predicates could not pick out scientifically useful kinds.

Idealizations also span realization kinds, at least in one important sense. What is more, they do so in a metaphysically acceptable manner. The ideally brittle solid is an idealization of most realization-restricted brittle kinds. Idealizations are simplifications of their targets: they represent what the target *would* look like *if* important factors were removed. The ideally brittle solid idealizes steels in this sense, because it leaves out a lot of things (ductile flow around the crack tip, the effect of ferrous grain size) that play a causal role in fracture. It also idealizes glasses—not by leaving out the same set of things (for glasses and steels differ) but by leaving out those factors unique to glasses.

The unique flexibility of idealization derives from its ability to leave out important casual features of kinds. This is what allows idealization to respect Grounding—remember, the principle that states that the causal powers of high-level kinds depend on the powers of their low-level realizers—while still covering a wide variety of realization kinds. Idealizing models do not purport to describe the world; their coverage of different kinds is not in virtue of shared actual causal powers, and hence we have left the scope of the grounding thesis.

This difference in project explains another point of comparison. MR predicates, we are sometimes told, are important because they pick out *similarities* between high-level entities, similarities that would not be apparent from a lower level science. Unfortunately, as we saw, we have reason to expect that there won't be any non-trivial similarities of this sort if we start crossing realization-restricted kinds.

Idealizing models, on the other hand, do pick out interesting similarities between high-level entities. This is possible because idealizations do not describe how things do act, but rather how they *would* act if they were made simpler in various ways. Again, this similarity relationship does not conflict with any of the principles that proved so problematic for MR kinds.

A final virtue of MR kinds is supposed to be their role in explaining and justifying the actual practice of working scientists. In particular, if we treat scientists as using MR kinds, we can see why they in practice spend a lot of time ignoring the details of lower-level sciences. Now, as 2 shows, when scientists are sorting out their ontology and engaged in research, they must be especially sensitive to the breadth of their proposed kinds. Failing to respect lower-level kinds will result in scientifically useless kinds. Attention to idealization does not justify scientific practice in *this* sense, then.

In another important sense, though, the special sciences do end up relatively autonomous. Although scientists are metaphysically restricted in their ontology, the use of idealization gives them a good deal of methodological autonomy from the lower-level sciences. Idealizations can be used and investigated without much concern for the lower-level sciences. This in turn makes idealizations proper objects of study in their own right—witness Griffith's work, which spawned a whole subfield devoted to studying the properties of variations on the ideally brittle solid.[16]

In that regard, the present account is one of a number of recent

---

[16] Jeffry Ramsey and Clark Glymour have discussed something like the autonomy of idealizing models (though within different frameworks) in their [Ramsey, 1995] and [Glymour, 1970], respectively.

proposals that try to situate the autonomy of the special sciences in the distinctive modal import of their postulates. Marc Lange, for example, notes that many biological laws would hold even in physically impossible circumstances. As such, they give us information about a range of counterfactual possibilities that outstrips the physically possible ([Lange, 2004] 105). So even if biology can be reduced to physics in a classical sense, it cannot be eliminated in favor of physics. The modal breadth of biological laws outstrips anything that physics could give us. Similarly so for idealizations.[17] A good idealization tells us how a number of apparently disparate phenomena would behave if simplified; this information does not readily reduce to any information that physics provides.

## 4.2 Idealizing Explanation and its virtues

Before we get to the last set of comparisons with MR kinds, it's worth taking a step back. Idealization should look strange if you're used to thinking about MR kinds. Idealizing models are mere possibilia. Talk about them is false of anything in the actual world. You can't use them to predict anything. Why bother talking about them at all?

The main use of idealizing models—at least when we care about the properties of the actual world—is in explanation. In particular, idealizations are typically used to explain the ceteris paribus laws that cover the (realization-restricted) kinds of particular special sciences.[18] When scientists talk about the ideal gas, it is usually in the context of explaining the ceteris paribus laws that cover actual gasses. The ideally brittle solid is never cited on its own to explain anything. Instead, it forms the important *beginning* of a slew of good explanations in fracture mechanics.

Of course, these explanations will tend to be fairly complicated. Unlike explanations involving realization-restricted kinds, we cannot explain a law simply by showing how it is derivable from the laws governing the idealization. The laws governing idealized models don't hold in the actual world, after all. Idealizing explanation must be more complicated that this. The details are beyond the scope of this

---

[17] Indeed, Lange explicitly mentions idealizations in biology as an example of his thesis ([Lange, 2004] 105).

[18] As opposed to explaining token events. I take it that idealizations are of little use for explaining token events, since usually we want to know just how that event came about—what caused it, usually [Salmon, 1984].

paper, but I can make a few remarks about how it usually goes. To get to a law we want to explain, we often have to do a fair amount of work. In particular, we have to *progressively elaborate* idealized models, adding back in the complexity left out at the beginning.

This is what happens with the ideally brittle solid. Suppose we want to explain, say, the relationship between the size of the grains (the mid-size structural features in many metals) in steel and fracture toughness.[19] Griffith's model cannot account for fracture toughness in steel, because it assumes that the only energy requirement opposing propagation is the surface energy required to make new crack faces. Ductile metals flow at a micro-scale as well as a macro-scale, however. This micro-scale flow also soaks up energy, and so absorbs strain energy as well. The omission of ductile flow is (one of) the reasons why Griffith's model fails to give correct predictions for crack behavior in ductile metals.

Irwin and Orowan first elaborated a (still idealized) model that took into account ductile flow around the crack tip.[20] In this hypothesized region, applied stress can cause irreversible, non-linear effects (like ductile yield). Further mathematical advances allowed their assumption of a simple, small plastic zone around the crack tip to be relaxed.[21] Subsequent work re-introduced the effects from the formation of microcracks ahead of the crack tip—these are small cracks that propagate against the main line of fracture, and often increase the rate of crack formation [Krasowsky, 1998]. Finally work has shown how ferrite grain size plays an important role in the formation of microcracks.[22]

As it turns out, this relationship is an interesting and useful ceteris paribus law covering brittle steels. But that means our job is now done: we have managed to move from the highly idealized ideally brittle solid to a pretty useful, accurate law of fracture mechanics. By progressively adding back in features Griffith left out, scientists have

---

[19] For a technical discussion, I recommend [Lawn and Wilshaw, 1975].

[20] See [Irwin, 1957] for Irwin's original approach; for a generalized treatment see ([Lawn and Wilshaw, 1975] Ch4).

[21] Particularly important was Rice's development of the Rice Line Integral, which allowed him to elaborate the Irwin-Orowan treatment to an arbitrary, path-independent region around the crack tip ([Lawn and Wilshaw, 1975], p80-81). Rice's treatment makes possible several of the advances below. Note though that Rice's treatment was both made possible and *justified* by the simplified treatment of Irwin and Orowan.

[22] This relationship has been summarized by Tokobori in [Yokobori, 1998].

explained the specific behavior of a single realization-restricted kind.

Although idealizing explanation is more complicated than explanation via straightforward properties, it's worth pointing out that the *advantages* of idealizing explanation are also the advantages that accrued to explanation via MR kinds. First, idealizations allow explanations of laws ceteris paribus that avoid dealing with the massive complexity of special science systems all at once. Complications can be added back in one step at a time, greatly simplifying what would otherwise be an intractably complex task.

Second, the *same* idealization can form the basis for explaining a number of different kinds of system. The behavior of many different realization-kinds of materials can be explained by starting with the ideally brittle solid. One will have to add back in different complexities for different realization-restricted kinds, of course. But a good idealization simplifies this process considerably. As Gilman puts it:

> "One of the lessons that can be learned from the history of Griffith's paper on fracture is how exceedingly influential a good fundamental idea can be. Zener has pointed out reliable theories make powerful and efficient tools because all that is needed to use them is "paper and pencil". Langmuir called ideas like that of Griffith 'divergent' because they started from a small base and spread in depth and scope. Thus they differ from 'convergent' ideas which start from a broad basis of facts and converge to a very specific event or conclusion." ([Gilman, 1998], xxii)

Rather than having to start over afresh with each realization kind, scientists can start with a simple, fruitful, and well-understood idealized model. This is why idealization is so common, and why scientists spend so much time looking for useful idealizations.

Further, many have argued that the best explanations of laws are ones that *unify* a number of diverse phenomena. By showing that Kepler's laws of motion are special cases of Newtonian laws (that cover a number of other phenomena), for example, we reduce the number of independently acceptable phenomena and bring some sort of cognitive unity to science.[23]

Idealizations aid our understanding of a diverse group of phenomena in a similar way. Of course, the commonalities that idealizations

---

[23] See Friedman's [Friedman, 1974] and Kitcher's [Kitcher, 1989] for canonical discussions, as well as [Newell, 1990] for an appeal for unification in psychology.

point to may be far removed from implementation details—they are similarities, again, that things *would* have if they were suitably simplified. Still, by unifying phenomena in this way, scientists gain explanatory advantage that they would lack if they focused on the details of specific systems.

Again, we see that idealized models have all of the features usually attributed to MR kinds. Idealizations simplify the world. They allow us to see rough commonalities between a number of different realization kinds. Indeed, they are applicable to a wide range of realization kinds, though variably so in each case. They are used often by the special sciences. These are just the advantages that proponents of MR kinds cite as so indispensable. The only difference is that idealizations can actually give us these advantages, while MR kinds can not.

# 5 Idealization and Ontology

Idealization gives us a way out of the impasse identified in section 3. Folks like Kim are right to say that there are no MR kinds. They are wrong to assume that all work in science must involve reference to realization-restricted laws and properties. Idealizations range over a number of different kinds, and they do so without any ontological tomfoolery. Indeed, it is in part this wide scope that makes idealizing a useful scientific strategy.

Similarly, folks like Fodor are right to say that many special science kinds look MR and are quite useful. They are wrong to think, however, that that for this reason we must think of them as legitimate kinds that figure in laws of nature. There are good reasons to think that we shouldn't. More importantly, we can do all of the scientific work with them that we want without treating them as such. Attention to idealization shows how we explain well without talking much about actual kinds.

Ultimately, what I suggest is a certain metaphysical strictness combined with a methodological liberality. If we are to be serious about causal powers and kinds, we must be serious about the constraints physicalism puts on the relationship between high-level kinds and their lower-level realizers. As a consequence, we must *also* take seriously the plurality of representational strategies available to scientists. Scientists need not always cite causal powers and laws when they are explaining the world. We cannot simply read the causal structure of the world just by looking at what nouns scientists use.

A final warning is therefore in order. Given the preceding, it is reasonable to assume that most special science terms are in fact ambiguous. Sometimes they refer to realization-restricted kinds. This will most often be the case when they figure in laws ceteris paribus that are used to explain token events. Other times, they will refer only to elements in an idealized model. This will often be the case when we are explaining the laws and processes characteristic to a special science.

That is quite odd, and should be marked: many good uses of mental terms don't pick out any actual entities. They are explanatory nevertheless. The consequences of this for traditional philosophy of mind are wide-ranging. The first step in figuring out the consequences, however, will be careful attention to the representational practices of working scientists. Only then can we sort out the ontological commitments of the special sciences.[24]

# References

[Callender and Cohen, 2006] Callender, C. and Cohen, J. (2006). There is no special problem about scientific representation. *Theoria*, 55:7–25.

[Cartwright, 1983] Cartwright, N. (1983). Do the laws of physics state the facts? In *How the Laws of Physics Lie*, pages 54–73. Oxford University Press.

[Cherepanov, 1998] Cherepanov, G. P., editor (1998). *Fracture: A Topical Encyclopedia of Current Knowledge.* Krieger Publishing Company.

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

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

[Fodor, 1997] Fodor, J. (1997). Special sciences: Still autonomous after all these years. *Philosophical perspectives: Mind, Causation, and World*, 11:149–163.

[Friedman, 1974] Friedman, M. (1974). Explanation and scientific understanding. *Journal of Philosophy*, 71(1):5–19.

---

[24] **Acknowledgments suppressed for blind review.**

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

[Gillett, 2003] Gillett, C. (2003). The metaphysics of realization, multiple realizability, and the special sciences. *The Journal of Philosophy*, 100(11):591–603.

[Gilman, 1998] Gilman, J. J. (1998). Alan Arnold Griffith (1893–1963): An appreciation. In [Cherepanov, 1998], pages xix–xxii.

[Glymour, 1970] Glymour, C. (1970). On some patterns of reduction. *Philosophy of Science*, 37(3):340–33.

[Godfrey-Smith, 2004] Godfrey-Smith, P. (2004). On folk psychology and mental representation. In *Representation in Mind: New Approaches to Mental Representation*, chapter 8. Elsevier Ltd.

[Gordon, 1984] Gordon, J. E. (1984). *The New Science of Strong Materials.* Viking Penguin, 2nd edition.

[Griffith, 1920] Griffith, A. A. (1920). The phenomena of rupture and flow in solids. In [Cherepanov, 1998], pages 2–29.

[Irwin, 1957] Irwin, G. (1957). Analysis of stresses and strains near the end of a crack traversing a plate. In [Cherepanov, 1998], pages 30–40.

[Jaworski, 2002] Jaworski, W. (2002). Multiple-realizability, explanation, and the disjunctive move. *Philosophical Studies*, 108:289–308.

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

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

[Kitcher, 1989] Kitcher, P. (1989). Explanation, unification, and the causal structure of the world. In [Kitcher and Salmon, 1989], pages 410–505.

[Kitcher and Salmon, 1989] Kitcher, P. and Salmon, W., editors (1989). *Scientific Explanation.* University of Minnesota Press.

[Krasowsky, 1998] Krasowsky, A. J. (1998). Local approach to fracture of structural materials. In [Cherepanov, 1998], pages 316–330.

[Lange, 2004] Lange, M. (2004). The autonomy of functional biology: a reply to rosenberg. *Biology and Philosophy*, 19:93–109.

[Lawn and Wilshaw, 1975] Lawn, B. R. and Wilshaw, T. R. (1975). *Fracture of Brittle Solids.* Cambridge University Press.

[Millikan, 1999] Millikan, R. (1999). Historical kinds and the ‘special sciences’. *Philosophical Studies*, 95:45–65.

[Newell, 1990] Newell, A. (1990). *Unified Theories of Cognition*. Harvard University Press.

[Oppenheim and Putnam, 1958] Oppenheim, P. and Putnam, H. (1958). Unity of science as a working hypothesis. *Minnesota Studies in the Philosophy of Science*, 2:3–36.

[Quine, 1960] Quine, W. V. O. (1960). *Word and Object*. MIT Press.

[Ramsey, 1995] Ramsey, J. L. (1995). Construction by reduction. *Philosophy of Science*, 62(1):1–20.

[Rosenberg, 2001] Rosenberg, A. (2001). Comments and criticism on multiple realization and the special sciences. *The Journal of Philosophy*, 98(7):365–373.

[Roylance, 1996] Roylance, D. (1996). *Mechanics of Materials*. John Wiley and Sons.

[Salmon, 1984] Salmon, W. (1984). *Scientific Explanation and the Causal Structure of the World*. Princeton University Press.

[Shapiro, 2000] Shapiro, L. (2000). Multiple realizations. *The Journal of Philosophy*, 97(12):635–654.

[Shoemaker, 2001] Shoemaker, S. (2001). Realization and mental causation. In Gillett, C. and Loewer, B., editors, *Physicalism and its Discontents*, pages 74–98. Cambridge University Press.

[Weisberg, 2006] Weisberg, M. (2006). Who is a modeler? *British Journal for Philosophy of Science (forthcoming)*.

[Yablo, 1992] Yablo, S. (1992). Mental causation. *The Philosophical Review*, 101(2):245–280.

[Yablo, 1998] Yablo, S. (1998). Does ontology rest on a mistake? *Proceedings of the Aristotelean Society*, 72:229–61.

[Yokobori, 1998] Yokobori, T. (1998). The unified philosophies in fracture. In [Cherepanov, 1998], pages 698–708.