Cognitive Science 28 (2004) 409–432

COGNITIVE SCIENCE

http://www.elsevier.com/locate/cogsci

Special Sciences: still a flawed argument after all these years

Todd Edwin Jones*

University of Nevada, 1404 S. Sixth St., Las Vegas, NV 89104, USA

Received 15 February 2002; received in revised form 26 June 2003; accepted 7 October 2003

Available online 5 March 2004

Abstract

Jerry Fodor has argued that the multiple realizability argument, as discussed in his original "Special Sciences" article, "refutes psychophysical reductionism once and for all." I argue that his argument in "Special Sciences" does no such thing. Furthermore, if one endorses the physicalism that most supporters of the "Special Sciences" view endorse, special science laws *must* be reducible, in principle. The compatibility of MR with reduction, however, need not threaten the autonomy of the special sciences.

Keywords: Reductionism; Special Sciences; Mental states; Physical states; Mind-body problem; Fodor

1. Introduction

Cognitive scientists often take for granted the idea that neurobiology or physics, by themselves, could not possibly provide a sufficient explanation for why we think the way we do. They believe they are on quite safe grounds in assuming this, since this has been the consensus view among cognitive scientists since at least the mid-seventies. One of the factors responsible for making this anti-reductionist stance the consensus view has been a much reprinted and widely cited article by the philosopher Jerry Fodor, entitled "Special Sciences, or The Disunity of Science as a Working Hypothesis." In "Special Sciences," Fodor argues that the longstanding dream of a single all-encompassing unified science that could be used to explain everything could never come to pass. Certain sciences like psychology, he argues, are

* Tel.: +1-702-895-4691; fax: +1-702-895-1279.
E-mail address: tjones@unlv.nevada.edu (T.E. Jones).

0364-0213/$ – see front matter
doi:10.1016/j.cogsci.2003.10.004

"autonomous" in that they investigate realms governed by unique special laws. Psychological laws, he argues, could be realized radically differently in numerous physical systems, in ways that make it impossible to derive psychological generalization from the general laws of physics. This is true, he argues, even if minds are ultimately made of merely physical substances. The view Fodor articulates in "Special Sciences," which enables people to endorse both physicalism and the autonomy of special sciences, has been very convincing to many. In a recent article entitled, "Special Sciences: Still Autonomous After All These Years," Fodor re-endorses the ideas of his previous article. He quotes Jaegwon Kim as saying, "the conventional wisdom in philosophy of mind [is] that psychological states are 'multiply realized' [and this fact] refutes psychophysical reductionism once and for all" and cheekily responds: "Despite the consensus, however, I am strongly inclined to think that psychological states are multiply realized and this fact refutes psychophysical reductionism once and for all" (1998, p. 9).

I will argue here that nothing in "Special Sciences" really gives us any reason to believe multiple realizability refutes psychophysical reductionism once and for all. Indeed, in contrast to the standard view, I argue the physicalism most advocates of the "Special Sciences" view endorse actually contains elements insuring that special science laws *must always* be reducible, in principle. "Special Sciences" does give us a vivid illustration of how reduction to physics will likely be a difficult daunting task. It does not, however, provide us with a good argument against the possibility of reduction. The special sciences may well be autonomous, in some sense, but this can not be due to their irreducibility.

2. Fodor's central illustration: Gresham's law

One of the striking things about "Special Sciences" is that, despite it being one of the classic sources of the anti-reductionist consensus, it is not easy to find its main argument against reduction. As far as I can tell, the central anti-reductionist argument is given only in a single long sentence towards the middle of the article:

> The reason it is unlikely that every kind corresponds to a physical kind is just that (a) interesting generalizations (e.g., counterfactual supporting generalizations) can often be made about events whose physical descriptions have nothing in common; (b) it is often the case that whether the physical descriptions of the events subsumed by such generalizations have anything in common is, in an obvious sense, entirely irrelevant to the truth of the generalizations, or to their interestingness, or to their degree of confirmation, or, indeed to any of their epistemologically important properties; and (c) the special sciences are very much in the business of supporting generalizations of this kind. (Fodor, 1993, p. 433)

The reason that Fodor doesn't try to make this anti-reductionist point with a long detailed argument becomes clear in the next sentence: "I take it that these remarks are obvious to the point of self-certification; they leap to the eye as soon as one makes the (apparently radical) move of taking the existence of the special sciences at all seriously." Fodor appears to think that it is *so* obvious that the special sciences have laws with predicates that are irreducible to certain types of physical realizers that he barely bothers to give any illustrations of these.

Fodor's main illustration of his central point is described in only the barest of outlines in his next several sentences:

> Suppose, for example that Gresham's 'law' really is true ... Gresham's law says something about what will happen in monetary exchanges under certain conditions .... But banal considerations suggest that a physical description that covers all such events must be wildly disjunctive. Some monetary exchanges involve strings of wampum. Some involve dollar bills. And some involve signing one's name to a check. What are the chances that a disjunction of physical predicates which covers all these events (i.e., a disjunctive predicate for the right hand side of a bridge law of the form 'x is a monetary exchange ↔ ... ') expresses a physical kind?

In relying on a bare bones example to make his point, Fodor is following in the footsteps of his mentor, Hilary Putnam. In the sixties, Putnam argued that one could easily see that certain laws were irreducible when one considered the following: a rigid square peg that is a certain length across cannot be put through a rigid round hole that is the same length across (Putnam, 1975a, p. 295). Fodor apparently thinks he does not have to describe what Gresham's law is or what makes it irreducible in any detail because it is just intuitively obvious that no physicalistic reduction of "monetary exchange" could ever be forthcoming. Like many of "Special Sciences" readers, I admit that I too feel the grip of an intuitive picture in which the property of being money—which can take every form from cowery shells to magnetic traces in computer accounts—could never be describable using rigid narrow terms of physics. Still, given the centrality of reduction as a scientific strategy, and given that the main purpose of the article is to show that multiple realizability refutes reduction, it would have been nice to augment an example where reduction is hard to conceive with a more step-by-step discussion of *why* there could be no reduction here. After all, someone might easily say, "well, I don't see how to do it, but there might well be some really smart fellow out there who can." Fortunately, I think it is fairly clear how a more step-by-step discussion of this point could go. Gresham's law states that 'bad money will drive out good money.' Money perceived as a storer of value or commanding a better foreign exchange rate (say gold) will tend to drop out of circulation in favor of money perceived as being usable only in domestic exchanges (say silver). Economic states of affairs meeting this description can be realized by innumerable different physical systems. Gresham's law is presumably true whether the money involved is coinage or computer traces, and true whatever the coins are made of, or whether they are large or small. It is this *multiplicity* that seems to make things problematic for physicalistic reduction. For every physical molecular account one could give detailing why, in this or that set of instances, *these* monetary exchanges helped dwindle *that* supply of a money type, there will be other sorts of exchanges of other types of money that led to a shrinking supply of still different types of money through different types of exchanges. This means that no physical account of how good money drives out bad will ever tell you why this is true of *every* case; no physical account will tell you why, or even that this *law* is true. There are therefore some true economic generalizations that cannot be explained by physics. This simple case is presumably analogous to numerous cases in other social sciences, as well as in psychology, or even biology. When we have a situation in which there can be innumerable realizers of a property in a true generalization, then the reduction of that generalization will be problematic because no reductive description will describe why *every* instance of that generalization holds true.

3. Three straightforward objections

In "Special Sciences," Fodor doesn't appear to believe that he needs to do much to flesh out the picture of how multiple realizability is supposed to refute reduction—that it did so was, for him, "obvious to the point of self-certification." But when we flesh out the discussion of why we should expect no reduction for multiply realizable laws, as we did above, several objections immediately suggest themselves. What our argument sketch shows is that, in many cases, no *particular* microphysical explanation can tell you that a certain macro-level generalization holds. But that doesn't tell you that there *can be no* microphysical explanation of the generalization. There are two ways a defender of reduction might try to explain the higher level generalization. The first is to say that while no particular microphysical explanation tells you why all instances of that macro-generalization are as they are, an exhaustive *disjunction* of the microphysical explanations of each of the cases of this generalization could. A second is to claim that we can explain why the different realizations of this generalization all hold true using abstract general terminology which can itself be ultimately defined with physical terminology.[1] If disjunctive or abstract physical accounts are possible, then the fact that each particular realization is somewhat different doesn't do anything to establish the impossibility of reduction. One might also object that the identity conditions for something's being a "monetary exchange" have little to do with coinage and much to do with the attitudes of the participants. So the fact then that money is multiply realized does little to indicate that the psychological states that provide the real underpinnings of Gresham's law are not reducible to neurology and ultimately physics.

Any one of these objections, if sustained, is sufficient to show that pointing to Gresham's law and other examples do not really "refute psychophysical reduction once and for all." In "Special Sciences" and elsewhere, Fodor shows some awareness of these objections and makes some efforts to overcome them. I believe that these efforts are unsuccessful. I will discuss each objection and Fodor's attempts to block them, below. It should become clear that, despite the conventional wisdom that "Special Sciences" helped bring about, multiple realizability and reduction are, in fact, compatible.

4. Disjunctive realizations and reductive explanations

One way to try to undermine claim that multiple realizability makes reduction impossible is to argue that, even if no *particular* microphysical explanation can tell you why all instances of a macro-generalization are as they are, an exhaustive *disjunction* of the microphysical explanations of each type of case of this generalization could. In an early paper espousing the multiple realizability argument, Hilary Putnam noticed that the MR argument might indeed be circumvented this way. But Putnam, *with no argument whatsoever*, dismissed such a disjunctive account, as "not a metaphysical option that can be taken seriously" (Putnam, 1975b, p. 437). In "Special Sciences," Fodor puts a little more effort into blocking the possibility of reduction through the use of disjunctive accounts. Take a law of the special sciences of the general form $S_1x \rightarrow S_2y$. Suppose someone could show that a higher level type, S_1x, is realized by the disjunction of lower level types, $P_1x \; V \; P_2x, \; V \ldots \; P_nx$. (In the following section we will consider *whether* such a disjunction can be given; in this section we are seeing what follows from assuming it *can* be.) Suppose that S_2y is realized by $P_1^*y\ V\ P_2^*y\ V \ldots\ P_n^*y$. Suppose that the laws of physics showed that for each P_nx, you would get a P_n^*y. You still, in Fodor's view, would not have a reduction to physics of the law $S_1x \to S_2y$. Why not? Because the most you could show with physics was that the generalizations $P_1x\ V\ P_2x,\ V \ldots\ P_nx \to P_1^*y\ V\ P_2^*y\ V \ldots\ P_n^*y$ was true. But this generalization is not a *law*, so it doesn't explain why $S_1x \to S_2y$ is lawful. Why isn't it a law? Because, Fodor argues, $P_1x\ V\ P_2x,\ V \ldots\ P_nx$ is not a *natural kind* term. It's just a list. To specify why it's not a natural kind term, Fodor argues that terms like that are not utilized in scientific laws. Now using lawfulness to specify what is or isn't a kind, and kindhood to specify what is and isn't a law puts Fodor in a circle. But he thinks that kindhood and lawfulness are intertwined such that he can't see any way out of the circle. And we can use this intertwined circle to see both that special science laws and special science kinds are irreducible. ("I don't know how to break out of this circle, but I think there are some interesting things to be said about the circle we are in" (p. 432)).

Now given that both the notions of law and of kind have remained frustratingly vague despite years of philosophical exploration, it's somewhat surprising that Fodor gives these notions so much weight in his argument that special science kinds can't reduce to disjunctive kinds. Worse for Fodor, however, is that stressing their intertwined circular definition enables critics of his argument simply to run the circle in the other direction: to argue that there are special laws about disjunctive kinds. In a pair of recent articles, "Multiple Realizability, Projectibility, and the Reality of Mental Properties," and "Who's Afraid of Disjunctive Properties, '' Louise Antony defends the idea that there can be true laws about disjunctive kinds, and that disjunctive kinds are real kinds. In these articles, Antony is mainly concerned to argue against Kim who believes, like Fodor, that a disjunction of heterogeneous properties couldn't really have any laws true of it, or couldn't be described by a projectible predicate. Antony argues that there might be plenty of true laws about disjunctions like "cows or bulls." There could be plenty of predicates describing disjunctive properties that are quite real kinds because they describe a unified set of things that has a "distinctive causal repertoire" (see also Clapp, 2001). She argues that "we ought, after all, to identify [high level] MR properties with their associated [disjunctive heterogeneous] Kim properties" (2003). She goes as far as to say that even Gresham's law whose irreducibility Fodor views as obvious *could* be identified with its disjunctive physical realizers: "Fodor's predicates—like the one constructed by disjoining each and every physicalistic description that applies to an event that is a monetary exchange—are ones that are necessarily coextensive with *projectible* predicates—in this example, with the predicate "monetary exchange" (2003).

Now Antony, it should be noted, does not believe that this identification of special science predicates with a disjunctive set of physical realizers counts as a reduction of special science predicates to physical ones. But she denies this is a reduction only by employing what I consider to be a somewhat idiosyncratic construal of "reduction." Antony writes "we need to think a little differently about the notion of "reduction" ... the question of reducibility should be construed as partially a semantic and partially epistemological question, rather than an ontological question" (Antony, 1999, p. 18). But in post-positivist (and post "linguistic turn") philosophy, the question of reduction has been widely regarded as a paradigmatically ontological question. Deniers of mind-brain identity theory were concerned to argue that what a particular mental state *is*, ontologically, was a kind of functional state, rather than a kind of brain state. It had to be functional because of *the way the world is*, not because of any epistemic problems of discovery, or the way humans use words. Antony argues that there are severe epistemic problems in trying to uncover disjunctive classes that realize high-level properties. I certainly agree. But, if we give "reductive identification" its widely held ontological reading, practical epistemic problems are irrelevant. If there is, *in principle*, a way to use physical terminology to describe what each of the realizers of a special science predicate can and cannot do, then that kind is reducible to physics, on this ontological construal of reduction.

I believe, however, that there is a much more damaging objection to Fodor's argument against reducing high-level properties to a disjunction of lower level ones. While the notion of what does and doesn't count as a law is notoriously vague, let's concede to Fodor, for the sake of argument, that, contra Antony, a generalization involving disjunctive properties does not count as a law. Let us also concede the disjunctive property the generalization describes is not any sort of "natural kind." Why should this matter at all to the question of reductive explanation? Recall that we are assuming, with Fodor, that the high-level predicates S_1x, and S_2y are coextensive with sets of particular disjunctive realizers, ($P_1x\ V\ P_2x,\ V \ldots\ P_nx$) and ($P_1^*y\ V\ P_2^*y\ V \ldots\ P_n^*y$). We are also assuming that we can use the laws of physics to show why it has to be that $P_1x\ V\ P_2x,\ V \ldots\ P_nx \rightarrow P_1^*y\ V\ P_2^*y\ V \ldots\ P_n^*y$. Why should it matter, then, whether $P_1x\ V\ P_2x,\ V \ldots\ P_nx \rightarrow P_1^*y\ V\ P_2^*y\ V \ldots\ P_n^*y$ really counts as a law? We would have shown how physics makes it the case that $S_1x \rightarrow S_2y$ is a true generalization. If we know why it's a true generalization, haven't we explained this generalization whether or not we count it as a law? Correspondingly, why should we care whether or not $P_1x\ V\ P_2x,\ V \ldots\ P_nx$ counts as "natural kind" (whatever that means)? Perhaps the generalization $S_1x \rightarrow S_2y$ really isn't a law, but is just something that happens to be true, given the particular physical facts we happen to have here. A lack of lawfulness wouldn't make the generalization any less true, or any less able to be explained by physics. To take a concrete example, lets assume that it is a true generalization that that every metal spoon now existing in the universe will bend when a combined pressure of over 50 kg is applied to each end. Such a generalization is multiply realizable since the type "metal spoon" can be realized by objects with very different physical make-ups. Such a generalization would probably not be intuitively considered a law. But, in principle, we could certainly use the laws of physics to explain why this generalization holds true. We would give a physical description of each existing metal spoon, and show why their construction is such that applying 50 kg of pressure would cause each to bend. If Fodor is bothered by using honorific descriptions "law" and "natural kind" for disjunctive properties and generalizations like $P_1x\ V\ P_2x,\ V \ldots\ P_nx \rightarrow P_1^*y\ V\ P_2^*y\ V \ldots\ P_n^*y$, why can't we just refrain from calling them that, but maintain we've succeeded in giving a reductive physical explanation of the relation between the predicates in an important special science generalization?

It appears, then, that the use of disjunctive realizers (lawlike or not) does not, by itself, mean that you can not use physical predicates and laws to show why special science generalizations are true. Still, there are a few ways in which Fodor and others might try to deny that showing this constitutes a reductive explanation. They might try to deny that showing how to derive special science generalizations from physics laws is really an explanation. Or they might say that even if it is an explanation, such explanation doesn't really constitute a reduction. Or they might say that even if it does constitute a reduction, it is not a reduction to physics. I believe, however, that none of these denials would be plausible.

We have been assuming in this section, along with Fodor, that one *could* use the laws of physics to derive the generalization $P_1x \vee P_2x, \vee \ldots P_nx \rightarrow P_1^*y \vee P_2^*y \vee \ldots P_n^*y$. When we know this, we know why it is the case that every S_1x leads to S_2y—since the $P_1x \vee P_2x, \vee \ldots P_nx$ disjunction *picks out each and every case of* S_1x and the $P_1^*y \vee P_2^*y \vee \ldots P_n^*y$ disjunction picks out every S_2y. This is certainly intuitively an *explanation* of why every S_1x leads to S_2y (for a concurring view, see Sober, 1999). And most models of explanation used in the philosophy of science routinely call this sort of account explanatory. On the standard Hempelian deductive-nomological theory of theoretical explanation, sets of lower level laws and bridge principles are combined to deductively derive (or show the high probability of) higher level generalizations. Physics laws about molecular motion and bridge principles linking molecular motion to gas temperature and pressure, for example, combine, on Hempel's theory, to show us why the higher level Graham's law concerning gas diffusion is true (see Hempel, 1966). Nothing in Hempel's notion of explanation says that you can't use a bridge law to link a kind term to a set of disjunctions. Deriving $S_1x \rightarrow S_2y$, from the facts that $P_1x \rightarrow P_1^*y$, and $P_2x \rightarrow P_2^*y \ldots$ and the bridge principle that $S_1x \leftrightarrow P_1x \vee P_2x, \vee \ldots P_nx$, then, fits the standard Hempelian model of theoretical explanation just fine. It fits with more contemporary theories of explanation as well. On Kitcher's unification theory of explanation (1981, 1989), a derivation is explanatory if the fact being derived can be derived from the smallest set of patterns of derivation that can derive the largest set of facts about the universe. If, as Fodor and other physicalists suppose, our world is composed entirely of fundamental particles moving in arrangements dictated by a small number of forces, then the smallest, most unifying sets of derivations would be ones that described facts about the world using low-level physics laws. But that is exactly what is done when the $S_1x \rightarrow S_2y$ generalization is derived from purely physical laws like $P_1x \rightarrow P_1^*y$, and $P_2x \rightarrow P_2^*y$. The derivation Fodor describes, then, passes the unification theory's test of explanatoriness as well.[2] Other theories of explanation hold that we explain F when we show what *causes F* to be true. The *F* in question here is the generalization $S_1x \rightarrow S_2y$. Let's assume that this is a causal generalization. Could the derivation Fodor describes constitute a causal explanation? Certainly it could. Why does S_1x cause S_2y? Presumably Fodor, like other physicalists believes that the higher level state S_1x has the causal powers that it does by inheriting the causal powers of its various realizers (see Kim, 1998). What causes S_1x to bring about S_2y, in other words, is that S_1x can be realized by P_1x or P_2x, etc., and that P_1x causes P_1^*y while P_2x realizations cause P_2^*y, etc. The existence of these causal laws tells us why, in every situation in which we have an S_1x, an S_2y state of one or another sort is caused to result. We see how these physical causal laws cause it to be the case that $S_1x \rightarrow S_2y$. That, on the causal theory, is an explanation of $S_1x \rightarrow S_2y$. On various models of explanation, and intuitively, the $Px \rightarrow P^*y$ laws do, indeed, provide us with a physical explanation of why $S_1x \rightarrow S_2y$.

But does such a physical explanation count as *a reduction* of $S_1x \rightarrow S_2y$? Certainly it does on the classic Nagel model reduction. Kim has pointed out that "the Nagel model of reduction is in effect the Hempelian D-N model of scientific explanation applied to intertheoretic contexts ... Nagelian reduction is accomplished in the derivation of the target theory from the base theory taken in conjunction with bridge laws as auxiliary premises" (1998, p. 25). As we've just described, Fodor's disjunctive scenario is just such a Hempelian explanation of a higher level generalization using lower level laws and bridge principles connecting them—a classic

Nagelian reduction. Now some have complained that Nagel model of reduction is too liberal, in allowing derivations using conditional bridge principles to count as reductions, rather than demanding full identity relations between the reducing and the reduced domain (Causey, 1977; Kim, 1998). I agree that the classic notion uses a somewhat liberal notion of reduction. But if we are willing to admit disjunctive laws as laws, or if the generalization's lawfulness is irrelevant as I have argued above, then there are no obstacles to talking about $S_1x \rightarrow S_2y$ reducing to $P_1x\ V P_2x,\ V \ldots\ P_nx \rightarrow P_1^*y\ V\ P_2^*y\ V \ldots\ P_n^*y$ according to even the most stringent full *identity* standard of reduction. Showing how the laws of physics make it that case that every instance of the predicate S_1x leads to S_2y also captures the central idea of what reduction is all about: ontological simplicity. The central intuitive idea of reduction is that we can take a smaller set of entities and laws (like those of physics) and use these to show why a much larger set of entities (such as those of chemistry, biology, or psychology) behave as they do. If we can use physics to show why the S_1xs (say from psychology) must always lead to S_2ys, then we have intuitively reduced this generalization to a combination of physical laws. We will have shown how what looked like a sui generis autonomous generalization was really the net effect of a set of more fundamental laws. This is exactly what reductionism was always meant to do.[3]

But if such a reduction could be accomplished, would it really be a reduction to physics? I see no reason it shouldn't be characterized as such. True, there may be no basic physics laws about the causes and effects of a given disjunctive list of properties; and perhaps no basic physical predicates could describe what such a disjunctive list shares. But that doesn't mean we can't take conjunctions or disjunctions of physical properties, and describe these *collections* of physical types with second and third order short-cut predicates. (After all there are no strictly physical laws about the sugar-type "monossaccharide"; but that doesn't mean we can't reduce monossaccharide to $C_6H_{12}O_6$ (glucose) or fructose or galactose, etc.) Such high-level predicates would just pick out various states of the world that we could, in principle, describe with long conjunctions or disjunctions of physical predicates. In the other direction, when we say that S_1x can be realized by physical states $P_1x\ V P_2x,\ V \ldots\ P_nx$, we are taking a non-physical predicate and describing what it picks out using the predicates of physics. S_1x, it could be said, just *is* this long list of different physical states. (Nothing more and nothing less if we have really given the complete list of disjuncts.) Once we have that list, we could theoretically ignore that we ever call this collection by its macro-level S_1x name. If we use physical laws to show why $P_1x\ V P_2x,\ V \ldots\ P_nx \rightarrow P_1^*y\ V\ P_2^*y\ V \ldots\ P_n^*y$, we are using physics alone to show why, when we have this or that physical state of affairs, we will invariably get this or that resultant state of physical affairs. We are giving a purely physical account of why this disjunctive set of physical states leads to that disjunctive sets of physical states. If we can say what the disjunctive list of physical brain states $P_1x\ V P_2x,\ V \ldots\ P_nx$, must do in certain circumstances, according to the laws of physics then why not call this a reductive physical explanation of what they do?[4]

Putnam and Fodor have both argued that even if we could identify all the disjunctive realizers of the entities in a higher level generalization, that still wouldn't count as reductively explaining with physics. We've seen, however, that there are no good reasons to deny that this would be an explanation, a reduction, or a purely physical depiction. Contrary to what was claimed in *Special Sciences*, if we can, in principle, know the disjunction of all a high-level predicate's physical realizers, and if we know the physics laws that connect predicates' realizers to insure a high-level generalization will be true, we can reduce that special science generalization to physics.

5. Infinite realizers

There is a different way, however, that Fodor could try to make the case that many special science generalizations are irreducible. We can only really reduce a predicate in the special science to predicates in physics if we can describe the class the special science predicate names with a physics-predicates-only description. But this might be impossible if the kind the special science predicate describes has an *infinite* number of physical realizers. If there are an infinite number of physical states that could be money, to use Fodor's example, no finite disjunctive description of various physical states could captures them all. The difficulty infinite realizers pose for reduction, I believe, is what Fodor has in mind when he stresses that true laws support counterfactuals: a law regarding a special science kind must apply not only to all actual disjunctive realizers of that kind, but all the realizers there could possibly be (1996, p. 433). Perhaps for some special science types, no disjunctive list could name all the possible physical realizers because there is no possibility of describing a list that is infinitely long. Perhaps the only way we can describe not only all the actual, but all of the infinite number of *possible* physical instances of the law's realizers is to use some more abstract non-physical characterizations. This is why, perhaps, we're better off referring to certain states as, say, "pain" rather than attempt the impossible task of enumerating an infinite number of physical brain states.

The need to use more abstract descriptions to describe a potentially infinite number of realizers, however, does not automatically show a reduction is impossible. The reductionist can counter that there is (or may be) abstract descriptions that are (a) purely physical, while having an infinite number of different possible types of realizers (terms like "force" and "particle" come to mind) or (b) defined in terms of predicates that are defined in terms of predicates, etc., that are ultimately defined in terms of purely physical descriptions. Let's call the class consisting of either of these descriptions, 'abstract physical/physically reducible.' Now consider the familiar generalization about water turning to steam at sea level at 100 °C. Explaining why this generalization is true usually involves a number of abstract predicates (e.g., 'volume of water,' 'mean molecular kinetic energy'). Many of these states described by these predicates are realizable in an infinite number of ways. This generalization is true (ceteris paribus) regardless of size of the volume, meaning that the generalization holds for an infinite number of different types of physical realizations of "a volume of water." Indeed, even for a *single* given volume of water, at the molecular level, there are an infinite number of types of ways in which the necessary mean molecular kinetic energy can be reached. The requisite mean molecular kinetic energy can be reached by every molecule moving at a certain speed, or with half that number moving at twice that speed. There are innumerable types of molecular states the volume could be in, but as long as it reaches the requisite mean molecular kinetic energy, the water is boiling and turning to steam. There is no problem, however, with giving a general explanation, using physics alone, of why all of the *infinite* sorts of physical realizers of "a volume of water" behave as they do when heated.[5] A class's having an infinite number of members does not mean that we can't give a finite description of what can make-up the class. A class's having an infinite number of physical types of members doesn't mean we can't have a finite description in physics terminology of what those class members can be. Throughout physics we have numerous cases where general abstract physical characterizations explain how an infinite number of type of physical arrangements can make a higher level generalization true (see Batterman, 2000). The existence of multiple realizability, then, needn't mean that it's unlikely that a system could also have a general, purely physical characterization.[6]

Why should we think, then, that, for a given multiply realizable special science law, there *could never be* an abstract physical/physically reducible characterization? Fodor gives us assertions and rhetorical questions, here: "What are the chances that a disjunction of physical predicates which covers all these events ... expresses a physical kind?" (p. 433). But he provides no real arguments. This is unsurprising. As I suggested before, Fodor is relying on the fact that it is very hard to picture any physical way of describing all possible cases of monetary exchange. But as I also suggested, it would be nice to have more to rely on here than an intuitive difficulty to conceive. Shapiro (2000) has recently pointed out, that what might seem intuitively to have an unruly irreducible multitude of realizers, might, upon inspection to have actually only a few physically describable realizers. Bechtel and Mundale (1999) have discussed how this might actually be the case regarding mental states that seem to have no shared physiological characteristics at first glance.

At this point, we have a stalemate. Faced with a seemingly irreducible special science generalization, a determined reductionist might reason, "The success of reduction in other areas of science, leads me to expect there will be a reduction here someday." At the same time, however, the Fodorian anti-reductionist will point out the current inability even to conceive of how a reduction might go. Yes, the Fodorian could say, there is the logical possibility that we *might* someday find a way of characterizing a special science type in a purely physical vocabulary, even though today we can't even picture how that would be possible. But if all reductionism has in its favor is the possibility that there *might* be a common physical characterization of a type's multiple realizers, then we have no good reason to believe that many special science laws are not irreducibly autonomous.

6. Beyond the stalemate—causal closure's threat to irreductibility

There is, however, a way to move past a mere stalemate. There would be little way of moving forward if the anti-reductionist simply insisted that a physical characterization was implausible while the reductionist insisted it might happen. But if there is reason to believe that there *must* be either a disjunctive or an abstract physical/physically reducible characterization, then that's a different story. And there is, indeed, reason to believe that there must be such characterizations for *all* properties involved in causal laws. The reason stems from a view endorsed by most materialists called "the causal closure principle."

Undoubtedly, one of the reasons that "MR refutes reduction" has become the consensus view in philosophy over the last forty years is because of its ability to facilitate a nice compromise in which one can both embrace physicalism, long a central intellectual position in the west, while at the same time believing in the autonomy of the mental (another central intellectual position in the west). Nevertheless, I will argue that this compromise position is an unstable one. At a minimum, believing in physicalism requires one to assent to at least two propositions. The first is that every entity is made of some combination of physical substances (and nothing more). The second concerns what causes these substances to do what they do. Termed "the causal closure of the physical domain," it has been described by Jaegwon Kim this way: "If you pick any physical event and trace out its causal ancestry or posterity, that will never take you outside the physical domain. That is, no causal chain will ever cross the boundary between the physical and the non-physical" (Kim, 1998, p. 40). The causal closure principle holds that no particle is made to do anything without being acted upon by those physical forces typically studied in physics. Most proponents of the MR argument do endorse physicalism because of all its advantages over substance dualism. This means they must hold that whenever something is caused there must be physical forces acting on physical material doing the causing. But they also hold, at the same time, that *general laws* about what must always happen can exist that can't be reduced to these purely physical causings.[7] I contend, however, that one cannot hold both views at the same time. The causal closure principle, when combined with some other plausible ideas, surprisingly makes a commitment to the irreducibility of the mental impossible for most physicalists. My contention is that committing to the causal closure principle must lead one to commit to also holding that physical reduction is always possible, despite multiple realizability. The conventional wisdom that one can be both a physicalist and believer in the irreducible autonomy of the mental is wishful thinking.

Consider any special science generalization of the form "S_1x, causes S_2y." Suppose that the S_1x we are dealing with is, like the S_1xs we discussed in Section 4, describable as a disjunction of various different physical states. What can make it the case that every realizer of S_1x, always causes a realizer of S_2y? The causal closure of physics says that *nothing* can cause anything except via physical forces. So a set of physical forces and physical forces alone must cause each realizer of S_1x to produce a realizer of S_2y. If there are a finite number of S_1xs, S_2ys, and a finite number of processes by which a realizer of one produces a realizer of another then we can simply enumerate all of the physical ways in which S_1x realizers produce S_2y realizers. We will, thus, have used physics alone (along with bridge principles) to *explain* the generalization 'S_1x, causes S_2y, ' thus showing how this generalization can be reduced to physics.

But suppose that our S_1x is infinitely realizable like the classes discussed in Section 5. How would we explain why each of its infinite number of realizers produce only realizers of S_2y? We have just seen how. We show how the laws of physics ensure that any physical structure meeting the criteria for being an S_1x will produce only S_2y structures. An S_1x could produce something else only by violating the laws of physics (or having something other than physical forces cause this result—which would violate causal closure). "S_1x, causes S_2y," in other words, is akin to a *theorem* that can be derived from the general axioms of physics applied to certain types of circumstances (and bridge principles linking the lower and higher order vocabularies). The generalization that "water boils" at 100 °C, for example, is not a basic law of physics, but it can be derived from them—despite the infinite number of states that constitute "boiling." Similarly, as we saw in footnote 6, we can use the laws of physics to explain why no rigid square peg can go through a round hole the same length across. These explanations are similar to ones used in the disjunctive cases discussed above, except that, here, we use abstract physical/physically reducible descriptions of the *general* types of states that the laws of physics act on, instead of disjunctive *lists of more specific* types of physical states. But here, too, when we use the laws of physics to show why any arrangement of physical entities and properties that constitutes a S_1x-type must *always* result in a S_2y-type, then we've used physics to explain why that generalization must always be true. We've thus reduced it to physics.

So we see two ways in which MR special science generalizations could be true without violating the causal closure principle that most physicalists endorse. In both of these ways the special science generalization *is* reducible to physics. The central question, then, is whether there is any way *other than these two* in which special science generalizations can be true. It's difficult to see how. Lawful MR generalizations describe circumstances in which *every possible* realization of S_1x will result in an S_2y. Why must they all do this? If the causal closure principle is true, there can't be any non-physical forces that ensure only S_2ys result. And since only physical forces cause anything, there can't be any "super/meta" forces that make sure that some combination of contingencies and physical forces or other will always be on hand to guarantee that S_1xs always result in S_2ys. If some combination of facts and forces always does guarantee it, then this must be because the laws of physics *alone* are such that nothing else could happen in these circumstances. Physicalism and causal closure allow there to be no mechanisms making MR generalizations true besides systematic physical forces operating on certain possible physical arrangements. A counterfactual supporting special science generalization, then, will be true only because the physics of the world is such that any physical realiser of S_1x is subject only to configurations of forces that allow only S_2ys to result. When we *describe* what it is about the physics of the world that only allows these results, we *explain* why S_1x always causes S_2y. If there are a finite number of physical laws under which the S_1xs *are constrained* to produce only S_2ys (as opposed to an infinite number of processes by which S_1xs *happen* to produce only S_2ys), then there must be a finite physical explanation of this generalization available, in principle (even if we can't give it in practice). We can explain special science generalizations in the reductive physical ways described above, but the causal closure principle allows us no other way of explaining them.

Now it is *logically possible* that every S_1x state always produces an S_2y state with nothing systematically causing this, and nothing explaining it. It could just happens to be the case that, while no physical laws guarantee this, an infinite number of physical conditions realizing the S_1x-type, through a combination of various physical forces and circumstances, always inevitably happen to produce a S_2y-type. This would be akin to a situation in which every coin flipped by someone's left thumb, just always happened to land heads up through various ordinary physical processes, without any physical laws guaranteeing it would (also a logical and physical possibility). But do non-reductive physicalists really want to base an argument for genuine irreducibility on the supposition that this logical possibility is the actual state of affairs? (For a similar argument, see Papineau, 1992.) I don't doubt for a minute that, in practice, we may often have great difficulty giving reductive physical explanations for any given special science generalizations. Among the many possible sources of our difficulties is the possibility that the generalization isn't actually true. Another is that physicalism and causal closure are false. Another is that the reductive explanation is complex and we haven't figured out how to give it yet. Each of these is a far more plausible account of an inability to reduce than the supposition that a given special science law is a brute complex pattern explained by nothing at all. But the idea that every past, present, and future S_1x will always happen to result in an S_2y, without anything constraining that to happen, is the only supposition that a believer in physicalism and causal closure is left with, if they deny reducibility. The non-reductive physicalist needs to realize that it is this implausible stance they must defend—rather than coolly assuming, as many cognitive scientists have for three decades, that “psychological states are multiply realized and this fact refutes psychophysical reductionism once and for all.”

In summary, if some multiply realizable S_1x always causes some S_2y to happen, something must make this happen. If there are no other forces but physical ones (as the causal closure principle suggests), then it must be physical forces and physical contingencies making this happen. But, short of a miracle, there are no physical ways of guaranteeing that something will always happen except via the laws of physics. And how the finite number of laws of physics make this happen, must be finitely describable using physics laws and facts alone. Multiple realization can make it difficult to say, in physical terms, why S_1xs always produce S_2ys. But the causal closure principle leads us to believe that if S_1xs always produce S_2ys, there will be a derivation from physical laws and circumstances that says *why* this is guaranteed. Contrary to the conventional wisdom “Special Sciences” helped usher in, every causal law, even multiply realized ones, must be reducible to physics.

Furthermore, the properties named in these causal laws must be reducible physics as well. If there is a property type such that a finite number of laws ensure that a certain end state type will always result even when the initial property type has an infinite number of realizers, these realizers must be constrained to be in arrangements which have *general* features enabling it to be the case that finitely describable physical laws alone guarantee that same result. The general features might be described by physical predicates (which can be multiply instantiated), predicates about kinds of physical arrangements (which can be multiply realized while still counting as the same kind of arrangement) and arrangements of these arrangements (also multiply realizable). But, again, if these potentially infinite number of ways to realize a property are constrained in a way such that a finite number of laws guarantees a certain resulting end state, then this property must have general abstract physical/physically reducible features that are constrainable by these laws alone. And finite physical laws and general physical features (even if they have infinite sorts of realizers) are describable by us, using a physical vocabulary. If my argument is correct, then, there must, in principle, be a reductive physical description of every causal property mentioned in a universal causal law. To put the point another way: Part of giving a physical explanation of why the property S_1x causes property S_2y is physically characterizing what S_1x could be. So if a *law* is physically reducible then the *properties mentioned in the law* must be as well.[8]

The situation, in the special sciences, then, is this: All the realizers of a special science property will be physical, if physicalism is true. The realizers will be finitely describable through brute enumeration, if we are talking about a finite disjunction of properties. They will be finitely describable using abstract physical/physically reducible descriptions when there are an infinite number of realizers. And why finite or infinite realizers all happen to cause the end states they do must be explainable with physical laws, if they are not miraculous coincidences, and if causal closure is true.

Many special science generalizations may also be mixtures of abstract physical/physically reducible and disjunctive types. An infinite number of realizers of a property might all share a general characteristic, but there might also be other types of realizers of the property that don’t have this general characteristic, but have a different one. A property such as this could be described using a *disjunction of general abstract physical/physically reducible* properties.

If the abstractions are reducible to physics, then so are their disjunctions. One might also come up with an abstract general description of a disjunction (or conjunction) of different sets of disjunctions of properties—an abstract description of sets of disjunctive properties. Since each of the disjunctions is reducible to physics, a shorthand description of various collections of these disjunctions is also reducible.[9]

The seemingly mild mannered doctrine of the causal closure of physics, then, can lead to quite dramatic consequences. Causal closure, along with other plausible premises, leads one to the conclusion that if a property is involved in a causal generalization where it always leads to a particular consequences, then it has to be a property that's constrained to be describable by physics, even if it has infinite realizers. While many scholars have believed in causal closure, few have noticed how potentially damaging it is for the MR argument against reducibility.[10] Causal properties involved in causal generalizations, no matter how multiply realized, must be reducible to physics. Now the causal closure principle does allow for the possibility of *non-causal* properties and non-causal generalizations that are irreducible. It is unclear, however, why cognitive scientists should be interested in such inert properties, even if they exist (though mathematicians might be). It's safe to say that few advocates of the autonomy of the special sciences would like that autonomy and irreducibility to come at the cost of holding that the properties focused on in these sciences were causally inert. One could also argue for irreducibility by denying that the causal closure principle is true. But the denial of a basic component of physicalism, however, is a price that most scholars would find too high to pay for the autonomy of the special sciences. Scholars who are physicalists, it seems, must believe that the properties they study are ultimately reducible to physics, despite multiple realizability.

7. Reduction and institutional facts

The above argument is a general one that shows that every causal generalization must have a physical reduction. It does not, of course, go very far in showing exactly how the reduction of any particular special science generalization must go. It would not be surprising, then, if those who couldn't see how a particular special science law could possibly be reduced would still be skeptical after seeing my argument. Someone might still ask, with Fodor, how could we possibly reduce Gresham's law to general physical principles when monetary exchanges can be done with items as diverse as wampum, dollar bills, and signed checks? Despite my general argument that some physical reduction must be there, it still seems hard to imagine how any physicalistic reduction could possibly describe the commonalties in the situations that Gresham's law applies to.

I believe that many social scientific laws like this seem so utterly irreducible because there is some (unintentional) intellectual sleight-of-hand going on. The identity conditions for a property like being money are quite different from the identity conditions for a property like being milk. What makes milk what it is are certain internal chemical properties. What makes something an example of money, or someone a citizen of Mexico, by contrast, depends almost entirely on factors *external* to the bearer of the property. If one is trying to find common physical features characteristic of such properties by looking at internal structuring, then, one is looking in the wrong place. The identifying features and causal powers, if any, of such properties stem from features external to the bearer of the property in question. A reduction could only properly be sought, then, among those external features. When a property seems incapable of being reduced, in this case and in many others, that may be because one is trying to do the reduction by wrongly focusing on the internal features of the property's bearer.

It is likely that money, national boundaries, land deeds, being famous, etc. belong to a fascinating class of properties that John Searle (1995) refers to as *institutional* facts. Institutional facts are those facts that are true because people believe they are true and believe that others believe them. Wide ties truly are out of style, for example, because enough people believe they are. Similarly, people count as being married because other members of that society believe that, after certain ceremonies, such people have certain rights and responsibilities. If people in that society didn't have beliefs about the institution of marriage, nothing about a person's internal structuring could make him or her married. Reducing the property of being a husband or being money to physics has an air of inconceivability about it because examples of the bearers of such properties seem to have no physical commonalties. But we have to remember that the important commonalties and causal powers to focus on would be in the *external* features creating such properties—in this case, the beliefs of the members of a society. While the coinage involved in Gresham's laws could be radically different in different cases, the *beliefs about exchange value* people have, which leads to this law being true, might well be very similar. And while reducing certain beliefs to physical neurological states might prove to be difficult or even impossible, it is certainly not inconceivable. Indeed, it's been at the center of numerous research projects in the behavioral and brain sciences for decades. There may be no descriptions of truths about money that are true about all possible types of coinage. But economic laws about money are not really statements about any type of coinage, they are statements about behavior involving arbitrarily agreed-on markers of institutional facts, facts created by shared beliefs. A large percentage of the properties and laws of the social sciences are undoubtedly institutional facts resulting from people's networks of beliefs about their obligations to each other in various circumstances. Can such beliefs be reduced to physics? Again, the idea that they could is not at all inconceivable, and again the argument in the previous section indicates that such a reduction must be an in-principle possibility for causal properties if physicalism is true. The apparent inconceivability that Fodor uses to make the case for the impossibility of reduction is partly a by-product of focusing on trying to reduce the wrong factors.

While he does not discuss it in the text, in a footnote Fodor acknowledges the line of argument that suggests that one should initially focus on psychology when discussing the reduction of economic phenomena to physics. He dismisses this argument, however, on grounds similar to his original dismissal of the possibility of reducing Gresham's law to physics. For the reduction to succeed, argues Fodor, economic notions like "commodity, labor pool, etc." would have to be reconstructed in the vocabulary of psychology. Fodor writes, "I think it's fair to say that there is no reason at all to suppose that such reconstructions can be provided; prima facie there is every reason to think that they can not" (p. 441). As before, however, this is a mere expression of skepticism, not an argument against the possibility of reduction. For all Fodor has shown, Gresham's law might or might not reduce to physics via psychology. And my arguments in the previous sections have shown that not only *might* there be such a reduction, but that there *must be*, if physicalism and the causal closure principle are true.

8. Autonomy with reduction: concluding remarks

The MR argument of "Special Sciences" has been very influential in convincing many scholars that there could be no reduction of psychological states to physical ones, and that psychology and many other sciences should therefore be considered "autonomous." We've seen, however that the MR argument, despite Fodor's claims, does not really show "once and for all" that physical reduction is impossible. Multiple realizability does show that physical explanations of the behavior of a set of *particular* realizers of a high-level generalization might not count as a reduction of that generalization. But that doesn't mean there can't be any reductive physical explanation. We've seen that if a generalization concerns a finite number of types of properties, physicalism is committed to the idea that we can explain the generalization using an exhaustive disjunction of a property's realizers. There are no reasons not to consider such disjunctive accounts as reductive physical explanations. If the generalization involves properties that have an infinite number of realizers, this must still have an abstract physical/physically reducible explanation, if the generalization is causal and the causal closure principle is true. And some laws that seem to be irreducible only seem so because we mistakenly focus on attributes that are not the real source of the causal powers of these laws.

But the fact that physical reduction is always possible if physicalism is true, need not mean that the special sciences couldn't be autonomous in a different sense. The fact that reductive physical explanations can always be given, in principle, does not mean they can always be given in practice. Nor does it mean such explanations are always preferable. Louise Antony's arguments about multiple realizers may not really show that reduction is impossible as a matter of metaphysical identification, but they do show what an epistemic nightmare we'd face if we really tried to explain properties in terms of physical properties. These epistemic difficulties give us good *methodological* grounds to look for evidence of the *truth* of such generalizations from sources other than physics, if there are such sources. We know already, however, that there are, indeed, such alternative sources. The very existence of special sciences such as psychology shows that there are numerous ways that we can uncover promising potential generalizations, test their truth and their scope, and explore their implications other than deriving such generalizations from first principles of physics.[11] The special sciences, then, can still be thought of as autonomous, in that they can be an important independent source of knowledge about true generalizations—knowledge that's extremely difficult to get from physics alone.

And there is a way in which the ability to derive MR generalizations from physics can actually strengthen our belief in the importance of the special sciences. Such derivations show us how there *can be* generalizations at a higher level that are robust and *stable,* due to physical laws and contingencies guaranteeing that they will hold. The special sciences may provide us with empirical evidence that observed S_1xs seem to correlate with S_2ys. But we have an additional source of evidence that that correlation is lawful and necessary when we see how it is mandated by physical forces. If stable high-level generalizations exist, it certainly behooves us to know that they exist, and to explore their implications. The existence of robust high-level generalizations that physical derivation can help to verify, then, also gives us another reason for developing high-level special sciences.

The autonomy of the special sciences, one of the central features that made the MR argument attractive, then, need not disappear when we see that the MR argument does not render reduction impossible. The other feature that made the MR argument attractive—that it allowed for physicalism—is, of course, only enhanced by showing that reduction to physics must always be possible, in principle. Under the sort of physicalism that advocates of the MR argument usually endorse, each particular event in the universe is explainable in terms of physical substances and forces. Every particular mind event is identical to a brain event, so the things we learn about the brain have the potential to immensely improve our understanding of the mind. At the same time, advocates of the MR argument have held that there were generalizations or laws about the mental whose existence could never be understood by examining physical principles alone. The argument I've given here suggests that these *generalizations*, as well as particular events, are understandable in terms of physical principles. A further understanding of the physical engineering of the brain should enable us to better understand psychological *laws* and generalizations as well as particular events. The physicalism that helped make the MR argument appealing can be reconceptualized as an enhanced, more widely encompassing physicalism than that which was possible in a world where reduction wasn't. Given that the autonomy of the special sciences needn't be threatened by the reducibility of MR generalizations, the possibility of *enhancing* our understanding of psychological generalizations through such reductive explanations should lead supporters of the MR argument to treat reducibility as welcome news.

The picture of reduction and autonomy that emerges from this reexamination of the multiple realizability argument is one that scholars in diverse realms should find amenable. Jerry Fodor is fond of saying that God can do anything he likes. He could have made a world with nothing but physical properties. He could have made a world that had both physical properties and other properties, governed by special laws of their own. I think it would have been nice for God to make a world in which everything is, indeed, governed by a single set of basic laws so that anyone who knows these (plus contingencies) could potentially figure out everything. But it would also have been very nice if such a world were also one in which there are lots of different ways of producing stable systems (e.g., well-designed biological and human made artifacts). It would be nicer still if the various disjunctive ways in which these entities could cause things could be described using robust higher level laws that one could learn even if one didn't have the ability or patience to derive an understanding of each generalization from an understanding of the basic laws. I think it's fortunate that we live in a world in which all of these conditions seem to be the case. The world we live in seems to be one in which it makes sense to pursue the understanding of nature both by trying to develop a comprehensive physical "theory of everything" *and* by developing special sciences where we explore stable multiply realized generalizations, difficult to derive from physics alone. The conventional wisdom that "Special Sciences" helped to create has been that the MR argument shows the impossibility of succeeding at the first pursuit. I think we should be relieved to see that the MR argument, never, in fact, showed this. It merely points out the difficulty of reduction, and gives us a reason to pursue, in parallel, collecting higher level "special science" truths about the world. We may never have a comprehensive "theory of everything" but we should not think that multiple realizability does anything to show that that's an impossibility. The argument that it does was flawed to begin with, and is still flawed after all these years. If physicalism is true, physical explanations for special science generalizations are out there to be found.

Notes

1. By ‘physical,’ here, I mean something more than ‘embodied’. I mean entity or property types that are describable using the terminology of physics. This brings up the question of which laws and properties count as those studied by physics. One must be careful here, for if one construes “physical forces” too narrowly—say, as only those described in *today’s* physics, then something that we would intuitively want to count as physical, might not count as physical if we were to discover it has microfeatures that are unknown by contemporary science. On the other hand, one doesn’t want to define “physical forces” so leniently that a witch’s hexing, if found to exist, could be construed as just another physical force. The best solution, I believe, is to construe ‘physical’ as those things having a family resemblance to those materials and forces described in today’s physics. (See Hellman and Thompson, 1975, for an articulation of this position.)
2. In other places, I have argued that the unification theory of explanation may look like it is committed to microphysical reduction, but that it needn’t be (Jones, 1995). My arguments were based on the assumption that some generalizations could not be given a microphysical derivation. If not, then the unification theory would have to be committed to just adding these generalizations to the store of patterns we explain things with. In the case at hand, however, we are looking at generalizations which Fodor is assuming that we *can* derive, using a disjunction of accounts.

   And it doesn’t matter if this derivation is one that we would not use a number of times. Whenever we explain a *single* event, we must do so with an account that is somewhat unique—but one that incorporates predicates and patterns that are used in numerous cases. Perhaps some generalizations are explained in the same manner as single events—partly unique, but using general physical predicates, laws, and derivation patterns.

   Note that if I am right in my arguments in Sections 5 and 6, the unification theory cannot make use of the assumption that some generalizations have no microphysical reduction. If this is the case, then, despite what Kitcher intended (and what I formerly argued) the unification theory of explanation will have to count microphysical explanations as correct ones.
3. It should be noted, however, that, like many of the terms used in this realm, “reduction” is a vague one, with different scholars construing it in a variety of different ways. Doubtless, there are some construals of what “reduction” requires that would not count the explanatory and ontological identification relations I discuss here as genuine reductions. We have already seen, for example, that Antony would resist calling certain identity relations reductions because there is no practical way we could give these disjunctive identifications. Other scholars would resist calling these identifications reductions because the terms of the high-level and the disjunctive low-level terms are not fully *intertranslatable*. Proponents of possible world semantics and proponents of meaning holism, for example, could maintain that the high-level terms denoted and connoted things that a disjunction of physical terms did not. Indeed, on some views of translation and linguistic meaning, reduction will be impossible simply because there *can be no* real translations from between non-cognate languages. At the same time, however, on the classical Nagelian notion and the stronger Kim notions of reduction, the disjunc-

tive accounts I describe here would count as genuine reductions. In the end, I am not much concerned with whether the low-level explanation of special science generalizations should count as reductions. We may not want to use the term "reduction." My main argument is that if it is possible, in principle, to use physics to explain why these high-level generalizations must be true, then the special sciences don't really have the inviolable autonomy that Fodor suggests they do.

4. It is sometimes suggested that the type of disjunctive account I describe cannot count as a reduction to physics because no physical predicates specify what all of the disjunctive realizers have in common. The unity of the high-level properties, in other words, has not been accounted for by physics. This need not be a worry for the reductionist. The reductionist's project here is to explain why the generalization "S_1x causes S_2y" is true. Fodor assumes that such accounts can be given by showing how each particular S_1x realizer is guaranteed by physical law to cause a particular S_2y realizer. But to give a physical explanation of this generalization, there doesn't need to be a *further* physical explanation of the similarity of the disjuncts. Indeed, there doesn't need to be any presumption of similarity. Now one might still wonder why the different physical states that realize S_1x are all designated by the same predicate. This may be an interesting question, but we don't need to give a *physical* explanation of this to give a reductive explanation of why the *generalization* is true. Perhaps these states are all given the same name because of various psychological or historical factors. Perhaps it is just a brute metaphysical fact that S_1x *just is* P_1x *V* P_2x, *V* . . . P_nx in the same way that it is a brute fact that water is H_2O and jade is either jadite or nephrite (see Fodor, 1998). Or maybe each of these different states are labeled in the same way because they behave in a similar manner. But you don't *need* physical similarity for things to behave in the same manner—so the similarity of behavior doesn't require a physical explanation. In a big world, you might just happen to have different things that, nevertheless, behave in a similar manner. Now the fact that S_1x causes S_2y is certainly *more likely* to be the result of a genuine physical similarity in the S_1x realizers than the result of these other possibilities. (Each of these other possibilities leaves the use of a common predicate for the S_1xs and the S_2ys unexplained. And while we can physically explain their connection, why the sets of realizers that are connected both have unifying labels is unexplained. Having unexplained brute facts like these isn't impossible, but it is unusual.) But when microphysical similarity is responsible for the truth of the generalization, there isn't any need to give a disjunctive account. We can give the kind of abstract reductive explanation we discuss in Section 5. A reductionist should certainly prefer to give these sorts of unifying accounts, if he or she can get them.

   Now one might worry that allowing there to be disjunctive kinds and generalizations will "open the floodgates." What's to prevent us from saying that there are complicated disjunctive "laws" about "photographs, or tin foil, or encyclopedias?" This needn't be a worry at all. The fact that we *can* talk about all manner of disjunctive kinds doesn't mean that any more than a few need be of any scientific or practical interest to us.

5. When a state can be given a general abstract physical description, one might wonder whether such a state really counts has being multiply realizable. Part of the difficulty of answering this question lies with the vagueness of the notion of "multiple realizability."

"Indeed, as far as I know," writes Lawrence Shapiro, "no philosopher has ever tried to complete the sentence, 'N and M are distinct realizations of T when and only when . . . ' " (2000, p. 636). We certainly want to avoid a construal of "multiple realization" that is so permissive that when there are two swatches of the same blue paint on my wall, then we say that that blueness is multiply realized. Realization, a relation between properties, is surely different from instantiation, a relation between properties and particulars. On the other hand, we should count the property of being "blue paint," a property that is realized by both oil and water based paint as being multiply realized. In my usage of "realization," a property still can be multiply realized, even if it has a single abstract physical description, as long as we could further divide those realizers into general sub-type properties. On this notion of realization, "boiling water" counts as multiply realized in the same way that Putnam's "square pegs" does (see footnote 6). A physical property is multiply realized, in my usage, not merely when the property has instantiations in many different particulars; it's multiply realized when the abstract physical characterization can be realized by discernibly different sub-type properties.

It is true that some theorists, notably Kim (1993, 1998), seem to favor a more stringent notion of realization than this. On the notion of realization Kim favors, properties that count as multiply realized must have realizers so different that there are unlikely to be any laws that cover both sorts of realizers (see footnote 7). Other theorists like Fodor (1998) and Antony (2003) (who holds the German *kuh* is multiply realized by cows and bulls) have a more relaxed notion. It's extremely important to note, however, that if the argument I am making in this section is correct, *it doesn't matter whether one uses a more stringent or a more relaxed notion of realization.* If one uses a more stringent notion, then *there really aren't multiply realizable laws*, and multiple realization doesn't pose the threat to the possibility of reduction that Fodor and Putnam think it does. The different properties involved in different laws will each get their own physical reduction. On the more relaxed notion, different types of realizers could each realize the same abstract property. Here, the reductive project would be one of trying to give a disjunctive or unifying physical description of all of the multiple realizers of that property. My argument in this section concludes that there must be a physical account of all causal properties, regardless of whether one construes the vague realization relation in a more strict or more loose way.

6. Even Putnam's canonical example of physical irreducibility seems able to be given a purely physical characterization. Why can't a square peg go through a round hole the same distance across? Using a rough idealization, we can explain it in this manner: Let's start by defining "square peg." Imagine an atom, A, separated from another atom, B, of the same type by a certain distance. Imagine a third atom, C, the same distance away from A, located at right angles from the line one could imagine connecting A and B. Imagine a fourth, D, at right angles to the A–C line, the same distance from C. Now imagine four more atoms, each at right angles, in the third dimension, from each of the original four. Imagine that within the four planes that these eight atoms define are atoms with a dense proximity to each other such that if equal amounts of pressure were exerted by other sets of atoms on each of these planes, the atoms would still not move closer together because of the natural electrical repulsion of protons of the same charge.

(This states the requirement that the pegs must be rigid; without this requirement, the generalization isn't true.) Here we have a description of the concept of " rigid spare peg" in the language of physics, defined in a way that is flexible enough to apply to square pegs of all different sizes and materials.

Now imagine another pair of atoms in space, E and F, half the distance apart as A and B. Imagine an area bounded by a set of a set of atoms, as close to each other as electric repulsion allows, each the same distance from atom E as atom F is, lying on the same plane as E and F. Imagine no (or few) other atoms within this bounded area, and imagine that there are a set of atoms packed around the twice E–F distance (the diameter) at the same density as the "square peg," radiating out to form a (circular) structure, say of the same length and width as the peg-structure (forming a "wheel with a hole in the center"). (Similar densities are required, if the generalization is to be a true one). This defines a round hole the same length in diameter as the square is across, of the same density. Like the definition of "square peg," this definition is flexible enough to apply to round holes of different sizes and materials.

Now imagine the center point of the square plane at one end of the rectangular peg. Place this atom next to atom E at the center of the hole, at the closest distance allowable before repulsive forces push them apart. Imagine applying pressure to the other end of the rectangle, using any amount of force up to the point where the electrostatic attraction between the atoms in either lattice is overcome (meaning the peg or the structure around the hole breaks). Up to that point, very few atoms of the peg will pass beyond the plane of the hole because (1) the set of atoms in each of the right angle "corners" of the peg that come into contact with the atoms surrounding the hole cannot go any distance forward because they are blocked by the electric repulsion between the protons of the hole-structure's atoms, and (2) the electromagnetic attraction (an abstract physical term describing a variety of ways of bonding) that links the atoms in the peg structure together prevents the non-corner atoms from moving beyond the plane at the surface of the hole. Using abstract physical definitions in this manner, finitely describable physical properties and laws can explain why none of the rigid square pegs of various sizes and materials will be able to pass through round holes surrounded by that material.

Very often, it's hard to see how we could give purely physical definitions to certain properties, since, as with "square," property descriptions can utilize flexible, non-stringent terms that are applicable to many different physical arrangements. It can be hard to see how the same properties can be described using the much more stringent terminology of physics. But it can sometimes be done. The trick to doing so, as we see above, is to successively re-describe higher order properties by using terms describing an *arrangement of a number of* (possibly disjunctive) lower order properties (e.g., start by substituting "area bounded by a pair of parallel lines of equal length, and a second pair located at a 90 degree angle from the first" for "square"). These lower order properties can, in turn, be replaced by arrangements of still lower order properties, until we are talking only of properties describable in the vocabulary of physics (e.g., a long string of molecules lying in the shortest path between two endpoint molecules for "line"). When this can be done, one has an abstract higher-order property that is ultimately definable in terms of a set of purely physical properties, i.e., reducible to physics.

7. While most non-reductive physicalists see the causal closure principle as an innocuous component of physicalism, some see it as an illegitimate concession to reductionism that denies higher level entities causal power or begs the question about the possibility of physical reduction. I believe that it is neither illegitimate nor innocuous. There is nothing inherently controversial about the causal closure principle for non-reductive physicalists. The causal closure principle is intimately related to the idea of physical supervenience. Numerous scholars, Fodor included, believe that the fact that mental states supervene on (are realized by) physical states explains how mental states can be causally efficacious, avoiding the dualists problem of how the physical interacts with the non-physical. When Kim talks about "the physical" he means the entities and properties discussed by physics. But this principle certainly enables large *collections of* physical entities and properties (e.g., rocks, trees, a neural states realizing the belief that P) to be causally efficacious. So the causal closure principle, by itself, does not deny mental or biological or geological causation. It also doesn't beg the question in favor of the reductionist, since, by itself, it says nothing about the identification of higher level properties with lower level ones. But the causal closure principle is far from innocuous. Kim (1998) points out that when you combine this principle with other plausible principles like "no overdetermination," mental causation is not ruled out, *but its autonomy is*. Physical realizers are really doing all the causal work.

   My own argument that I develop in the text says that causal closure means that causal generalizations, like everything else, hold true because of physical facts alone. If the physical causes of these generalizations holding are explainable (as I argue they are), then all generalizations are reducible to physics. One way in which my argument seems different from Kim's is that, in *Mind in a Physical World* he focuses on the causal sufficiency of the physical realizers involved in particular instances of mental causation. He does not focus on how physical laws could explain higher order special science laws (or, conversely, how physicalistic reduction of multiply realized laws are possible). Looking at his "Multiple Realization and the Metaphysics of Reduction," (1993) we see why he doesn't focus on higher level laws. In this paper Kim argues that truly multiply realizable properties are "unprojectible" and are not likely to *have* systematic causal laws true of them. If he is right, one does not have to give a physical account of such laws, since they don't really exist. But my argument concludes that even if there are higher order laws, these, too, are explainable by physics alone.
8. I can imagine the following objection, here: if I am right about what I said about disjunctive realizers in footnote 4, there doesn't need to be any commonality to the various realizers of the generalization "S_1x causes S_2y,"—so there doesn't need to be a common finite explanation of why this generalization is true. My response is that the two kinds of cases are importantly different. The cases we are talking about in Sections 5 and 6 are ones in which 'S_1x causes S_2y' is a counterfactual supporting law with an infinite number of realizers. There must be a unified explanation of how this law results from the laws of physics—for nothing else can say *why every possible* S_1x is constrained to produce nothing but S_2ys. It can't be the case that each of the infinite number of realizers S_1x all just happens to produce similar effects. And the predicate S_1x can't apply to an infinite number of past, present, future, and counterfactual cases because of a set of historical

accidents. Here, there really must be a common property—and nothing but physics can explain why it always produces the effects it does. In the disjunctive account, we don't need such unity because we are not talking about an infinite number of possible S_1xs or S_2ys. The disjunctive and abstract physical/physically reducible explanations of multiply realizable generalizations are very different from each other. The main point is they are both kinds of reductions—reductions of generalizations that Fodor has argued are not reducible.

9. If there are an infinite disjunction of abstracts, or an abstract description of an infinite number of disjuncts that produce lawful effects, then these are reducible for the same reasons that we have been describing in this section.
10. Some, however, have noticed this, or come close. Among the earliest is David Lewis (1983), who argued (rather quickly) in his 1966 paper, "if nonphysical phenomena are ruled out by our confidence in physical explanation, only neural states are left" (p. 106).
11. Such epistemic difficulties are there for abstract physical/physically reducible general accounts as well as disjunctive ones. And there are resource conservation reasons for preferring an abstract non-physical to an abstract-physical account, in addition to epistemic ones. The amount of resources it takes to derive a generalization from physics may be for more costly to an agent than storing the information in another form (see Jones, 2003).

Acknowledgments

I would like to thank, Dave Beisecker, Andy Clark, Guy Dove, Ian Dove, Jackie Maloy, and two anonymous referees for many helpful comments and suggestions on an earlier draft.

References

Antony, L. (1999). Multiple realizability, projectibility, and the reality of mental properties. *Philosophical Topics, 26*(1), 1–24.

Antony, L. (2003). Who's afraid of disjunctive properties? *Philosophical Issues, 13*, 1–21.

Batterman, R. (2000). Multiple realizability and universality. *British Journal for the Philosophy of Science*, *51*, 115–145.

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

Causey, R. (1977). *The unity of science*. Holland: Dordrecht.

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

Fodor, J. (1993). Special Sciences. In R. Boyd, P. Gasper, & J. D. Trout (Eds.), *The philosophy of science* (pp. 429–441). Cambridge, MA: MIT Press.

Fodor, J. (1998). Special Sciences: Still autonomous after all these years. In J. Fodor (Ed.), *Critical condition* (pp. 9–24). Cambridge, MA: MIT Press.

Hellman, G., & Thompson, F. (1975). Physicalism: Ontology, determination, and reduction. *Journal of Philosophy*, *72*, 551–564.

Hempel, C. (1966). *Philosophy of natural science*. Englewood Cliffs, NJ: Prentice Hall.

Jones, T. (1995). Reductionism and the unification theory of explanation. *Philosophy of Science, 61*, 21–30.

Jones, T. (2003). The virtues of non-reduction, even when reduction is a virtue. *Philosophical Forum*, *34*(2), 121–140.
Kim, J. (1993). Multiple realization and the metaphysics of reduction. In J. Kim (Ed.), *Supervenience and mind* (pp. 309–335). Cambridge: Cambridge University Press.
Kim, J. (1998). *Mind in a physical world*. Cambridge, MA: MIT Press.
Kitcher, P. (1981). Explanatory unification. *Philosophy of Science*, *48*, 505–531.
Kitcher, P. (1989). Explanatory unification and the causal Structure of the world. In P. Kitcher & W. Salmon (Eds.), *Scientific explanation* (pp. 410–508). Minneapolis: University of Minnesota Press.
Lewis, D. (1983). An argument for the identity theory. In D. Lewis (Ed.), *Philosophical papers: Volume I* (pp. 99–107). Oxford: Oxford University Press.
Papineau, D. (1992). Irreducibility and teleology. In D. Charles & K. Lennon (Eds.), *Reduction, explanation, and realism*. Oxford: Oxford University Press.
Putnam, H. (1975a). Philosophy and our mental life. In H. Putnam (Ed.), *Mind, language, and reality* (pp. 292–303). Cambridge: Cambridge University Press.
Putnam, H. (1975b). The nature of mental states. In H. Putnam (Ed.), *Mind, language, and reality* (pp. 429–440). Cambridge: Cambridge University Press.
Searle, J. (1995). *The construction of social reality*. New York: The Free Press.
Shapiro, L. (2000). Multiple realizations. *Journal of Philosophy*, *97*(12), 635–654.
Sober, E. (1999). The multiple realizability argument against reductionism. *Philosophy of Science*, *66*, 542–564.