Identity, Variability, and Multiple Realization in the Special Sciences

Lawrence A. Shapiro, University of Wisconsin-Madison

Thomas W. Polger, University of Cincinnati

Abstract. Compositional variation and variability in nature is abundant. This fact is often thought to entail that multiple realization is also ubiquitous. In particular, compositional variability among cognitive creatures is thought to provide conclusive evidence against the mind-brain type identity theory. In this chapter we argue that the type identity theory, properly understood, is compatible with a wide range of compositional and constitutional variation and variability. Similarly, contrary to received wisdom, variation poses no threat to reductionist ventures. Multiple realization as we understand it, requires a specific pattern of variation. Multiple realization is not self-contradictory; the kinds of variation that qualify as multiple realization are not impossible, but they are less common in general than is widely supposed.

Identity, Variability, and Multiple Realization in the Special Sciences

Lawrence A. Shapiro, University of Wisconsin-Madison
Thomas W. Polger, University of Cincinnati

> From out of all the many particulars comes oneness, and out of oneness come all the many particulars.
>
> —Heraclitus

I. Introduction

Issues of identity and reduction have monopolized much of the philosopher of mind's time over the past several decades. Interestingly, while investigations of these topics have proceeded at a steady rate, the motivations for doing so have shifted. When the early identity theorists, e.g. U. T. Place, Herbert Feigl, and J. J. C. Smart, first gave voice to the idea that mental events might be identical to brain processes, they had as their intended foil the view that minds are immaterial substances. But very few philosophers of mind today take this proposal seriously. Why, then, the continued interest in identity and reduction? The concern, as philosophers like Hilary Putnam and Jerry Fodor have expressed it, is that a victory for identity or reduction is a defeat for psychology. For if minds are physical, or if mental events are physical events, then psychologists might as well disassemble their laboratories, making room for the neuroscientists and molecular biologists who are in a better position to explain those phenomena once misdescribed as "psychological." The worry nowadays is not that locating thought in immaterial souls will make psychology intractable, but that locating thoughts in material brains will make it otiose.

We find irony in the transfiguration of identity and reduction from heroes—Immaterial substances? Souls? Come on!—to villains. Too much has been made of the claim that the reduction of mind to brain undermines the legitimacy and autonomy of psychology; too little scrutiny to the claim that the diversity of creatures that exhibit mental state kinds makes reduction impossible. And far too much credence has been given to the combined thesis that it is a good thing that reduction is impossible because it is, in any event, undesirable. The idea that—should reduction go through—psychologists must vacate the premises so as to let the real scientists do their work, is baseless. We begin with a brief discussion of the early identity theorists. Place, Feigl, and Smart need not be thought of as endorsing some of the monolithic claims that are often associated with identity and reduction. Moreover, although they were primarily concerned with distinguishing their view from dualist alternatives, they also anticipated some of the issues for which Putnam and Fodor are widely credited. Seeing how Place, Feigl, and Smart thought about scientific identities will help us understand the identity theory's relationship to reductionism and the extent to which the possibility of multiple realization poses a threat. We then reconsider the commitments of reductionism. Prospects for reduction depend very much on how one understands the goals and requirements of reduction. Even when the focus is on Ernest Nagel's well-worn conception of reduction, different ways of interpreting the role of so-called bridge laws make a difference to whether the reduction of psychology to neuroscience appears plausible. Finally, we turn to the topic of multiple realization. As we shall see, the thesis of multiple realization is surprisingly hard to articulate in a way that renders it non-trivial but at the same time worrisome for the identity theory.

In the end, given all the uncertainties surrounding the proper conceptions of identity theory, reduction, and multiple realization, one might feel a temptation to bury the lot. But that

would be a mistake. On reasonable interpretations of these ideas, the identity theory is true and psychology is nevertheless autonomous.

II. The Identity Theory

According to the textbook view, Place, Feigl, and Smart, the fathers of the identity theory, argued that mental *kinds* are identical to physical *kinds*. This claim is no doubt stronger than they could have fully appreciated in the days before Kripke's argument that scientific identities are necessary and a posteriori. In the post-Kripkean world, to say, e.g., that c-fibers are identical to pain is to say that in any possible world where an organism is experiencing pain, that organism is also in possession of active c-fibers. This entails, obviously, that organisms that differ from human beings in lacking c-fibers cannot feel pain. They *necessarily* cannot feel pain.

One factor to consider when interpreting the commitments of the identity theorists is that they took themselves to be challenging dualism or, more generally, the belief that physical states, rather than somehow or other constituting mental ones, are merely *correlated* with them, suggesting that the mental exists independently of the physical. Place for example, warns against conceiving of mental events as taking place in a "mysterious internal environment (1956: 55). Likewise, Feigl inveighs against theories that stipulate an interaction between neurophysiological properties and "a totally (or partially) immaterial 'self' (1958: Part II). Similarly, Smart targets those who would understand mental reports to be referring to "an irreducible psychical something" (1959: 60). The intent of these authors was clearly to offer an alternative to the belief that minds are "outside" or "other than" entities accessible to the sciences, merely tagging along in parallel with brain processes. However, success in this endeavor does not require establishing identities between mental and physical *kinds*. As later critics have been at pains to emphasize,

one might think it would be enough to identify mental and physical tokens. So why would the early identity theorists adopt the stronger claim?

Quite plausibly, U. T. Place would not endorse the type-identity claim in its post-Kripkean form. One reason is that he seems to have been genuinely committed to the idea that the mind-brain relation is contingent. However, once Kripke cleared the way for a category of the necessary a posteriori, the identity theorist has the option of giving up the contingency claim (Lycan 1987: 19-21; Polger 2010). And the arguments employed by Feigl and Smart suggest that they should take that option. Place, on the other hand, never construed scientific identities in quite the same way as the other two (1960, especially footnote 5). And, although all three believed psychophysical identities to be contingent, Place explicated the relation on the model of material composition, which is asymmetric and (he seems to have thought) contingent. He himself holds that this compositional "identity" is the best explanation for the correlation of observations of mental states with observations of brain states. The 'is' of psychophysical identities, Place tells us, is neither the 'is' of definition ("A square is an equilateral rectangle"), nor of predication ("Her hat is red"), but of composition: "Her hat is a bundle of straw tied together with string" (1956: PPP).

Feigl and Smart, rather than explaining the correlation between mind and brain in terms of composition, avail themselves of the apparatus of co-referring expressions – just as Kripke later would. The distinction between the reference of a term and the meaning of a sentence or the beliefs of a speaker is crucial to many of Smart's responses to the identity theory's detractors. "Slugabeds," who never wake in time to see the Morning Star and are thus ignorant that Venus might be so labeled, can nevertheless refer to that planet with the expression 'Evening Star'. Similarly, 'after-image' and 'brain process of sort X' might refer to the same thing, and thus

even those who lack all neurophysiological knowledge (e.g. illiterate peasants) can, when reporting the experience of after-images, in fact be referring to brain processes of a particular sort (1959: 146).

In contrast to Place, Smart seems like a plausible candidate for endorsing a post-Kripkean identity theory, one that takes mind-brain identities to be necessary a posteriori. He claims that the identity between mental states and physical states is 'strict,' meaning that it is not simply one involving spatial and temporal coincidence, nor merely qualitative similarity. Lightning, for example, is strictly identical with an electrical discharge because 'lightning' and 'electric discharge' refer to one and the same thing. And he seems to have no special dedication to the claim that the relation is contingent, beyond adherence to the positivist framework that links necessity to analyticity.[1] Although the case for Feigl is perhaps less clear, the fact that he argues for the identity on the grounds that mental and physical terms are co-referential suggests that, if he adopted Kripke's framework, he too could endorse necessary a posteriori identities.

Saying that Smart and Feigl could endorse necessary identities is one thing, that they would or should is another. But there are reasons to think that they would or should. Most tellingly, both Feigl and Smart were concerned that a full defense of physicalist identity theory should rule out "nomological danglers" (Feigl 1958: 428; Smart 1959: 142). In Smart's work this comes out most clearly in his response to Max Black's famous Objection #3. There Smart argues that mentalistic terms do not pick out mental states by mentalistic properties, but rather by their "topic neutral" properties. Apparent reference to mental particulars or properties, e.g., a yellowish-orange after image, should instead be understood comparatively: "There is something going on which is like what is going on when I have my eyes open, am awake, and there is an

[1] For this reason he denies that the identities are necessary, for he is explicit that identity statements are not analytic (see, again, Lycan 1987 and Polger 2010).

orange illuminated in good light in front of me" (1959: 149). The proposal that mentalistic language is topic neutral is intended to eliminate the need to postulate special "psychical" properties to be the reference fixers for mentalistic terms.; and the proposal is that instead the mentalistic terms pick out similarity classes of events whose members turn out to be events of kinds that occur in brains—that is, brain event kinds. These considerations suggest that Smart and Feigl, at least, were right to endorse type-identities.

It is common to portray the identity theory as committed to a view of the world according to which mental kinds are identified with homogenous brain state kinds, and brain state kinds with homogenous biochemical kinds, and so on,… all the way down (Figure 1). If so, then any heterogeneity at any step of this "stovepipe" reduction would falsify the identity theory. We are all aware of the ubiquity of variation in nature—across species, between individuals at a time, in individual composition across time, between protein variants, between isotopes,… and all the way down. So the identity theory might appear obviously false.

Yet, the "stovepipe" view is absurd and to think that Place, Feigl, Smart or any identity theorist has ever held it is uncharitable at best. Identity theorists know as well as anyone that we live in a Heraclitean world, where variation and change are ubiquitous. They cannot reasonably be attributed the view that the truth of the identity theory requires that nature be compartmentalized into homogenous columns of reality. This is not to say that there might not be some kinds of variability that would falsify the identity theory—only that falsifying the identity theory is not so simple as finding some variation somewhere. In short, the identity theory has to be compatible with at least some extent of compositional variation.

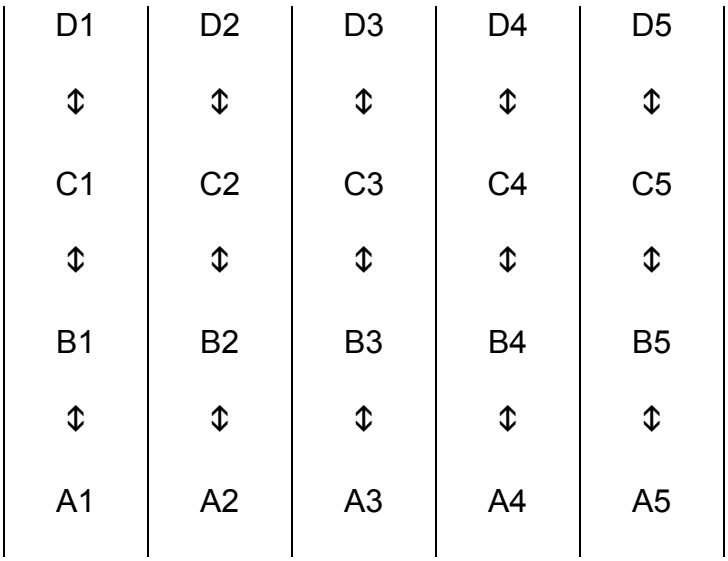

Figure 1. The stovepipe picture of reduction. Note that the levels here correspond to scientific taxonomies, which may or may not be related mereologically.

This no doubt is startling news to those many philosophers who have been nursed since Philosophy 101 on the fable that any variability decisively refutes the identity theory. If we believed the fable, we should be surprised to discover that Place, Feigl, and Smart all discussed cases of compositional variation. Feigl, for example, considers that an identity might be between a single predicate on the mental side and a disjunction of predicates on the physical side—precisely the possibility that Putnam says we "do not have to take seriously" (1967/1975: 437). Feigl has no qualms with this maneuver, noting only that the resulting identity would not be "logically necessary":

> we repudiate the *logical* translatability thesis not because of the possibility, definitely contemplated, of a one-many ψ–ϕ correspondence.

> One could always formulate such a correspondence with the help of a general equivalence between statements containing single ψ–predicates on the one side and disjunctions of statements containing several and various ϕ-predicates on the other. It is rather the logical necessity of the equivalence which is here rejected. (1958: 391)

According to Feigl, the identity theory demands only that whatever the truth, whether the correspondence turns out to be one-one or one-many, the connection between mind and brain be construed as "logically" contingent. And, of course, in pre-Kripkean times the insistence on contingency is essential to the empirical and a posteriori nature of the relationships. Contemporary identity theorists can agree that the identity statements are not analytic or a priori, while allowing that the relation they describe holds with necessity.

Smart, too, says that his identity theory was always intended to be compatible with some kinds of compositional variation:

> Compare topiary, making use of an analogy exploited by Quine in a different connection. In English country gardens the tops of box hedges are often cut in various shapes, for example peacock shapes. One might make generalizations about peacock shapes on box hedges, and one might say that all the imitation peacocks on a particular hedge have the same shape. However if we approach the two imitation peacocks and peer into them to note the precise shapes of the twigs that make them up we will find differences. Whether we say that two things are similar or not is a matter of abstractness of description. (Smart 2007)

Clearly, Smart regards the presence of internal or compositional variations among members of a type as (as in the topiary example) irrelevant to the question of whether they are type-identical.

Variation of any kind does not automatically render a type incompatible with identification—some disjunctions (unions) of variants can be kinds. Nothing in the original articulation of the identity theory requires the existence of identities between mental states and brain states described at any arbitrary level of abstraction. The textbook fable fails to distinguish between ubiquitous compositional variability and the special sort of variability that underwrites claims of multiple realization.

When re-reading the original papers of Place, Smart, and Feigl, one cannot help but suspect that the canonical interpretation of their ideas has been unfair. Far from being a slow and wounded stag that the arrow of multiple realization might easily bring down, the identity theory, as originally presented, proposed a versatile and prescient conception of the relation between the mind and the brain. When Putnam later claimed that "if we can find even one psychological predicate which can clearly be applied to both a mammal and an octopus (say 'hungry'), but whose physical-chemical 'correlate' is different in the two cases, the brain-state theory has collapsed," (1967: 77), one should feel bewilderment. Putnam never identifies the brain-state theorists he associates with this view. Whether there ever were any is a mystery.

III. Reduction

If the phenomenon of compositional variation was not news to the identity theorists, then what explains the impact of Putnam's and Fodor's multiple realization arguments? The answer is that although compositional variation had been anticipated, the use to which Putnam and Fodor put the phenomenon had not been. Putnam and Fodor worried that by identifying psychological kinds with neuroscientific kinds, the identity theorist would undermine the legitimacy of

psychological kinds as such. Putnam and Fodor thereby introduce a new foil for the identity theory: so-called non-reductive physicalism.

What's so scary about reduction? Above we challenged the common view that the identity theorists were committed to one-to-one matching between perfectly natural mental and physical kinds. But kinds, even natural kinds, are almost never unstructured entities. The members of kinds are not indiscernible from one another—sometimes they are not even indiscernible with respect to their kind-relevant properties. Consequently we should not expect them to line up as the "stovepipe" picture of reduction would suggest. This point should be obvious, but the literature about kinds and reduction suggests otherwise. For example, in an influential paper Joe LaPorte writes, with respect to isotopes, "Contrary to previous beliefs, not all atoms of any given element, or atomic number, are type identical" (1996: 118). But this claim is confusing, if not downright confused. Of course, atoms of an elemental type are, *ipso facto*, type identical—they are identical (the same) with respect to their elemental type. What they are not is indiscernible, or qualitatively identical. But the type-identity theorist does not claim that all members of the kind gold are indiscernible when she claims that the kind *gold* is identical to the kind *element with the atomic number 79*. The identity theorist is not committed to stovepipe reductionism.

Here attending to the details of the identity theory *qua* reductionist theory of mind is important. The *locus classicus* of the current "anti-reductionist consensus" is Fodor's 1974 "Special Sciences (Or the Disunity of Science as a Working Hypothesis)." Whereas Putnam's target was narrowly confined to the claim of mind-brain identity, in Fodor's hands compositional variation—multiple realization, loosely speaking – derails *any* attempt to reduce *any* special science. The centerpiece of Fodor's argument is an attack on the tenability of so-called bridge

laws which, he assumes, are necessary for any successful reduction. Although the exact significance of bridge laws for reduction is a topic to which we must soon turn, for now it suffices to see them as demanding one-to-one relations between the kinds of higher-level and lower-level sciences. As such, they appear susceptible to the multiple realization argument for the same reason that Putnam thought the identity theory was. Indeed, Fodor's argument is widely seen as a generalization of Putnam's.

Yet, even before considering more closely Fodor's reasoning, a peculiarity is worth noting. The subtitle of Fodor's paper, "The Disunity of Science as a Working Hypothesis," would suggest that the theory of reduction in his sights was that which Oppenheim and Putnam introduced in their seminal "Unity of Science as a Working Hypothesis" (1958). If this truly was Fodor's target, his criticisms do not come close to hitting it (Shapiro 2004).

Nothing in Fodor's attack on bridge laws should have been of any concern to reductionists who adopted Oppenheim and Putnam's conception of scientific unity. Oppenheim and Putnam saw reduction as a remedy to "'Hypotheses' such as Psychism and Neo-Vitalism [which] assert that the various objects studied by contemporary science have special parts or attributes" (1958: 12). The "special parts or attributes" that Oppenheim and Putnam had in mind—immaterial psyches, entelechies, vital forces, etc.—were, like Feigl's nomological danglers, excrescences that, we should hope, a simpler theory could do as well without. Reduction was a test of this idea, for if successful, reduction would show that the resources of physical science alone were enough to explain the phenomena for which psyches, entelechies and other entities that "lack any clear scientific meaning" had been recruited.

To be sure, the existence of bridge laws between psychological or biological properties, on the one hand, and physical properties on the other would be enough to dispel the possibility of

immaterial psyches or entelechies. Indeed, such would be *more* than enough; and is clearly more than Oppenheim and Putnam ever intended. The central concept in bringing unity to the sciences is not that of a bridge law, but that of a *micro-reduction*; and the "essential feature of a *micro*-reduction is that the branch B1 deals with the parts of the objects dealt with by B2" (1958: 6). Showing that the relationship between kinds in the reducing theory to those of the reduced theory is that of part to whole is sufficient for showing, Oppenheim and Putnam believe, that laws of the former can be explained by laws of latter, with no "danglers" in the remainder.

But what if pain has many possible instantiations?[2] What if thousands of different sorts of acid can function as genes? From Oppenheim and Putnam's perspective, understanding the challenge from compositional variation is difficult. Reduction, on their view, is primarily an explanatory enterprise. One rids the sciences of supernatural entities when one shows how from more basic processes, less basic phenomena occur: how, e.g, from chemical processes, life begins; or from neural processes, perception occurs. Compositional variation clearly poses no obstruction to explanatory successes of these sorts. This much is obvious from the fact that the presumed multiple realizability of biological and psychological phenomena has obviously done nothing to impede our progress in understanding them. But no part of this project depends on finding one-to-one correspondences between the kinds or predicates of various sciences.

One response to the observations above is to insist that although Fodor's "Disunity of Science as a Working Hypothesis" may indeed be poorly (sub)titled, suggesting an attack on Oppenheim and Putnam's theory of reduction when Nagel's theory was its actual target, Nagel's

[2] Of course mental states do not have brain states or physical states as parts, in the mereological sense. But this is only to emphasize that the phenomena of compositional variation and multiple realization are distinct—the former does not entail the latter.[0]

theory is more compelling, or more received, and so it is against Nagel's theory that Fodor's objections should be measured.

Our initial reaction to this point is perplexity. If Oppenheim and Putnam's unification model is not Fodor's target, this only serves to highlight the problem at hand, namely, that of explaining what sort of reduction is in play. If one concludes, on the basis of Fodor's argument or any other, that "reduction" is impossible, one must say something about the sort of reduction at stake. By the time of "Special Sciences" the Nagelian model of reduction was already reaching the end of its shelf life, and alternatives were readily available (cf. Schaffner 1967). The continued popularity of the Nagelian model in the philosophy of mind is a mystery, not an explanation.

So often lost in discussions of multiple realization and reduction is an awareness of the motivation behind reductionist projects. The consequent error is to suppose that if multiple realization via compositional variation dispatches Nagelian reduction, it is thought similarly to disarm *any* plausible conception of reduction—as if the motivation for reduction might not be met in some way other than that which Nagel described. The point is similar to that which we made above regarding compositional variation and the identity theory. If one takes Putnam's multiple realization argument (1967) as the definitive response to the identity theory, why does one not have to justify the choice to focus on a version of the identity theory that insists on physically homogenous neural kinds?

But for the sake of argument, let us suppose that Nagel's is the only theory of reduction worthy of consideration. We can still ask whether the phenomenon of compositional variability has the dire consequences that Fodor imagined. An answer to this question in part depends on how extensive compositional variation is in the world – an issue we examine in the following

section. But the answer also requires careful consideration of how Nagel actually understood the commitments of reductionism, an issue concerning which there is very little consensus.

As early as 1979, Robert Richardson argued for the compatibility of reductionism with multiple realization, claiming that Nagel would have been satisfied if the bridge laws that linked the predicates of the reducing (or *primary*) and reduced (or *secondary*) science were one-way conditionals rather than bi-conditionals. We admit that Richardson's interpretation is open to scrutiny.[3] Nevertheless, other commentators on Nagel seem to agree that his theory of reduction has the flexibility to accommodate easily the possibility of compositional variation.

Colin Klein (2009), for instance, argues that rather than endorsing a universal sort of reduction*ism*, Nagel is better read as an advocate of temporally localized reductions.[4] The distinction acknowledges the historical fact that, as Nagel says, "though contemporary thermodynamics is undoubtedly reducible to a statistical mechanics postdating 1866 (the year in which Boltzmann succeeded in giving a statistical interpretation for the second law of thermodynamics with the help of certain statistical hypotheses), that secondary science is not reducible to the mechanics of 1700)" (1961: 362). Talk of reduction*ism* is too vague: what we at best possess are *episodes* of reduction – reductions at a particular time, given the resources at that time of a primary science to explain the laws of a secondary science.

Recognizing the temporal indexicality of reduction moves one to see reduction not as a metaphysical relationship between properties, for metaphysical relationships like identity do not come and go with the ebb and flow of scientific knowledge, but rather as a measure of the

[3] Richardson (1979) cites a passage from a footnote in which Nagel notes that the connection between terms in the reduced and reducing theories is "not necessarily biconditional in form, and may for example be only a one-way conditional" (1961: 355). However, as Richardson recognizes in a later paper (Richardson 1982), Nagel then goes on to deny that one-way conditionals suffice for the condition of derivability, which, Nagel thought, was necessary for reduction.

[4] See Marras (2002) for yet another reading of Nagel that commits his theory of reduction to nothing more than one-way conditionals between the primary and secondary sciences.

explanatory power of one science to express the insights of another. Eighteenth century mechanics, lacking the representational resources to describe thermodynamical phenomena, could not reduce something like the Boyle-Charles' law. Only with the more sophisticated statistical tools developed in the later part of the nineteenth century could reduction of the Boyle-Charles' law proceed. This is to emphasize, once more, that reduction is not an end to itself, but rather a method in service of explanatory goals.

The lesson Klein draws from this observation is that reduction should not be conceived as requiring bridge laws of the sort that philosophers like Fodor and Kim attribute to Nagel. Instead, reduction requires only that the primary science have the explanatory and expressive resources to refer to the kinds that predicates in the secondary science pick out. But this requirement is hardly so stringent that it would find difficulty in accommodating compositional variability. As Klein notes, a chemical kind like salt should be indisputably reducible. This fact, however, does not entail that there is a homogenous and non-disjunctive kind in a more basic science with which salt can be uniquely identified. For every kind of salt, there is some physical description that allows the physicist to pick out the same compound to which the chemist refers. Once physics lacked these resources and explanations of the behavior of salt at a level more basic than chemistry was impossible. This is no longer true: the physicist is now able to provide explanations for why particular generalizations about salt are true. A similar story can be told for elemental isotopes.[5]

Simply put, one needn't *identify* the kinds of the secondary science with those of the primary science in order to explain why laws of the former hold. One needs only some way of talking about the kinds of the secondary science from within the primary science. Understood

[5] But see below, concerning the ways in which higher level sciences idealize. It may well be that there are not any strong generalizations about salts, but that all such generalizations are abstractions that idealize over the various salts (isotopes).

this way, compositional variation, rather than interfering with reduction, shows only that the primary science contains ample representational resources to capture all manner of generalizations about kinds that, from the perspective of the secondary science, are indistinguishable.

To summarize, in this section we have seen that, as in our earlier discussion of the identity theory, philosophers have mischaracterized reduction, painting with broad strokes a picture that is far better rendered with a sharp point. At its most basic, reduction is about explanation – the explanation of the distinctive phenomena ("laws" if you like) of one science using the resources of another. Within these confines, reduction can be interpreted in various ways. Consequently, abundant reason exists to be skeptical of refutations of reduction that treat it as a monolithic endeavor in which stovepipe bridge laws play a central and metaphysically demanding role.

IV. Multiple Realization, All Over Again

At this point the position we are advocating may seem somewhat paradoxical. Contemporary identity theorists, one might suppose, are surely committed to the idea that identities are one-to-one relations, holding across all possible worlds. So whatever might be said about the vexed notions of reduction, compositional variation still presents a prima facie problem for any identity theory. That is, multiple realization seems like it should be a problem for the identity theory even if it is not a problem for reductions as such.[6]

Evaluation of this claim requires first the recognition that different sciences generalize over distinctive domains of kinds. This suggests forsaking the imprecise question "Is kind X of

[6] Sometimes even our occasional comrade in arms, Bob Richardson, seems to take this line (e.g., Richardson 2008). But Bob also agrees with us that multiple realization of this sort is no obstacle to reductions.

science S reducible (or multiply realizable)?" in favor of the more specific question "Is kind X of science S reducible to kinds *in science T*?" The fact that the kinds of fundamental physics are particles whose various members swarm in Heraclitean clouds in the general vicinity of purported macrophysical objects (if that is a fact) may be simply irrelevant to the question of whether some psychological kinds are identical or reducible to neuroscientific kinds. If we want to know whether A=B, the relations of A or B to some third kind C are not germane.

Now you might think that A's and B's relations to C are at least evidentially relevant. But that would be the case only if we already had reason to think that B-types are sensitive to C-variations—if we already have a reason to think that C-differences ramify into B-differences. In the most salient cases, we do not have reason to think so. In fact, the advocate of multiple realization "all the way down" is committed to the view that the proprietary kinds of non-basic sciences typically screen-off the variability of kinds from those more basic sciences on which the special science kinds depend. It is the reductionist or identity theorist who contends that the variability in more basic kinds is—in some particular cases—reflected in the identity conditions of less basic kinds. But as we were at pains to point out in the previous sections, the identity theorist is not committed to the columnar "stovepipe" picture of nature, such that any variability anywhere would defeat the worldview. The identity theorist is only making a claim about the relations between kinds at two levels (Figure 2).

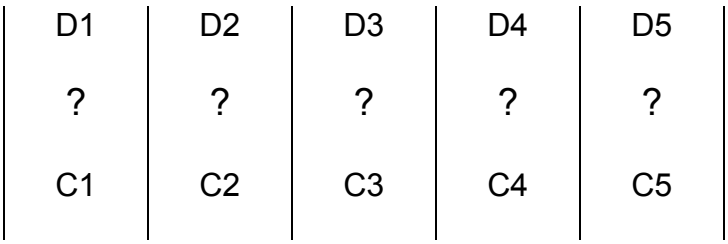

Figure 2. The question of identity is very limited: Does the taxonomy of one science cross-classify that of another, or not?

Keeping central the scientific context of issues related to identity and reduction highlights two other aspects of the identity theorists' view. One, which we discussed earlier, is that scientific kinds are not simple and unstructured. The members of scientific kinds are not indiscernible from one another: they have many individual differences in both their kind-irrelevant and kind-relevant properties. Eyes vary in their mass, but also in their size, ability to focus, light-collection capacity, and so forth. The mere fact of these variations is not reason to think that the kinds of one science cannot be identified with the kinds of another any more than it is evidence against the existence of scientific kinds at all.

The point is not that the segment of the world portrayed in Figure 2 could not be a part of a stovepipe world as portrayed in Figure 1. It could be. But our world is not such a world, and that is no problem for identifying A and B kinds. The rest of the world may be as illustrated in Figure 3, or may be entirely irregular for that matter, without ipso facto undermining the identification of some kinds with others. Of course some kinds of variation will be trouble for identities, and those are the sort that we are interested in when we consider multiple realization. Not just any sort of variation will do.

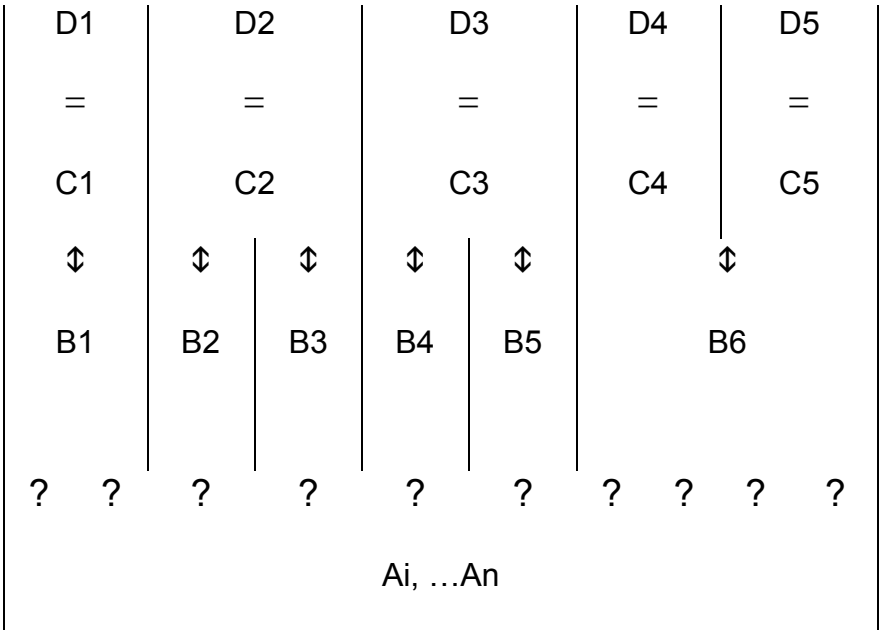

Figure 3. A non-stovepipe picture of the world that includes type identities, and leaves open what kinds of variation might ultimately be found

A second important issue concerns the extent of reduction and identity. Should the success of the identity theory require reduction of just a single kind from a given science? Many? All? Here, again, attention to the scientific motivations for identity is necessary. As against the dualist, the scientific identity theorist needs to show that there are no mental kinds that are not reducible to physical kinds of some sort or other. As we noted, the truth of the type identity theory is sufficient but (perhaps) not necessary to complete that task. As against the non-reductive physicalist hypothesis, plausibly a few or even one successful identification is sufficient. The availability of reductions or identities, according to the anti-reductionist, undermines the cognitive and psychological sciences. In this context, one good mind-brain

identity is enough to vindicate the identity theorist, and a few would be more than enough. For our part, we think the identity theorist shouldn't worry about extreme philosophical theses. The interesting scientific question is whether there are mind-brain type identities that are explanatorily or methodologically fruitful—that help us to understand connections between mental and neuroscientific kinds that would remain obscure without the insight provided by the identifications.

Finally, reminding ourselves that the identity question is always posed about actual scientific kinds alerts us to some of the practical details of scientific theorizing. Before setting out to inquire about identifications, we need to be sure that the kinds at stake are those to which the target sciences are ontologically committed. There are many non-committing apparent existential generalizations in the sciences, but some of these may be *façon de parler*, only instrumentally useful, or—quite often—simply neutral about commitment.[7] There is no chemical kind to which centers of gravity can be identified or reduced. But centers of gravity are not multiply realized by chemical kinds. Centers of gravity are merely useful abstractions. We also needn't worry ourselves over what we might call objects of convenience—those whose apparent postulation is merely a rule of thumb (dirt, say; or weeds).

One source of apparent variability arises from a mismatch of the granularity of the kinds of the sciences compared (Bechtel and Mundale 1999). Obviously there is no single neural kind associated with all pains, because the class of pains is complicated and various. Particular pain events are highly structured in space, time, and quality; and different pain events are associated with different stimuli, different time courses, different receptors, and so on (see, e.g., Polger and Sufka 2006.) Similarly we should not expect any simple neural identification involving the

[7] This is one sense in which those explanations can be autonomous without, *pace* Fodor, demanding irreducibility: a science can be neutral about reduction or even ontological seriousness of its proprietary kinds.

experience of color *simpliciter*—much less, as some have supposed, conscious experience of any sort. Such expectations are only slightly less ridiculous than the expectation of an identification of *stuff* with a kind proprietary to physics, and on par with hopes for the identification of *animal* with a kind proprietary to organic chemistry. Questions of grain may also arise when we consider identifications between kinds that come in determinate and determinable forms—indeed, the pain example is just such a case.[8]

A related source of apparent variability, one that is sometimes hard to discern from grain or determinacy issues, has been almost entirely unappreciated. Scientific kinds are frequently abstractions or idealizations. In forming scientific kinds, one often ignores or "subtracts" variations among the members, even kind-relevant variations, and even potentially kind-busting anomalies. In these cases, the presence of variability in the composition of these kinds, say neuroscientific variability among psychological kinds or isotope variability among chemical kinds, does not demonstrate multiple realization because the kind might not be "ontologically committing." We see instances of this in discussions of scientific models, where the imputed kinds are in truth idealizations of actual kinds (Klein 2008; Haug forthcoming).

Taking stock for a moment: Before we can evaluate that a case of compositional variability falsifies some kind identity claim we need assurances that:

(1) we are ontologically committed to the purportedly multiply realized kind: it should not be an idealization, an instrumental kind, or a merely heuristic kind.

(2) the variability is in the proprietary scientific kinds whose identity is being considered, not in some other kinds.

(3) the grain of the kinds matches.

[8] We think it's a mistake to think of determinates as multiply realizing (or realizing) their determinable properties. For reasons, see Funkhouser (2006), Walter (2007), and Haug (2010).

(4) the variation is not just a matter of individual differences, as between determinates of a determinable.

(5) the location and extent of the variation is incompatible with the scope of the identity claim.

Once we've convinced ourselves that the observed or hypothesized variability is at least a candidate for being a case of multiple realization, then we can finally apply some precise criteria.

The criteria that we endorse begin with the slogan: Same but different. The checklist above is designed to get us this far, to ensure that we have at hand a case of sameness with difference. But as we have argued before and mentioned above, even then we do not have an example of multiple realization unless the differences are *relevant* differences (Shapiro 2000, 2004, 2008). Most of the differences within members of any kind are irrelevant differences, differences that play no part in the kind membership: the colors of various corkscrews, the mass of various dolphins, the rarity of various stones, the electrical charge of books. The differences that count as multiple realizations of a kind—differences with respect to a scientific taxonomy that would block its identification with some other—must be differences that are relevant to the second. Color is a relevant quality of gemstones but not of corkscrews, so color differences are candidates for individuating multiple realizations of gemstones but not of corkscrews.

Moreover, multiple realization requires that the relevant differences nevertheless do not ramify into differences in the multiply realized kind. That is, a kind in one science is multiply realized by various kinds in another science when they are both *relevantly different* and *differently the same* (Shapiro 2008). Two kinds of cases are covered by this second requirement. Suppose that we have one kind b_i in science B, and we want to know whether it is multiply realized by kinds a_1-a_n in science A. First, perhaps obviously, the differences in the kinds a_1-a_n

must be such that they are all still classified as b_i in science B (they are differently the same). Otherwise, instead of multiple realization we have an example of kind-splitting, where, e.g. a_1-a_j realize b_i and a_k-a_n realize some kind other than b_i.(see Kim 1972, Craver 2004, Polger 2002, 2004).

Second, even if a_1-a_n are differently the same with respect to being of kind b_i, they won't count as multiple realizations of b_i if their variation merely explains individual differences within kind b_i – they must be relevantly different. Realizations of b_i may differ despite not differing in the ways in which they are being b_i. Possibly we could make this argument with Kim's example of the gemological kind jade and the mineral kinds jadeite and nephrite. Jaedite and nephrite do not have exactly the same mineralogical properties, but they are very similar. The judgment that jade is not a natural kind reflects the discovery that jadeite and nephrite are *not* differently the same – their differences suffice to render them different from the perspective of geological science.[9]

Consider next our favorite toy example, corkscrews. Corkscrews differ in many kind-relevant ways. We imagine that a science of corkscrews, were there one, would individuate corkscrews according to the mechanical principles by which they exert force on the bottle or cork. So we suppose that, for example, waiter's corkscrews, double-lever corkscrews, and simple hand screws count as genuine multiple realizations of the kind *corkscrew*. They are differently the same—they differ in how they transmit the force that overcomes the friction between the cork and the bottle and yet these differences do not force the sort of kind-splitting present in the jade case. But not every difference within or between kinds of corkscrews is significant. Double lever corkscrews may differ in color or composition while remaining of a

[9] More likely jade fails to pass the initial checklist, and is not a serious scientific kind.

kind (Shapiro 2000, 2004). Furthermore, there may be differences within kinds of corkscrews even with respect to those properties that are "corkscrew relevant", but which do not amount to cases of multiple realization. For example, the length of the lever arms of a double lever corkscrew is clearly relevant to how much force it can apply to the cork, and thus to its performance *qua* corkscrew. But short-levered and long-levered corkscrews are not different kinds of corkscrews. These are just individual differences within the kind *double-lever corkscrew*. And, of course, lever length is a continuous value, so there are indefinitely many variations. These are relevant differences that contribute to sameness (being a double levered corkscrew), but contribute to the sameness differently—by constituting individual differences among double-levered corkscrews. This is not to say that no lever-length difference would be kind-changing. It seems likely that a lever length of zero would be a different kind of thing; but in that case it seems likely that the "multiply realized" kind has been split into *double-lever corkscrews* (non-zero lever length) and *simple-corkscrews* (lever length of zero).

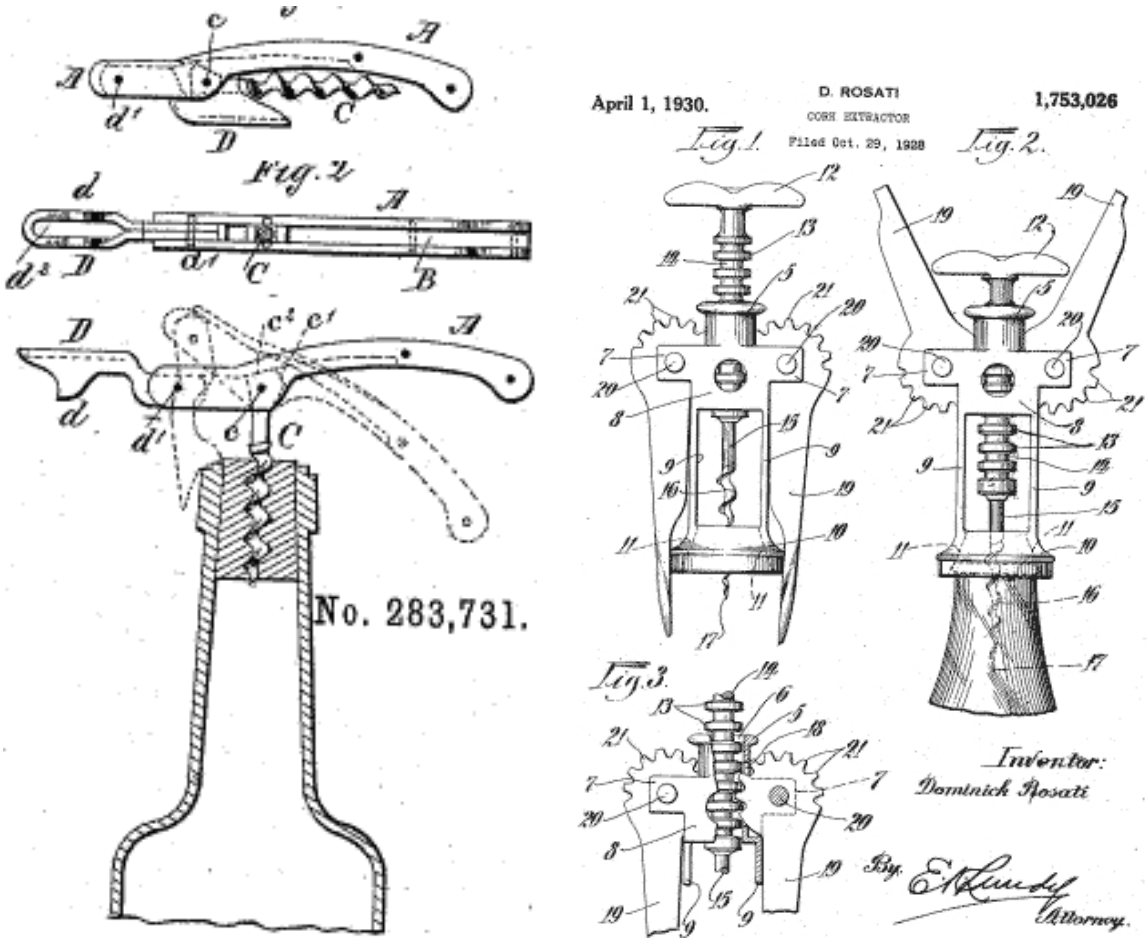

Figure 4. Waiters and winged corkscrews

This kind of consideration partly explains why, *pace* Aizawa (e.g., with Gillett 2009), we think that camera eyes with retinal cone cells that differ in spectral sensitivities due to their possession of different opsins (photosensitive chemicals in the cell bodies) do not count as multiple realizations of the kind eye. Even if the eye-opsin kind identity cleared our checklist above, the differences caused by the different opsins would still likely be individual differences from point of view of the science that classifies eyes.[10] That's because the biological taxonomy

[10] It doesn't clear the checklist.

that individuates kinds of eyes distinguishes them by their morphology and how they focus light, not by the chemistry of their photoreceptors (Shapiro 2004; Polger 2009).

We therefore adopt a four part criterion (Shapiro 2008, Polger 2009). Multiple realization of one kind by another occurs when:

(i) two entities A and B are classified commonly by system S1, and

(ii) A and B are classified distinctly by system S2. And in addition,

(iii) the factors about A and B that lead them to be differentially classified by S2 must be among those that lead them to be commonly classified by S1, and

(iv) the relevant S2-variation between A and B must be greater than the S1 individual differences between A and B.

Here we use clause (iii) to capture the idea that multiple realization requires A and B to not merely be different, but to be "differently the same"—to be different in ways that are relevant to their sameness. Winged and waiter's corkscrews are different in ways that contribute to their cork-removing capacities; camera and compound eyes are different in ways that are relevant to their light-sensing capacities. Differently colored corkscrews are the same with respect to their cork-removing capacities—they are samely the same, not differently the same. And remembering the factors of granularity and abstraction, we can say the following: Considered coarsely, camera eyes with different photoreceptive chemicals in their retinal cones are all doing the same thing in the same ways, so they are not differently the same; but considered finely, they are sensitive to different ranges and peaks of spectral stimulation, so they are differently different, not differently the same.

And we use (iv) to formulate the "differently the same" part of the criterion in terms of quantitative difference: for multiple realization, the difference must be "more" than mere individual difference. But the quantitative criterion is not strictly accurate, and could be misleading. For the crucial aspect is not that multiple realization requires a large difference—sometimes small differences may contribute to multiple realizations—but that the variation must not merely map onto individual differences.[11]

Finally, that we formulate our criterion in terms of the sameness and difference of objects with respect to scientific taxonomies deserves emphasis. Thus, like the kind categories themselves, sameness or difference with respect to a scientific kind is a messy matter both in practice and in principle. (This, of course, is what leads to abstraction, idealization, and ontologically non-committing categories.) Whether two things belong to one kind or not, within a given science, is frequently a difficult question for investigators, and may depend as much on the explanatory and methodological utility (and promises) of the kind or the science as on any criteria for kind membership that might be fixed in advance. Sameness and difference are relative judgments — investigators must decide, given their particular and often various purposes, when judgments of sameness and difference are applicable (Shapiro 2000, Polger 2008, 2009).

We have argued elsewhere that once one understands the many facets involved in multiple realization that its significance becomes dubious. At the very least, we lack the kinds of evidence for multiple realization that many philosophers assume must be available merely because of the scientific common-sense that the world is full of variability. And from an

[11] It is useful to remember that the S1 and S2 classificatory systems may operate at the same or different mereological levels. An anvil and a corkscrew may be classified similarly by the things-that-can-be-paperweights science and differently by the things-that-can-be-anchors science (same level) and differently by the things-that-react-with-acids science (lower-level).

empirical point of view we should expect multiple realization to be a fairly rare phenomenon. Here we've emphasized that because the identity theorist's view can be relatively modest and multiple realization is quite demanding, the identity theory needn't have much to worry about. Similarly, the explanatory goals of reduction, allowing as they do numerous approaches to reduction, are often quite consistent with multiple realization – whether rare or not. In short, life in a Heraclitean world should hardly dampen the enthusiastic pursuit of type identifications and reductions. And it's a good thing, too—because identity theorists and reductionists have always known that we live in such a world.

V. Acknowledgments

Brian McLaughlin told us that he thought that there have never been more than five identity theorists at any point in time. We would like to thank the other three for fighting the good fight with us. We are grateful to Simone Gozzano and Christopher Hill for the opportunity to write about our favorite topic. And we have benefited greatly from discussing these matters with Fred Adams, Ken Aizawa, John Bickle, Carl Craver, Carrie Figdor, Carl Gillett, Doug Keaton, Sarah Paul, Bob Richardson, Alan Sidelle, Elliott Sober, Jackie Sullivan, and Sven Walter. Tom Polger's work is supported in part by the Charles P. Taft Research Center for the Humanities at the University of Cincinnati.

VI. References

Aizawa, K. and C. Gillett. 2009. Levels, Individual Variation and Massive Multiple Realization in Neurobiology. In J. Bickle (Ed.) *Oxford Handbook of Philosophy and Neuroscience* (Oxford: Oxford University Press): 539-581.

Bechtel, W. and J. Mundale. 1999. Multiple realizability revisited: Linking cognitive and neural states. *Philosophy of Science*, 66: 175-207.

Bechtel, W. and R. McCauley. 1999. Heuristic identity theory (or back to the future): The mind-body problem against the background of research strategies in cognitive neuroscience. In the *Proceedings of the 21st Annual Meeting of the Cognitive Science Society* (Mahwah, NJ: Lawrence Erlbaum Associates): 67-72.

Block, N. (ed.) 1980. *Readings in Philosophy of Psychology, Volume One*. Cambridge, MA: Harvard University Press.

Block, N. and J. Fodor. 1972. What psychological states are not, *Philosophical Review* 81: 159-181.

Clapp, L. 2001. Disjunctive Properties: Multiple Realizations, *Journal of Philosophy*, 98: 111-136.

Craver, C. 2004. Dissociable Realization and Kind Splitting. *Philosophy Of Science* 71 (5):960-971.

Feigl, H. 1958. The 'Mental' and the 'physical'. In Feigl, Maxwell, and Scriven (eds), *Minnesota Studies in the Philosophy of Science*, 2: 370-497.

Fodor, J. 1974. Special sciences, or the disunity of science as a working hypothesis. *Synthese* 28: 97-115. Reprinted in Block 1980.

Fodor, J. 1997. Special sciences: Still autonomous after all these years. In J. Tomberlin (Ed), *Philosophical Perspectives 11: Mind, Causation, and World* (Boston: Blackwell Publishers, 1997.)

Funkhouser, E. 2006. The Determinable-Determinate Relation. *Noûs* 40: 548-569.

Haug, M. 2010. Realization, Determination, and Mechanisms. *Philosophical Studies*. 150: 313-330.

Haug, M. forthcoming. Abstraction and Explanatory Relevance, or Why Do the Special Sciences Exist? *Philosophy of Science*.

Kim, J. 1972. Phenomenal Properties, Psychophysical Laws, and Identity Theory. *Monist* 56, 2: 177-192. Excerpted in Block (1980) under the title, "Physicalism and the Multiple Realizability of Mental States."

Kim, J. 1989. The Myth of Nonreductive Materialism. *Proceedings and Addresses of the American Philosophical Association* 63, 3: 31-47. In Kim (1993).

Kim, J. 1998. *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation*. Cambridge, MA: MIT Press.

Klein, C. 2008. An Ideal Solution to Disputes about Multiply Realized Kinds. *Philosophical Studies* 140 (2): 161-177.

Klein, C. 2009. Reduction without Reductionism: A Defence of Nagel on Connectability. *Philosophical Quarterly* 59 (234): 39-53

Klein, C. UM. Multiple Realizability and the Semantic View of Theories.

Kripke, S. 1971. Identity and Necessity. In M. Munitz (Ed), *Identity and Individuation* (New York: New York University Press, 1971.)

Kripke, S. 1972/80. *Naming and Necessity*. Cambridge, MA: Harvard University Press.

LaPorte, J. 1996. Chemical Kind Term Reference and the Discovery of Essence. *Noûs* 30 (1):112-132.

Lewis, D. 1966. An Argument for the Identity Theory. *Journal of Philosophy*, 63: 17-25.

Lewis, D. 1969. Review of Art, Mind, and Religion. Journal of Philosophy 66: 23-35. Excerpted in Block (1980) as "Review of Putnam."

Lewis, D. 1970. How to Define Theoretical Terms. Journal of Philosophy 68: 203-211.

Lewis, D. 1972. Psychophysical and Theoretical Identifications. *The Australasian Journal of Philosophy* 50: 249-258.

Lycan, W. 1987. *Consciousness*. Cambridge, MA: The MIT Press.

Marras, A. 2002. Kim on Reduction. *Erkenntnis* 57: 231-257.

Nagel, E. 1961. *The Structure of Science: Problems in the Logic of Scientific Explanation*. London: Routledge and Kegan Paul.

Oppenheim, P. and H. Putnam, 1958. The unity of science as a working hypothesis. In H. Feigl et al., eds., *Minnesota Studies in the Philosophy of Science*, vol. 2, Minneapolis: Minnesota University Press.

Place, U. T. 1988. Thirty Years On—Is Consciousness Still a Brain Process? *Australasian Journal of Philosophy*, 66: 208-219.

Place, U.T. 1956. Is Consciousness a Brain Process? *British Journal of Psychology*, 47: 44-50.

Place, U.T. 1960. Materialism as a Scientific Hypothesis. *Philosophical Review*, 69: 101-104.

Polger, T. 2002. Putnam's intuition, *Philosophical Studies*, 109, 2: 143-170.

Polger, T. 2004. *Natural Minds*, Cambridge, MA: The MIT Press.

Polger, T. 2007. Some Metaphysical Anxieties of Reductionism. In *The Matter of the Mind: Philosophical Essays on Psychology, Neuroscience and Reduction*, M. Schouten and H. Looren de Jong (eds.), Blackwell Publishers.

Polger, T. 2009a. Evaluating the Evidence for Multiple Realization. *Synthese* 167 (3): 457-472.

Polger, T. 2009b. Two Confusions Concerning Multiple Realizability. *Philosophy of Science* 75 (5): 537-547.

Polger, T. 2009c. Identity Theories. *Philosophy Compass*,4 (4): 1-13.

Polger, T. 2010. Are Sensations Still Brain Processes? *Philosophical Psychology*.

Polger, T. and K. Sufka. 2006. Closing the Gap on Pain: Mechanism, Theory, and Fit. In *New Essays on the Nature of Pain and the Methodology of its Study*, M. Aydede (ed.). Cambridge, MA: The MIT Press.

Putnam, H. 1967/1975. The Nature of mental states, in H. Putnam (Ed.) *Mind, Language and Reality: Philosophical Papers, Volume 2* (New York: Cambridge University Press, 1975).

Putnam, H. 1974. Philosophy and our mental life. In H. Putnam (Ed.) *Mind, Language and Reality: Philosophical Papers, Volume 2*. (New York: Cambridge University Press, 1975).

Richardson, R. 1979. Functionalism and Reductionism. *Philosophy of Science* 46 (4):533-58.

Richardson, R. 1982. How Not to Reduce a Functional Psychology. *Philosophy of Science* 49 (1):125-37.

Richardson, R. 2008. Autonomy and Multiple Realization. *Philosophy of Science* 75 (5):526-536.

Salmon, N. 1979. How *Not* to Derive Essentialism from the Theory of Reference. *Journal of Philosophy* 76: 703–725.

Salmon, N. 1982. *Reference and Essence*, Oxford: Basil Blackwell.

Schaffner, K. 1967. Approaches to Reductionism. *Philosophy of Science* 34: 137-147.

Schwartz, S. 2002. Kinds, General Terms, and Rigidity. *Philosophical Studies* 109: 265–277.

Shapiro, L. 2000. Multiple realizations. *The Journal of Philosophy*, 97, 635-654.

Shapiro, L. 2004. *The Mind Incarnate*. Cambridge, MA: The MIT Press.

Shapiro, L. 2008. How to Test for Multiple Realization. *Philosophy of Science* 75: 514-525.

Smart J. J. C. 1959. Sensations and Brain Processes. *Philosophical Review*, LXVIII: 141-156.

Smart, J. J. C. 1961. Further Remarks on Sensations and Brain Processes. *Philosophical Review*, 70 (3): 406-407.

Smart, J. J. C. 2007. The Identity Theory of Mind. Stanford Encyclopedia of Philosophy. <http://plato.stanford.edu/entries/mind-identity/>.

Soames, S. 2002. *Beyond Rigidity*. New York: Oxford University Press.

Soames, S. 2004. Knowledge of Manifest Natural Kinds. *Facta Philosophica* 6: 159–181.

Walter, S. 2007. Determinables, Determinates, and Causal Relevance. *Canadian Journal of Philosophy*, 37: 217-244.