The functional unity of special science kinds

Daniel A. Weiskopf

Abstract: The view that special science properties are multiply realizable has been attacked in recent years by Shapiro, Bechtel and Mundale, Polger, and others. Focusing on psychological and neuroscientific properties, I argue that these attacks are unsuccessful. By drawing on interspecies physiological comparisons I show that diverse physical mechanisms can converge on common functional properties at multiple levels. This is illustrated with examples from the psychophysics and neuroscience of early vision. This convergence is compatible with the existence of general constraints on the evolution of cognitive systems, and does not involve any *ad hoc* typing of coarse-grained higher level properties. The mechanisms that realize these common higher level properties are really distinct by the criteria laid down by critics of multiple realizability. Finally, I present an account of how such functional properties might constitute special science kinds by playing a central explanatory role in a range of cognitive models. Behavioral science kinds in particular are the functionally defined constituents picked out by our most successful models of the multilevel systems and mechanisms that explain cognitive capacities.

1 Reduction versus elimination, redux

Realization is a relation between a property Ψ at one level of organization and a property Φ or family of properties Φ_1-Φ_n at a lower level of organization. According to the Multiple Realizability (MR) thesis, psychological properties, as well as the properties in the domain of many other special sciences, are multiply realizable. MR is at least supposed to be true of properties that are defined by some purpose, capacity, or contribution they make to some end—generally, by their functional role. Where there are interestingly different ways of playing the role that defines the property Ψ, then Ψ has different realizations. For Ψ to be *multiply* realizable, the Φs must belong to distinct kinds, as defined by some independent taxonomy.

The claim that psychology is multiply realizable played a crucial role in the rise of functionalism in philosophy of mind, and it has long been argued that the special sciences in general deal with kinds that are multiply realizable. Against this consensus, however, Lawrence Shapiro ([2000]) has argued that the MR thesis is not even coherent. Consider the Φs that purportedly realize Ψ. Either:

(1) these realizers differ in their causally relevant properties, or

(2) they do not.

'Causally relevant properties' are those that enable something having Φ to perform the function of a Ψ. What makes something a corkscrew, for instance, is the fact that it has certain causally relevant properties C that make it able to perform the function of a corkscrew. Consider two corkscrews, one a winged and one a waiter's model, that satisfy the relevant functional specification but do so by having different causal powers C_1 and C_2. This corresponds to case (1) above. In such a situation, the Φs are different kinds, since kinds are individuated in terms of their causal powers and the inductive generalizations they underwrite. 'But if they are different kinds then they are not the same kind and so we do not have a case in which a single kind has multiple realizations' (Shapiro [2000], p. 647). That is, if the Φs possess different causally relevant properties, then Ψ *itself* does not constitute a kind, and hence one higher-level kind isn't being multiply realized. But if (2) is the case, then they do belong to the same kind, hence they are not different realizations and the MR thesis is false in this instance as well.

This is Shapiro's dilemma. His argument is closely related to one presented by Jaegwon Kim ([1992]), who also concludes that multiply realized properties cannot be kinds, on the grounds that realizers belonging to distinct lower level kinds lack the causal unity needed to realize the same higher level kind. While my reply will be to Shapiro's formulation of this problem, it can also be seen as addressing Kim's closely related worries. Here I will assume that if Φ_1 and Φ_2 are different independently certified kinds then we have genuine MR. This is at least a sufficient condition on realizers' being distinct. So the solution to Shapiro's (and Kim's) dilemma that I'll offer involves defending type-(1) cases. What I deny is that the Φs being different kinds entails that Ψ is *not* a kind. The main burden of this paper is to show how functional properties that are heterogeneously realized can constitute kinds in psychology and neuroscience.

As I've noted, realization is a relation among properties (or possibly instances of properties—the debate between these positions won't matter here). On the view I adopt, properties are comparatively cheap. To a first approximation, where we have a meaningful and non-contradictory predicate we have a property that is the truthmaker for claims involving that predicate. Being red is a property expressed by ‘is red’, being a table is a property expressed by ‘is a table’, and so on. There are also such things as the property of being five feet from the Eiffel Tower, or the property of being either an electron or a bumblebee. And of course many properties are unnamed by any predicates of any language.

One can adopt a more restrictive ontology of properties than this. Some properties seem to make no difference to the objects that have them; consider the property of being *headsup*, which everything has in the event that my coin lands heads up, and lacks otherwise. Are such ‘Cambridge properties’ properties at all? One might propose that genuine properties correspond only to causal powers—capacities to produce certain effects under certain circumstances—which would remove such useless properties as being headsup from our ontology. This causal powers view of properties has been developed extensively in recent years (Heil [2003]; Shoemaker [2003]). I won’t enter into the debate over whether we should be restrictive or liberal with respect to properties, or whether we should be restrictive in the way that causal theorists of properties propose, since what matters for my purposes is only that there is a distinction between properties and kinds.

Kinds, by contrast with properties, are precious. Even on a causal powers view of properties, not all properties correspond to kinds. Whether Ψ is a kind or not depends on whether there is a sufficiently large and interesting body of empirical regularities in which Ψ itself is implicated. For there to be a kind there must be a rich cluster of properties that reliably co-occur with something’s being Ψ, where these properties do not cluster together by chance but by the operation of some governing principle or mechanism. The main point is that only some properties (or property clusters) correspond to kinds. Kinds are non-accidental loci of inductive potential. This is a relatively generic definition of kindhood; there may be many different ways of filling in the details (see, e.g., Boyd [1991]; Millikan [1999]). Both physics and the special sciences aim to model the structure and behavior of presumptive kinds: electrons, hurricanes, pyramidal cells, beta endorphins, visual memory buffers, lungfish, etc. The debate over multiple realizability has focused on kinds in this sense—groupings that are the distinguished subject matter of the special sciences.

The debate thus bears on two related questions. First, are any properties genuinely multiply realizable? And second, if some are, are any multiply realized properties kinds? It might seem that if Φ_1 and Φ_2 are causally different ways of realizing Ψ that this would in itself show that they did not participate in any (nonanalytic) common regularities and hence could not belong to a common kind. But this isn't obviously true, since, as we will see, distinct neurobiological mechanisms can still give rise to shared functional and psychological properties and generalizations. I will argue that for some causally relevant properties C_1 and C_2, there are kind-making empirical commonalities that hold among things having those distinct properties. We can see this by looking at an example that Shapiro himself discusses: the case of compound versus camera eyes.

2 The eyes of others

Arthropod compound eyes and vertebrate camera eyes are all eyes in virtue of falling under the functional description 'organs for seeing'. This is a purportedly analytic generalization about eyes—indeed, perhaps the only one. But different mechanisms are involved in the production of sight in each kind of eye; hence, by the anti-MR argument, eyes should not be a single kind. However, it turns out that both kinds of eyes can display similar psychophysical phenomena despite having different neurobiological and optical properties.

The main phenomenon of interest is the fact that the visual systems of many unrelated species allow them to experience *Mach bands*: regions of especially high or low brightness that occur just after the ends of a brightness gradient. These bands appear as light or dark stripes in an image. Mach was the first to investigate the phenomenon in detail and to show that it was a psychophysical rather than an objectively physical effect (Ratliff [1965]). Bands are variable in their thickness, brightness, and position, and can be induced to change their appearance under the influence of a wide range of contextual factors. While perception of Mach bands occurs in many organisms, including primates, cats, and horseshoe crabs (*Limulus polyphemus*), the neural circuits that underlie it differ radically across species. This can be illustrated with respect to the *Limulus* eye and the mammalian eye.

Much of the early work on the physiology of Mach band perception used *Limulus* as a model organism. The lateral eyes of the crab are compound eyes composed of ~1000 cones that terminate in structures known as ommatidia (Battelle [2006]). One of the most detailed analyses of the wiring pattern of the *Limulus* retina is given by Fehrenbach ([1985]), who maps every branch and synapse among five ommatidia extracted from a single organism. The ‘retina’ of these eyes is extremely primitive—the ommatidia contain photoreceptive cells that depolarize a central eccentric cell, the axons of which bunch together to form the optic tract. Eccentric cell axons also distribute collaterals to their neighbors in adjacent ommatidia. These interwoven branching collaterals form an anatomical structure known as the lateral plexus of the eye. The function of the collateral of the lateral plexus is to enable the activity of one ommatidium to be inhibited by activity in adjacent ones (Hartline and Ratliff [1957]). So highly stimulated photoreceptors tend to dampen the afferent signals coming from surrounding cells: when one is strongly 'on', its neighbors are nudged towards being 'off'. This lateral inhibition mechanism enhances contrast and sharpens perception of edges, thereby serving as an early stage in processing the visual image. This is achieved, as Fehrenbach ([1985]) notes, despite the fact that the plexus itself constitutes a 'seemingly random tangle of neuronal fibres' (p. 244).

According to one widely accepted theory—indeed, arguably the standard theory—this same mechanism also explains the perception of Mach bands. Where there is a luminance gradient, a set of receptors that laterally inhibit each other will display patterns of activation that correspond to the perceived brightness of the Mach bands themselves. Neural activity will be slightly greater on the more luminous side of the gradient and slightly lower on the less luminous side (Palmer [1999], pp. 115-7). This pattern is a qualitative match for the phenomenology, which suggests that this neural mechanism is at least implicated in the phenomenology. This inference embodies one regularly used heuristic for mapping neural states onto conscious experience. Where a set of neurons carries information about some property, e.g., luminance, and where activation patterns in those neurons correlate with some aspect of experience, e.g., brightness, those neurons are plausibly part of the causal process that eventuates in the conscious experience itself. Such patterns of neural activity account for the fact that experience does not map detected luminance directly onto perceived brightness. Perceiving Mach bands, then, is a kind of side effect of having a visual system that uses lateral inhibition to perform useful visual processing.

Mammalian eyes, while physically and optically different from compound eyes, also contain inhibitory mechanisms that produce Mach bands. In contrast to the loose organization of the lateral plexus, mammalian retinas are tightly organized into distinct layers: outer plexiform, inner nuclear, and inner plexiform. They also use a vastly greater range of cell types than does the *Limulus* eye. Photoreceptive cells feed into a complex network containing horizontal, amacrine, and bipolar cells, finally terminating at ganglion cells that project to higher regions. (These classifications are the broadest ones available: to indicate the complexity here, there are at least two types of horizontal cell in primates, and over 40 types of amacrine cells in vertebrates.) While there are many possible loci for lateral inhibition in the retina, it seems to occur initially in the horizontal cells linking adjacent rods and cones. These cells have highly specific connectivity patterns, as opposed to the near-random wiring of *Limulus* (Field and Chichilnisky [2007]; Sterling [1998]). Horizontal cells are also electrically coupled to their neighbors by gap junction synapses, which means that they can act as a single 'sheet' that filters incoming neural signals. Activity in one photoreceptor can result in inhibited activity in adjacent receptors and their forward projections via both feedback and feedforward modulation by horizontal cells. This is also the earliest known stage at which characteristic center-surround antagonistic fields appear in visual processing (Mangel [1991]).

This brief comparison of the biology of vertebrate and horseshoe crab retinas is sufficient to show that they incorporate significantly different neural mechanisms. A mechanism can usefully be thought of as 'a set of entities and activities organized such that they exhibit the phenomenon to be explained' (Craver [2007], p. 5). The phenomenon of interest here is producing lateral inhibition, and thereby producing both contrast enhancement as well as Mach bands and a host of other related psychophysical phenomena. The component entities of interest are primarily the neurons themselves and their spatial organization relative to each other. Their activities consist of how they respond to various kinds of stimulation and how their firing patterns interact to produce the complex phenomenon of mutual lateral inhibition. These activities involve various sorts of graded potentials and action potentials, mediated by both chemical synapses using different neurotransmitter types (histamine in *Limulus*, GABA in mammals) and by electrical synapses. The dynamics of these activities and interactions are different in each mechanism; eccentric cells are self-inhibiting as well as neighbor-inhibiting, whereas horizontal cells aren't, for example. Finally, these mechanisms enable different sorts of interventions, depending on the specific types of neurons involved, their particular wiring pattern, and the chemical and electrical messengers that they use to carry out their characteristic activities. It would certainly be impossible to simply 'swap out' the mechanism of lateral inhibition in *Limulus* and install it in a vertebrate retina and expect it to work. All of this suggests that by commonly accepted neurobiological criteria, these mechanisms count as belonging to different kinds.

One note of caution on the preceding story: while it is standard practice to explain Mach bands by appealing to lateral inhibition in the retina, this explanation may not turn out to be correct. Palmer ([1999], p. 117) notes that the distances over which inhibition operates may be less than the distances over which simultaneous contrast effects can be found. And Pessoa ([1996]) argues that simple inhibition-based models cannot account for the fact that Mach bands do not seem to be present in scenes involving sharp luminance steps, rather than gradual luminance ramps. He surveys a range of alternative, non-inhibition based models that aim to capture Mach bands and related psychophysical phenomena. The issue, then, may be more complex and unsettled than the standard account makes it appear.

Still, if lateral inhibition turns out not to be the main process that produces Mach band perception, the points to be made in this discussion still stand. First, we would still have common lateral inhibition mechanisms across species. This kind of processing unit itself is multiply realizable. So there is still a question about how the same functional unit can be realized in different kinds of neural mechanism. Second, whatever mechanisms do account for the perception of Mach bands across species will presumably also be ones for which the question of multiple realizability occurs. *Mutatis mutandis*, the same questions can be asked about them.

Comparing the mammalian retina with the horseshoe crab retina serves to make several points. First, both sorts of creatures display a common psychophysical effect; second, they possess substantially different underlying neurobiological mechanisms; third, these mechanisms seem to implement the same sort of functional property, namely lateral inhibition; fourth, according to the standard account this latter property is what explains the presence of the psychophysical effect. So what we seem to have is two structures satisfying the functional specification of eyes that have psychophysical commonalities but different underlying neural realizations. That is, we have what appears to be a case of actual multiple realization.

While these are not exhaustive of the possible kinds of eyes, they illustrate the point that distinct kinds at one level can have numerous other common causal properties, and therefore constitute a higher-level kind. And they share these properties despite a number of other seemingly significant differences between them. Thus Shapiro’s inference from different causal properties to the absence of a higher-level kind is not generally sound. Whether a functionally defined property also constitutes a kind is something to be decided on a case-by-case basis.

3 Unity through constraint?

Interestingly, Shapiro considers the case of lateral inhibition, but draws the opposite conclusion from it (Shapiro [2004], pp. 117-20). Rather than concluding that the use of this common strategy of visual processing in different species shows its multiple realizability, he suggests that it shows that the evolution of visual systems occurs under tight constraints. These constraints mean that there will likely be only one (or at most a few) ways of building an organic system that can process visual images. And this is potentially in tension with MR, which allows that there should be a wide diversity of evolved mechanisms for each functional capacity.

Lateral inhibition serves a useful function in vision: it sharpens contrasts and aids in discrimination of closely spaced stimuli. Undoubtedly this accounts for its recurrence in evolution. But this recurrence does not undermine MR, since it's also clear from the preceding example that different species carry out lateral inhibition using physiologically distinct mechanisms. Artificial visual systems give further examples of this diversity. There are a number of silicon-based implementations of fragments of early vision that specifically include lateral inhibition among adjacent units (Mead and Mahowald [1988]; Wolpert and Micheli-Tzanakou [1993]). These silicon retinas have electrical response profiles that are similar to the neural response profiles seen in biological retinas. But the mechanisms are distinct: the parts are hard-wired integrated circuit arrays that comprise a much simpler sheet-like structure than the multilayered vertebrate retina, and which contains different components carrying out different activities. For instance, there are no neurotransmitters in a silicon retina, and hence it supports different styles of intervention—electrical stimulation affects integrated circuits differently than it does horizontal cells, and compounds that block neurotransmitter release have no effect at all.

Indeed, lateral inhibition occurs in multiple sensory modalities within single species. Human vision, touch, and audition all contain mechanisms of lateral inhibition (Békésy [1967]).

But these mechanisms are also distinct. In the case of somatosensory stimulation, lateral inhibition is implemented by the activity of GABA-ergic inhibitory interneurons in the dorsal column nucleus (Mariño, Martinez, and Candeo [1999]; Willis and Coggeshall [2003], Ch. 12). These neurons are morphologically and physiologically different from those that carry out lateral inhibition in the mammalian retina. The same sort of case that the comparative physiological data make for multiple realization can very plausibly also be made intraspecifically using cross-sensory physiological data.

What this suggests is that there is a common *functional* characteristic that recurs across different species (and artificial systems) as well as within individual species. But the presence of a particular functional characteristic does not necessarily entail the presence of any particular physical mechanism. That isn’t to say that there might not be physical constraints on how nervous systems must be constructed if they’re going to realize terrestrial psychological capacities. Many of these constraints, however, are themselves organizational or functional, and these can’t be assumed to be physically identical in all cases; at least not unless we simply adopt a taxonomy of physical mechanisms that co-classifies as physically similar anything that satisfies the relevant functional specification. But, apart from the fact that this flies in the face of the independently justified neurobiological taxonomy of those mechanisms, this is just to beg the question against the MR thesis.

Shapiro’s other examples of constraints on the realization of cognition support this point. For instance, he argues that humanlike cognition requires sensory systems that transduce information into usable neural signals, receptors that tile sensory surfaces in varying densities, topographic maps in primary sensory areas, and broadly modular organization in the brain itself (Shapiro [2004], pp. 105-38). But it is clear that there are at best rather abstract similarities among, say, the diverse photoreceptors in the retina and the various tactile, thermal, and chemical receptors that mediate touch. Even if they possess broadly similar receptive field organization and project to topologically organized regions of primary sensory cortex, the fine-grained detail of these neural structures will differ, just as the neural mechanisms of lateral inhibition differ across sensory systems.

So one can agree with Shapiro that there are constraints on constructing terrestrial psychologies, but also maintain that these constraints are mainly functional. Since it is possible to build many different kinds of neural mechanisms within these constraints, this is compatible with multiple realizability.

Shapiro might argue that although the examples canvassed here do provide some evidence of multiply realized properties, this is not the norm in terrestrial psychology. Multiple realization might be existent, but *rare*. This challenge is impossible to answer without surveying many more examples drawn from different domains. But keep in mind that Shapiro thinks that such features as topographic maps of sensory spaces will be widespread, since they are likely to be produced by the general constraints on evolved neural systems. If is true, as I've argued, that being a topographic map is a multiply realizable property, then MR will be as widespread as maps are. And the same goes for his other examples of design under constraint, many of which are standard fare in comparative neurophysiology. So if the arguments given here are correct, MR *will* turn out to be close to the norm in terrestrial cognitive systems. Deciding this question will ultimately have to wait for more fully developed models of such systems, however.

4 Justifying taxonomic mismatches

It might seem, however, that the view developed so far runs afoul of an argument advanced by Bechtel and Mundale ([1999]). They suggest that the MR thesis only appears plausible because we allow Ψ to be individuated coarsely and the various Φs to be individuated finely. In the classic case of pain, we lump together human and octopus pain as being the same sort of phenomenal state, but then focus on the numerous differences between the neurobiological characteristics of the two species. Against this approach, though, they say: ‘if the grain size is kept constant, then the claim that psychological states are in fact multiply realized looks far less plausible. One can adopt either a coarse or a fine grain, but as long as one uses a comparable grain on both the brain and mind side, the mapping between them will be correspondingly systematic’ (Bechtel and Mundale [1999], p. 202).

It’s true that one can produce the appearance of MR by altering grain sizes in a somewhat arbitrary fashion. But grain sizes in and of themselves are not the main issue when considering whether we have a case of MR. For one can have a coarse/coarse grain scheme, and still have MR if there are many possible coarsely individuated lower-level realizing mechanisms available. And the same goes for a fine/fine grain scheme. What is wanted is a justification for adopting not merely a grain-mismatched taxonomy, but what I will call an *unaligned* taxonomy. Two taxonomies are aligned when there is a systematic, one-to-one mapping from higher level taxa onto lower level taxa, and unaligned otherwise. Unaligned taxonomies correspond to the classic picture of MR, on which a single upper level category may be mapped onto many lower level ones. So put, what is needed is a *principled* case for adopting different higher and lower taxonomic schemes in some particular contexts; that is, for employing unaligned taxonomies.

Whether we should adopt an unaligned taxonomy depends on whether the various lower level properties are independently certified as being theoretically ‘interesting’ in the relevant domain. Suppose we decide that *Limulus* eyes and cat eyes do not really both perform lateral inhibition—that there is not the same property instantiated in both organisms, just as we might decide that humans and octopi don't really instantiate the same property *pain*. Rather, there are two different properties, inhibition$_1$ and inhibition$_2$, picking out two different (in fine-grained terms) functional capacities had by creatures with correspondingly different visual systems. These capacities might differ in any number of ways: their spatial and temporal response properties, the strength of their inhibition, the circumstances that trigger them or inhibit them, etc. Focusing on these differences might lead us to say that the visual systems of different creatures do not really do the *same* things.

Of course, anything that performs inhibition$_{1/2}$ performs inhibition simpliciter, since these capacities are related as determinate and determinable. But it's important to note that inhibition simpliciter (the determinable property) and inhibition$_{1/2}$ (the determinate properties) play different explanatory roles. Explaining why a system experiences Mach bands may involve mentioning the fact that it implements lateral inhibition. And mentioning lateral inhibition might be relevant if we don't know the mechanism for producing Mach bands or if there are several different types of mechanisms that could produce them and we want to know which one is involved in this particular case. Inhibition simpliciter is a generic sort of capacity that can be realized in many more precise forms, all of which produce this psychophysical phenomenon. But mentioning these precise forms isn't necessary for the explanation. If we posit a generic sort of inhibitory mechanism without specifying the precise, fine-grained characteristics it has, we can still explain how Mach bands arise. This is a case where we have a coarse-coarse alignment between higher and lower taxonomy.

Explaining why the organism experiences bands with *these* precise characteristics, though, may involve adverting to inhibition$_{1/2}$, since those constitute the particular psychophysical capacities the system has. The fine-grained details of the psychophysics presumably derive from the details of the inhibition functions themselves. And explaining *how* a system instantiates inhibition$_{1/2}$ may involve referring to the particular fine-grained neurobiological mechanism at work. The neural mechanism in question produces precisely the right sort of functional capacity, which in turn gives rise to the precise psychophysical profile of the organism. These cases all involve matching fine-grained taxonomies at lower and higher levels, and hence are also cases of aligned taxonomies.

But, importantly, explaining how a system instantiates inhibition simpliciter might involve drawing attention *either* to the particular fine-grained mechanism at work in that organism, *or* to the range of possible mechanisms that can bring about that sort of capacity. The latter case involves a taxonomic mismatch of the sort that Bechtel and Mundale warn against—that is, a one-many mapping between higher and lower taxa. But adopting this mixed taxonomy is anything but a mistake. It can prove heuristically indispensible, for example, once one has characterized the general function of a cell type or brain region to then propose a range of possible lower level mechanisms that might realize that function, then proceed to rule them out on the basis of side effects, predicted responses to various interventions, predicted anatomical consequences, and so on. Craver ([2007]) notes that in giving mechanistic explanations we frequently are in the position of giving numerous competing ‘how-possibly’ accounts, all of which describe mechanisms that would be sufficient to produce the phenomenon in question, but which we don’t have the evidence to decide among. When a range of how-possibly mechanistic accounts is on the table to explain a single phenomenon, we are explicitly taking seriously the possibility of multiple realization.

Unaligned taxonomies of this sort, then, can serve a crucial heuristic role in discovering mechanisms. It is almost never true that we know there to be only one way of building a structure that has a particular function. Rather, given a certain function, we can imagine many possible ways to build something that does that. And since this kind of dual-level inquiry works with both fine grained and more broad taxonomies simultaneously, the MR advocate can adopt such a mixed taxonomy without its being *ad hoc*. (See Section 6 for more discussion of how coarse-grained higher level taxonomies can be useful in describing the structure of multilevel mechanisms.)

5 Further conditions on multiple realization

Recently, Polger ([2009]) and Shapiro ([forthcoming]) have articulated some further conditions on when we have genuine evidence for MR. Their conditions focus on when putative multiple realizers are sufficiently diverse.

Shapiro makes the following proposal: ‘DD: The differences in realizers that are relevant to MR should not be differences that cause only differences in the realized properties’ ([forthcoming], p. 9). That is, where two realizers are different this may cause differences in the realized property itself, but must also have other consequences. His example involves two watches, one that keeps time slowly and another that runs fast. Both watches contain a spring that is part of the timekeeping mechanism, but its length is slightly different in each watch. The difference in the length of the spring accounts for the difference in their timekeeping properties.

This difference, being *only* relevant to whether the watch is slow or fast in its timekeeping, is not enough to make the realizers really different. So the slow and fast watches do not count as multiple realizations. Comparing these spring-governed watches with a digital watch makes this clear. The digital watch contains a substantially different sort of mechanism from the mechanical watches with their springs and gears. Genuine MR depends on realizing kinds being different in a variety of respects beyond what they contribute to instantiation of the realized property itself.

Polger's conditions are expressed differently, but make essentially the same point. He says that to conclude that two entities A and B multiply realize a property, we must have evidence that they are:

> (i) classified commonly by system S1, and (ii) classified distinctly by system S2. In addition, (iii) the facts 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. (Polger [2009], p. 464)

Here S1 and S2 are taxonomic systems for classifying entities. We can think of these as picking out families of properties at some level of organization. MR requires that A and B be S1-similar and S2-different. Polger's condition (iv) is intended to capture the idea that realizers cannot be trivial variations of one another. So where A and B both display similar versions of a realized property—for example, inhibition$_1$ and inhibition$_2$—the realizing (S2) properties that each one possesses must do more than produce just the differences in the realized (S1) properties themselves. That is, A and B must be 'more different' at the level of the realizers than they are at the level of the realized properties. Shapiro's slow and fast watches illustrate how this condition can fail to be met: their low-level differences only produce the minor variations in their timekeeping properties, but have no wider ramifications for how they function.

Shapiro's DD and Polger's (i)-(iv) aim to make the same point. Two realizations of the same higher level property must be different in ways that go beyond just making the higher-level property slightly different in each instance. In short, the realizers must be *really* different in each case for us to have a case of genuine MR. Call this the Really Different Realizers (RDR) principle.

The RDR principle is meant to rule out arbitrarily adopting a taxonomic scheme according to which we have different realizers wherever two instances of some high-level property differ slightly in some way. The example of the slow and fast watches is just such a case. Polger also argues that cases of neural plasticity and reorganization of sensory cortex doesn't count as multiple realization, since the lower level properties in question are not distinct enough. In a series of widely discussed experiments by Sharma et al. ([2000]), ferrets had the regions of their brains that would have developed into a normal auditory cortex rewired so that they received input from the visual system. Surprisingly, these ferrets are eventually able to see, albeit less well than normal ferrets can. Auditory cortex is sufficiently plastic that it can assume the function of the visual cortex when given the right input. But this doesn't support MR, since when a ferret has its auditory cortex rewired so that it serves the function of enabling it to see, the rewired cortical region behaves, Polger claims, just like the normal ferret's visual cortex would have. There are no interesting lower-level differences between the rewired auditory cortex and the normal visual cortex, so these are not two Really Different Realizers of the same higher-level property.

One might think that RDR poses a threat to the lateral inhibition case. Comparing the Limulus retina, the vertebrate retina, and the various silicon retinas, all of which implement lateral inhibition, we can see obvious high-level similarities that speak in favor of classifying them together. Consider, for instance, the broadly similar response profiles they give to contrast stimuli. Moreover, where these high-level properties are slightly different from one realization to another, these differences are often explainable in terms of the lower-level mechanisms that implement inhibition in each system. So slight variations in the way inhibition is carried out can be traced to differences in the underlying mechanism. These cases might be seen as analogous to the example of the slow and fast watches, differing only in ways that bear on the particular shape of the inhibition function.

But the comparison doesn't hold, since these systems *also* differ with respect to any number of properties that have nothing to do with the differences in how they implement lateral inhibition per se. This can be seen by looking at the diversity of cell types, neurotransmitters, kinds of synapse (electrical vs. chemical), etc., as well as the different spatial and temporal organization of the parts and their activities. These realizing properties give rise to a range of system- and species-specific characteristics beyond merely the differences in the realized property itself. The differences in the biological retinas are already more substantial than those that would be produced by taking a single part and modifying it slightly—entirely new entities and activities are involved across species. These differences become still more extreme when we consider retinas that are built from integrated circuits.

Perhaps none of these mechanisms are as different from each other as digital watches are from mechanical watches, but all that is needed is that there be more dissimilarity than in the simple case of the slow and fast watch. Extending the analogy, we might think of them as being like mechanical watches put together from different parts—gears, springs, levers, wheels, numbers of jewels, and so on of different sizes, compositions, and shapes in entirely different spatial and temporal arrangements. So the RDR condition is met in the present case.

6 Explanatory taxonomy and the special sciences

Recall that according to the Shapiro-Kim dilemma, having diverse causally relevant properties precludes a group of entities from being a multiply realized kind. The case of Mach band experience across diverse visual systems is intended to raise doubts about this claim. What it shows is that diverse lower-level mechanisms can converge on common functional traits at various levels of organization. Moreover, the assumption that nature can be organized in this way plays a key heuristic role in discovering mechanisms. Now I will lay out one way in which these functional groupings might be seen as kinds.

On Shapiro's view, special science categories have a fundamentally taxonomic function: they 'collect and order the domain of a special science in a way that facilitates its investigation' ([2000], p. 654). Functional concepts fix a range of 'analytic' truths about things that fall under them. This is what backs Shapiro's claim that the only generalization about eyes as such is the analytic one that eyes are for seeing. But this simple dichotomy between empirical and analytic generalizations doesn't hold. 'Analytic' links can give rise to empirical generalizations.

For instance, though it might be analytic that eyes are for seeing, it's surely not analytic that seeing involves receiving ambient light or other radiation. In fact, some ancient theories of vision denied this, advocating instead the view that vision involves extramission, or the emitting of light (or some sort of ray) from the eye itself as part of the causal process of seeing. This mistaken belief may even be found among educated contemporary adults (Winer, Cottrell, Gregg, Fournier, and Bica [2002]). The fact that seeing requires taking in ambient light is a general empirical fact about seeing; indeed, it was a significant discovery (Lindberg [1976]). So it is a true empirical generalization about eyes that they take in ambient light. Even if this is only *partially* empirical (because it relies on an analytic premise), this is enough to make it a non-analytic truth.

Allegedly analytic generalizations run into further problems. For example, what counts as an eye is not always clear. In their comprehensive book on animal eyes, Land and Nilsson ([2002]) conclude that not all behaviorally relevant light-sensitive organs are properly eyes. Rather, they claim, eyes are devices that contribute to the construction of spatial images. Earthworms have simple light sensitive organs and hence can ‘see’ in some sense, but this sensitivity is limited. They are unable to use information about differences in light intensity coming from different directions. What matters to whether to something is an eye is that it can use this information about differences in light intensity to produce a spatial image of the environment. The concept of an eye as used in comparative vision science, then, is not an analytic or *a priori* deliverance, but one constructed by looking at salient differences in the organs of different species. So even analytic generalizations may rely on comparative empirical work that requires drawing lines to separate interestingly different sorts of structures—in this case, simple light sensors versus proper eyes.

It’s unclear, then, that the analytic/empirical distinction Shapiro relies on can bear much weight. This complaint aside, it is a further problem that not all functional concepts pick out categories that are equally interesting. ‘Things that can be knocked over with a feather’ is a functional category, but I doubt whether it is apt for a special science. Shapiro’s view leaves us with no way of explaining why there should be such a difference, if the categories that special sciences construct are not themselves kinds but rather mere taxonomic or organizational conveniences. What is the *point* of grouping things together by function if the grouping itself does not pick out an inductively potent kind?

The difference between trivial and non-trivial functional categories matters to theory construction because discovering the *right* functional components out of which to build an organism's control systems is itself a seriously challenging task. Much of the work of building theories, models, and simulations in the biological, behavioral, and psychological sciences involves finding the appropriate concepts to use in analyzing a system. On the view I propose, functional categories are kinds when they are appropriate and useful for constructing explanations of how a system comes to exercise particular psychological and behavioral capacities.

Consider central pattern generators (CPGs) as an example. CPGs are units that produce regular oscillations endogenously or in response to input. There are many different ways to assemble such circuits. They can be built out of multi-neuron arrays of varying sizes using inhibitory interneurons, or out of local dendrodendritic connections. These structures differ in their size, location, temporal characteristics, and many other physical/neural properties. But at a functional level they are all recognizable as CPGs. Wherever some rhythmic bodily activity needs to be produced (walking, swimming, digesting), CPGs are a useful device to have in one's engineering toolkit.

Clearly, the concept of a CPG is a functionally defined concept. It is also, by the arguments of the preceding sections, in principle multiply realizable. And even CPGs within particular species can be multiply realized. Prinz, Bucher, and Marder ([2004]) ran a series of two million computational simulations of the central pattern generator that produces pyloric rhythms in the stomatogastric ganglion of the lobster. In each simulation they varied the number, types, and strengths of the particular neural connections in the three-neuron network. Fully 20% of the simulated networks produced patterns of activation that were broadly similar to those produced in actual lobster ganglion networks. On narrower criteria of matching experimentally observed burst patterns, 11% of these 20% (452,516 networks in all) produced appropriate output patterns. So even on relatively fine-grained criteria for input-output equivalence, there are over 450,000 ways of constructing a network of cells that produces the right pyloric rhythms for governing digestion in lobsters.

In an experiment reminiscent of the familiar ‘silicon brain’ thought experiment, Szücs et al. ([2000]) replaced a single neuron in an extracted pyloric ganglion network with an electronic neuron that models various observed behaviors (e.g., chaotic vs. regular burst patterns) of biological neurons. When the electronic neuron was inserted in place of one of the original biological neurons, the network could be induced to produce the same overall burst pattern. Hybrid silicon-biological networks such as this simple one provide another empirical demonstration of how these functional components can be realized by a variety of mechanisms.

Despite being multiply realizable, CPGs have *explanatory value*: positing such circuits illuminates how certain kinds of behavior and observed neural activity might take place. These abstract structures play an explanatory role in answering these questions independent of information about their realizers. Knowing that an organism’s control systems contain a CPG in a certain location helps to explain certain of its capacities Examples are the aforementioned stomatogastric ganglion of the lobster, which regulates digestion, and control of swimming in the lamprey (Grillner, Wallén, and Brodin [1991]). Both of these crucially depend on CPGs. The CPG that governs lamprey swimming, like the pyloric network of the lobster, has also been shown to be realizable by a variety of underlying neural networks. Patel, Murray, and Hallam ([2005]) used genetic algorithms to evolve a range of distinct connectionist systems that produce the same burst patterns as empirically observed in biological networks. In virtue of producing certain sorts of effects, functional units such as CPGs can be situated within a larger system of control structures, which in turn explain the capacities in question.

Understanding how an organism possesses a range of capacities depends on seeing its inner organization as containing such units. The same is true of lateral inhibition, which is an abstract mechanism for producing a range of effects in sensory processing. So we can ask: how does the lamprey swim? how does the lobster digest? how do humans walk with a regular gait? I suggest that, in the behavioral sciences, it is by recurrently playing this sort of role in abstract control systems that functional categories earn their status as kinds. A functional grouping that can be used to produce a range of useful effects in constructing a complex control mechanism is a *functional kind* with respect to the behavioral sciences.

This view of functional kinds turns on a certain idea about what these special sciences aim to explain and how these explanations proceed. The central idea is that we aim to explain the possession of various capacities by a creature: the capacity to make visual discriminations, to fly, to add numbers, to find its way back to the nest after wandering in the featureless desert, etc. We explain these capacities by constructing models of the internal control structures that could produce them. These models typically describe the creature as containing a variety of systems, structures, processes, resources, and sources of activity that interact in a predictable way to produce certain effects under a certain range of conditions. The capacity in question is explained by the presence of these systems operating under certain well-specified conditions. This picture of psychological explanation as focusing on capacities and proceeding by structural decomposition is defended by Craver ([2006]), Cummins ([2000]), Glennan ([2005]), Love and Gureckis ([2007]), and Piccinini ([forthcoming]), among others.

The interesting functionally defined categories, then, constitute recurrent building blocks of cognitive systems. They explain the possession of various capacities of those systems without reference to specific realizing structures. They may in turn be explained by the presence of further functional units at lower levels, or by physical mechanisms. One major task in understanding cognition is finding the right decomposition of a system into abstract control units and constituents. Logic, computation theory, cybernetics and control theory, dynamical systems theory, and neural network theory provide examples of how accounts of such control units might be developed. And insofar as such functional categories can usefully be applied to modeling cognition, they count as kinds.

From this perspective, debates between, for instance, symbolic versus connectionist approaches to cognitive modeling constitute debates about what psychological kinds there are. To the extent that an empirically well-validated model of cognition as a whole (or some cognitive capacity in particular) makes use of the apparatus of connectionist networks, these networks and their constituents (units, weights, etc.) are psychological kinds. It might be that networks are only useful for explaining some such capacities—perhaps only perceptual pattern matching and certain aspects of skill learning, for instance. Or they might be more widely applied to language processing and acquisition, higher reasoning, etc. (McClelland and Rogers [2003]). The greater the range of capacities that can be explained by appeal to a set of functionally defined structures like a connectionist network, the more explanatorily useful such networks become.

This is a notion of kindhood that is predicated on relational characteristics of functionally grouped entities; roughly, how widely they can be used in our best explanatory models. Moreover, it constitutes a significant shift of emphasis from other notions of kinds. It's normally thought that a class inches towards kind status to the extent that it is mentioned in a range of empirical generalizations; in particular, as it supports a greater number of inductive inferences. As Millikan puts it, 'a natural kind corresponds not just to a projectable predicate, but must figure as the subject of many empirical generalizations' ([1999], p. 48). On this account, a group K is a kind to the extent that there is a true set of claims such as: things that are K are F, G, H, etc., where F, G, H, etc. are empirically discovered properties that 'cluster' with being K. The more such properties that cluster around K, the more of a kind K becomes.

On the present account, a class becomes a kind not in virtue of directly participating in empirical generalizations—not in virtue of there being many empirical truths about Ks as such—but rather in virtue of playing a role in a range of models of cognitive and behavioral capacities. Whether K is a kind is a matter of how important K is in these models, plus how widely applicable such models turn out to be (how widely these models can be projected across distinct types of cognizer). These factors, possibly weighted in some fashion, determine the 'kind score' for a functional category. So, for instance, if central pattern generators of various sorts are important parts of our best-confirmed models of the behavior of a range of organisms, then CPGs count as a kind for the purposes of behavioral explanation. Similarly for lateral inhibition mechanisms. In both of these cases, we have good reasons to think that many cognitive capacities will be built out of such parts. Lateral inhibition is re-used in many sensory modalities in many different organisms, and CPGs are part of many of our best accounts of motor behavior across species (and in many robotic systems—see Ijspeert [2008]). These components can be ‘plugged in’ to many different systems to explain a variety of different capacities.

To take another example, consider a cognitive resource such as a short-term memory buffer. Numerous models of human and animal cognition posit such structures; moreover, they play an important role in the design of manufactured digital computers. Being a memory buffer is multiply realizable—cyclical firing activity in a neural population or a refreshing a pattern in a set of electrical registers can both maintain information for a short period of time. Despite any differences that might attach to these distinct realizations, memory buffers are a kind for the purpose of computational psychology insofar as there are empirically well-supported models of cognition and behavior that involve them. Yet it is doubtful that there are many general inductive regularities of the form: all short-term memory buffers are F, G, H, etc. Abstractly defined functional categories like memory stores earn their credentials by participating in a range of models that are themselves empirically validated.

A final example from comparative psychology concerns kinds of representations rather than resources (e.g., memory buffers) or mechanisms (e.g., CPGs). Human infants as well as many nonhuman species have the ability to rapidly compare and make judgments about numerical quantities across sensory modalities (Gallistel [1990]; Gallistel and Gelman [2000]). According to many researchers, this ability relies on the presence of an abstract analog magnitude scale, sometimes called an ‘accumulator’. This representation tends to be more accurate for small quantities, but less accurate as quantities increase in size. So rats required to press a button a certain number of times will tend to hit the right number of presses within a tighter range for smaller numbers than for higher numbers. An analog scale also cannot easily represent fine distinctions between quantities. So infants habituated to displays containing a certain number of dots will dishabituate to displays containing significantly more or fewer dots (8 vs. 16), but will fail to treat displays with a similar number of dots (8 vs. 10) as novel (Xu and Spelke [2000]; Xu [2003]). This suggests that they can handle approximate quantities but not precise numbers.

Many models of numerical competence hypothesize that these capacities involve operations over an analog representational scale. These operations appear to obey Weber's law, which states roughly that the discriminability of two numbers is a function of their ratio. If these models prove to be correct, it would show that part of the basic representational toolkit for some human and animal cognitive capacities includes analog magnitude scales, no matter how these representations are physically realized. Representations, resources, processors, and the mechanisms built out of them, then, are all potentially multiply realizable kinds.

The key shift in emphasis here is from requiring categories to play a role in lawlike empirical generalizations to requiring that they play a role in well-supported models. The sort of models that are especially relevant are models that aim to explain capacities by giving a functional decomposition of some portion of the creature that possesses those capacities. This shift in what constitutes a kind meshes particularly well with the recent emphasis on mechanistic and model-based explanation as opposed to nomic or law-based explanation in the sciences (Bechtel [2008]; Craver [2007]; Cummins [2000]; Piccinini [forthcoming]). If the normal form of psychological and neurobiological explanation involves positing mechanisms and offering models of their internal structure in order to explain an organism's various capacities, then a notion of kindhood that is tied to the inductive behavior of a category with respect to laws or other regularities seems inappropriate, since there are not likely to be any *laws* involving these kinds. Rather, there are models that crucially rely on them to produce the necessary effects.

The notion of multiple realizability has traditionally been developed in the context of the nomic-causal framework of explanation in the special sciences, in which kinds were identified with groupings that participated in natural laws. I have laid out a contrasting notion of kinds that identifies them with functional components of explanatory models. One might object that this amounts to changing the subject of the debate. But there is no necessary link between debates over MR and the nomic conception of kinds (or the allied notion of explanation as essentially involving laws). The nomic conception of scientific explanation that held sway for many decades has been, if not replaced, then at least supplemented by a family of model-based conceptions akin to the one adopted here. In these conceptions, explanation involves modeling the functional structures that interact to produce cognition and behavior. It has been claimed by many—including Shapiro and Kim—that these functional categories cannot be kinds. Part of the purpose of this discussion is to argue to the contrary, and thereby to vindicate this widely successful explanatory strategy. If kinds are fundamentally the real explanatory structures at work in a certain domain, the functional categories posited by well-confirmed models should qualify.

By the same token, multiple realizability seems to be a hallmark of the special sciences. Accordingly, the defense of functional kinds goes hand in hand with a defense of their multiple realizability. The philosophical significance of such a defense is precisely the same as in the case of the older, nomic conception of kinds and explanation: to show how special science properties and entities can be constructed from physically and causally diverse underlying mechanisms. This motivation remains unchanged even though our conception of kinds in the special sciences might change.

Two points should be made about this conception of functional kinds. First, it is a graded matter whether a functional category is a kind. How significant a role must a category play in a model to count as a kind? How many capacities must it underpin? How widely applicable must the model be? These questions do not have general answers. But this does not represent a change from the standard conception of kindhood, since exactly the same questions can be asked about natural kinds. How many empirical generalizations must a category participate in? How important do these need to be? How widely projectable must these generalizations be? (Millikan even suggests that Californians may qualify as a sort of rough kind if the criteria are loosened sufficiently!) This isn't to say that these questions are simple to answer, only that the questions are not ones that are unique to the notion of functional kinds developed here.

One might object at this point that a functional kind as described here is in fact just a sort of complex causal kind. To be able to satisfy a functional specification is to have at least a certain set of causal capacities, since this is how functions themselves are described. And if this is so, can't the Shapiro-Kim dilemma just be pressed against this notion as well? Such an objection would have to turn on the implausibility of there being kind-making commonalities among causally diverse mechanisms. But if these kind-making commonalities are supposed to involve those mechanisms participating in a range of causal *laws*, I have already conceded that this is not likely. This is part of what motivates the shift to the model-based notion of kinds. The sort of causal relations that qualify a grouping as a kind, on this view, are relations that enable a category to play a recurrently useful role in a range of models. This involves—indeed, requires—having certain complex causal powers. But diverse kinds of lower-level mechanisms can overlap in the causal powers that are required to fulfill the functional specification of the components of cognitive models. This is what the extended example of lateral inhibition (Section 2) and the discussion of functional constraints on cognitive design (Section 3) are intended to help establish. So the Shapiro-Kim objection doesn't seem obviously applicable once we adopt the view of kinds on offer here.

Second, we should keep in mind that capacities are typically explained by *multilevel* mechanisms. There may be an initial high-level decomposition of a system into components such as a boxes-and-arrows flowchart diagram, which specifies only the functions of the various subcomponents and their interconnections. A further decomposition might specify the precise internal functioning of these subsystems and give a more fine-grained analysis of their communications and control connections. And so on—this process of explaining a capacity by positing a set of connected subsystems, then giving a further explanation of the functioning of those subsystems, can be carried out until one reaches the level at which one gives a specification of the particular physical structures that realize the system as a whole.

Indeed, it is a bit misleading to talk about a strict functional/physical contrast at all. A physical system can be described with varying degrees of abstraction. Each one of these abstract descriptions picks out what I have been calling an abstract structure of some sort. A description of how a system carries out a certain function is abstract to the extent that it omits certain details about precisely how this function is executed. These details may be filled in by further descriptions that are less abstract, including by descriptions of physical components and their organization. There is not just one 'functional level', but rather a whole host of intermediate structures at varying degrees of abstraction from the underlying physical components (Craver [2009]; Lycan [1981]). And there are bound to be co-evolutionary relationships among descriptions of the same system at different levels—lower-level discoveries can affect higher-level taxonomies and vice-versa (Keeley [2000]).

What the existence of a hierarchy of functional levels implies is that there will be a hierarchy of kinds in psychology, from very abstract kinds of systems to completely specified blueprints for how those systems are constructed from their most basic elements. This in turn has consequences for the multiple realization of the system as a whole. If we view the system from the perspective of the most fine-grained functional specification, it is comparatively unlikely that it will be multiply realized, since the lowest level of functional abstraction may be highly constrained in terms of the physical structures that realize it. Viewing the system from the perspective of higher levels of functional abstraction will loosen these constraints and allow more latitude for multiple realization. This practice of successive abstraction further illustrates the legitimacy of taxonomic mismatches such as those that Bechtel and Mundale ([1999]) seemed to rule out (see Section 4).

But even if the system as a whole, described at the lowest level of functional specificity, is not multiply realizable, there may still be functional kinds at this level. For instance, it might turn out that the most basic functional elements in the system are the equivalent of elementary logic gates. Even if the way that these gates are assembled to form the basic wiring diagram of the system is so precisely specified that there are very few ways (or only one way) to build a physical device that conforms to this plan, the logic gates themselves (AND-gate, NOR-gate, etc.) are the recurrent building blocks that are responsible for the functioning of every capacity exercised by the system. They therefore play a crucial role in our models that aim to explain these capacities, and so should count as kinds.

The relationship between kindhood and levels of functional abstraction is a complex one. What this example illustrates is that, since explanations of capacities will usually be multilevel, and since multilevel explanations will involve many layers of abstraction, there will be a variety of functional kinds corresponding to the types of mechanisms that can be built from the basic components at each of these levels. These (relatively) basic components will in turn be decomposed into complexes at the next level; where this process terminates depends on when, in the present context, we have decided that the capacity in question is sufficiently well understood for our purposes.

7 Conclusions

Obviously there is more to say about the nature of functional kinds and the subject matter of the special sciences more generally. I have dealt here only with functionally defined special science properties. For all I have said, there may be multiply realizable properties that are not functional properties. If there are, they may require a different sort of explanation than the one given here (Gillett [2007]; Pereboom [2002]). But this restricted focus is justified in light of the fact that much of the discussion of multiple realizability has focused on functionalism and functional properties. This also explains the decision to focus on psychological and neuroscientific kinds, since these are the ones most frequently defended by functionalists and attacked by their opponents.

There are two morals of this discussion. First, there are legitimate instances of multiple realization of functionally specified properties, and these violate no strictures on acceptable scientific taxonomy (for further examples, see Aizawa and Gillett [2009]; Couch [forthcoming]; Keeley [2000]). Second, despite the causal and structural diversity in these realizers, there is a sense in which they constitute kinds, at least with respect to the behavioral sciences. This latter point is what is truly at issue between Shapiro and Kim and their critics. Developing a notion of kinds that allows them to be multiply realizable is a crucial part of explaining how the properties described by the special sciences find their place in the world. I take it as a hopeful sign that this account also seems to mesh with explanatory practice in many parts of psychology and neuroscience, where explanations often involve constructing abstract models of the functional structures that underlie various capacities of organisms. If the practice of science is implicitly committed to multiple realization, the metaphysics of science should follow its lead.

Daniel A. Weiskopf

Department of Philosophy

Georgia State University

P.O. Box 4089

Atlanta, GA 30302-4089

dweiskopf@gsu.edu

Acknowledgements

Thanks to Bill Bechtel, Gualtiero Piccinini, Tom Polger, Larry Shapiro, three anonymous reviewers for this journal, and audiences at the Austrian Ludwig Wittgenstein Society (Kirchberg, Austria, 2008) and the Society for Philosophy and Psychology (Bloomington, IN, 2009) for helpful discussion and comments.

References

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

Battelle, B.-A. [2006]: ‘The eyes of *Limulus polyphemus* (Xiphosura, Chelicerata) and their afferent and efferent projections’, *Arthropod Structure and Development*, **35**, pp. 261-274.

Bechtel, W. [2008]: *Mental Mechanisms: Philosophical Perspectives on Cognitive Neuroscience*. New York: Routledge.

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

Békésy, G. [1967]: ‘Mach band type lateral inhibition in different sense organs’, *Journal of General Physiology*, **50**, pp. 519-532.

Boyd, R. [1991]: ‘Realism, anti-foundationalism, and the enthusiasm for natural kinds’, *Philosophical Studies*, **61**, pp. 127-148.

Couch, M. B. [forthcoming]: ‘Multiple realization in comparative perspective’, *Biology and Philosophy*.

Craver, C. [2006]: ‘When mechanistic models explain’, *Synthese*, **153**, pp. 355-376.

Craver, C. [2007]: *Explaining the Brain*. Oxford: Oxford University Press.

Craver, C. [2009]: ‘Mechanisms and natural kinds’, *Philosophical Psychology*, **22**, pp. 575-594.

Cummins, R. [2000]: '"How does it work?" vs. "What are the laws?"', in F. Keil and R. Wilson (*eds*), *Explanation and Cognition*, Cambridge: MIT Press, pp. 117-45.

Fehrenbach, W. H. [1985]: 'Anatomical circuitry of lateral inhibition in the eye of the horseshoe crab, *Limulus Polyphemus*', *Proceedings of the Royal Society of London, Series B,* **225**, pp. 219-49.

Field, G. D., and Chichilnisky, E. J. [2007]: 'Information processing in the primate retina: Circuitry and coding', *Annual Review of Neuroscience,* **30**, pp. 1-30.

Gallistel, C. R. [1990]: *The Organization of Learning*. Cambridge: MIT Press.

Gallistel, C. R., and Gelman, R. [2000]: 'Non-verbal numerical cognition: From reals to integers', *Trends in Cognitive Science,* **4**, pp. 59-65.

Gillett , C. [2007]: 'A mechanist manifesto for the philosophy of mind: The third way for functionalists', *Journal of Philosophical Research,* **32**, pp. 21-42.

Glennan, S. [2005]: 'Modeling mechanisms', *Studies in History and Philosophy of Biology and Biomedical Sciences,* **36**, pp. 443-464.

Grillner, S., Wallén, P., and Brodin, L. [1991]: 'Neuronal network generating locomotor behavior in lamprey: Circuitry, transmitters, membrane properties, and simulation', *Annual Review of Neuroscience,* **14**, pp. 169-99.

Hartline, H. K., and Ratliff, F. [1957]: 'Inhibitory interaction of receptor units in the eye of *Limulus*', *Journal of General Physiology,* **40**, pp. 357-376.

Heil, J. [2003]: ‘Multiply realized properties’, in S. Walter and H.-D. Heckmann (*eds*), *Physicalism and Mental Causation*, Exeter, UK: Imprint Academic, pp. 11-30.

Ijspeert, A. J. [2008]: ‘Central pattern generators for locomotion control in animals and robots: A review’, *Neural Networks,* **21**, pp. 641-53.

Keeley, B. L. [2000]: ‘Shocking lessons from electric fish’, *Philosophy of Science,* **67**, pp. 444-65.

Kim, J. [1992]: ‘Multiple realization and the metaphysics of reduction’, *Philosophy and Phenomenological Research,* **52**, pp. 1-26.

Lindberg, D. C. [1976]: *Theories of Vision from al-Kindi to Kepler*. Chicago: University of Chicago Press.

Love, B. C., and Gureckis, T. M. [2007]: ‘Models in search of a brain’, *Cognitive, Affective, and Behavioral Neuroscience,* **7**, pp. 90-108.

Lycan, W. G. [1981]: ‘Form, function, and feel’, *Journal of Philosophy*, **78**, pp. 24-50.

McClelland, J. L., and Rogers, T. T. [2003]: ‘The parallel distributed processing approach to semantic cognition’, *Nature Reviews Neuroscience,* **4**, pp. 310-22.

Mangel, S. C. [1991]: ‘Analysis of the horizontal cell contribution to the receptive field surround of ganglion structures in the rabbit retina’, *Journal of Physiology,* **442**, pp. 211-34.

Mariño, J., Martinez, L., and Candeo A. [1999]: ‘Sensorimotor integration at the dorsal column nuclei’, *News in Physiological Science,* **14**, pp. 231-37.

Mead, C. A., and Mahawold, M. A. [1988]: ‘A silicon model of early visual processing’, *Neural Networks*, **1**, pp. 91-7.

Millikan, R. [1999]: ‘Historical kinds and the “special sciences”’, *Philosophical Studies*, **95**, 45-65.

Moini, A. [2000]: *Vision Chips*. Dordrecht: Kluwer.

Palmer, S. E. [1999]: *Vision Science: From Photons to Phenomenology*. Cambridge, MA: MIT Press.

Patel, L. N., Murray, A., and Hallam, J. [2005]: ‘Increased swimming control with evolved lamprey CPG controllers’, *IEEE International Joint Conference on Neural Networks*, pp. 2195-200.

Pereboom, D. [2002]: ‘Robust nonreductive materialism’, *Journal of Philosophy*, **99**, pp. 499-531.

Pessoa, L. [1996]: ‘Mach bands: How many models are possible? Recent experimental findings and modeling attempts’, *Vision Research*, **36**, pp. 3205-3227.

Piccinini, G. [forthcoming]: ‘The mind as neural software? Understanding functionalism, computationalism, and computational functionalism’, *Philosophy and Phenomenological Research*.

Polger, T. [2009]: ‘Evaluating the evidence for multiple realization’, *Synthese*, **167**, pp. 457-72.

Prinz, A. A., Bucher, D., and Marder, E. [2004]: Similar network activity from disparate circuit parameters. *Nature Neuroscience*, 7, 1345-52.

Ratliff, F. [1965]: *Mach Bands: Quantitative Studies on Neural Networks in the Retina*. San Francisco: Holden-Day.

Shapiro, L. [2000]: ‘Multiple realizations’, *Journal of Philosophy*, **97**, 635-54.

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

Shapiro, L. [forthcoming]: ‘How to test for multiple realization’, *Philosophy of Science*.

Sharma, J., Angelucci, A., and Sur, M. [2000]: ‘Induction of visual orientation modules in auditory cortex’, *Nature*, **404**, pp. 841-7.

Shoemaker, S. [2003]: ‘Causality and properties’, in S. Shoemaker, *Identity, Cause, and Mind* (2nd ed.), Oxford: Oxford University Press, pp. 206-33.

Sterling, P. [1998]: ‘Retina’, in G. Shepherd (*ed.*), *The Synaptic Organization of the Brain* (4th ed.), Oxford: Oxford University Press, pp. 205-54.

Willis, W. D., and Coggeshall, R. E. [2003]: *Sensory Mechanisms of the Spinal Cord: Volume 2: Ascending Sensory Tracts and Their Descending Control* (3rd ed.).

Winer, G. A., Cottrell, J. E., Gregg, V., Fournier, J. S., and Bica, L. A. [2002]: ‘Fundamentally misunderstanding visual perception’, *American Psychologist*, **57**, pp. 417-24.

Wolpert, S. and Micheli-Tzanakou, E. [1993]: ‘Silicon models of lateral inhibition’, *IEEE Transactions on Neural Networks*, **4**, pp. 955-61.

Xu, F. [2003]: ‘Numerosity discrimination in infants: Evidence for two systems of representation’, *Cognition*, **89**, B15-B25.

Xu, F., and Spelke, E. [2000]: ‘Large number discrimination in 6-month-old infants’, *Cognition*, **74**, B1-B11.