Penultimate version
To appear in P. Thagard (editor)
Handbook of the Philosophy of Psychology
and Cognitive Science
Kluwer, 2006

Realization*

Robert A. Wilson
Carl F. Craver

---

1. A Servant of Two Masters

For the greater part of the last 50 years, it has been common for philosophers of mind and cognitive scientists to invoke the notion of realization in discussing the relationship between the mind and the brain. In traditional philosophy of mind, mental states are said to be realized, instantiated, or implemented in brain states. Artificial intelligence is sometimes described as the attempt either to model or to actually construct systems that realize some of the same psychological abilities that we and other living creatures possess. The claim that specific psychological

* The authors would like to acknowledge the Society for Philosophy and Psychology for providing a venue (their 2001 meeting in Cincinnati) for the start of their collaboration. Rob Wilson's work on this paper has been supported by SSHRC Individual Grant 410-2001-0061 from the Social Sciences and Humanities Research Council of Canada.

capacities, such as the capacity to understand spoken language, might be realized by different individuals in different ways, has been presupposed by psychologists and cognitive neuroscientists, who design clever experiments in which measures of dependent variables— reaction times, error rates, or localized metabolic activity, for example— provide evidence about what the neurological realizers for specific psychological abilities are.

As common and as it is to speak of realization and realizability, these notions have only recently been scrutinized under the philosophical microscope. Much of this work has been critical. Some have identified putative problems with standard views of realization (Wilson 2001, 2004; Gillett 2002); others have challenged the widespread commitment to the thesis that the mental is multiply realized in the physical or biological (Bechtel and Mundale 1999; Shapiro 2000, 2004). But along with the critiques, some positive views have emerged, and we now have a better understanding of several of the desiderata that any view of realization must satisfy.

One thing that has already become apparent is that the concept of realization serves two different masters. A caricature of the two will help to distinguish two different sets of desiderata on accounts of realization.

On the one hand, there is the *Metaphysician of Mind*, concerned primarily with, to use C.S. Lewis's phrase, the place of mind in the world order. With the exception of the recently burgeoning area of consciousness studies, analytic philosophy of mind has been predominantly physicalist (in one sense or the other) since the classic statements of the mind-brain identity theory by J.J.C. Smart and U.T. Place in the 1950s. But this broad commitment to physicalism does not take the Metaphysician of Mind as far down the path to understanding the mind as one might hope, and has given rise to its own problems. Consider several of these. Is the mind strictly identical to the brain, or should physicalists endorse some other kind of relationship between the two? Wherever one stands on this issue, what room does this leave for the reality of the mind, for genuine *mental causation*, for an understanding of consciousness and content, or for distinctive aspects of our grasp of the mental, such as first-person knowledge?

In serving this master, the concept of realization has been slotted into a particular network of technical concepts, such as supervenience, metaphysical sufficiency, and nomic necessity. That network of concepts is also partly constituted by a range of by now familiar "isms"--physicalism, functionalism, computationalism, and reductionism being the four most common. The Metaphysician of Mind uses conceptual analysis as a way of exploring entailments and tensions between positions that one might adopt using those concepts. This approach is exemplified by the work of Jaegwon Kim. Best known for his sustained work on supervenience and mind over the past 30 years, Kim has more recently turned his attention to realization in extending his critique of non-reductionist views of the mind.

But the Metaphysician of Mind is not the only master that the concept of realization must serve. It must also serve the *Cognitive Scientist*. For the Cognitive Scientist, it is not the place of the concept of realization amongst any network of concepts that needs to be understood, but how specific psychological functions and capacities are or can be realized by particular psychological and neurological structures and mechanisms. These realizations are best explored through the construction of models or schemata of the corresponding processes, models or schemata that can be specified at various degrees of abstraction. These schemata range from the very abstract functional decompositions ("boxology", to its detractors) that one finds in much of cognitive psychology and artificial intelligence, to biochemically detailed accounts of specific neural pathways. For the Cognitive Scientist, the realization of the mental is to be investigated through such familiar strategies as localization of function and physical decomposition. Such strategies start with the cognitive capacities or behaviors of an organism (plus or minus a bit) and then proceed to explain them in terms of the capacities or behaviors of its parts (e.g., parts of the central nervous system). This is how, for example, cognitive neuroscientists and psychologists investigate short-term memory or visual shape recognition.

In saying that the concept of realization must serve the Cognitive Scientist, we are helping ourselves to an assumption that is widespread among Metaphysicians of Mind and philosophers of cognitive science. This assumption is that although cognitive scientists seldom use the term "realization", much of what they say and do can be reconstructed with the help of this term. Psychologists and cognitive neuroscientists are much more likely to talk of the *neural correlates of*, the *neural substrates of*, the *neural mechanisms for*, or the *implementation of* psychological capacities than to talk of how those capacities are "realized" by specific mechanisms and processes. As Jaegwon Kim says, acknowledging that the meaning of "realize" has not been fully explained in the philosophical literature in which it features, "you will not go far astray if you read 'P realizes M' as 'P is a neural substrate, or correlate, of M" (1996: 102, fn.4). So the idea is that these ways of talking about the relationship between the psychological and the physical, prevalent in the discourse of scientists themselves, can be viewed as less metaphysically committed invocations of the concept of realization.

The slack between talk of neural correlates and mechanisms within the cognitive sciences and philosophical talk of realization deserves at least a brief comment. The short answer to the question "How do cognitive scientists conceptualize realization?" is: "They don't". Or, to put it more accurately: they do, but almost exclusively when they are attempting to sketch a broader location for their particular views, or when interacting directly with their philosophical interlocutors. Given that, the assumption that we can gloss the scientific talk of the physical correlates of, the underlying mechanisms for, or localized structures causally implicated in, the operation of some particular psychological capacity in terms of the notion of realization may not be cost-free. It may require that we adjust our view of what appeals to "realization" can do, or leave us with a philosophically emaciated view of realization (cf. Polger 2004).

It is our view that realization is and will continue to be, a servant to these two masters, and that this is not altogether a bad thing. But two masters they are, and each make distinct demands on the concepts that serve them. Consider the following list of desiderata that would surely make any Metaphysician of Mind happy. They would like a view of realization that:

- elucidates the relationship between functionalism, physicalism, and reductionism
- enables us to at least clarify what mental causation and multiple realizability involve (if not tell us whether they occur in the domain of cognition)
- points the way to an explanation of how organized networks of neurons give rise to the full range of mental phenomena

They might, of course, want more, but this would be at least a start.

Contrast these desiderata with those that a Cognitive Scientist could not just live with but live for. They would like a view of the realization of psychological capacity, P, in neural structures N, that:

- identifies the N that are relevant to the operation of P
- reveals what variation exists in the relationship between P and N across different populations
- provides a step-by-step account of just how P is realized by N in any particular instance

Again, as a Cognitive Scientist, one might expect more—one might like an account of realization to guide the design of experiments, or to allow a clear distinction between adequate and inadequate realizing explanations. But it seems that this would be a minimal list of desiderata for any Cognitive Scientist.

The trivial point about these two sets of desiderata is that they are different. Less trivially and more speculatively, projects that draw on a concept satisfying one set of desiderata may do little to satisfy the other. For example, the experiments that show that the hippocampus is involved in episodic memory will likely tell us little about how physicalism, functionalism, and reductionism are related, and will leave us none the wiser about mental causation or multiple realization in general. Conversely, suppose that we had a "deep explanation" of just why neural networks give rise to particular mental phenomena. Even such an apparent breakthrough for the Metaphysician of Mind may rely on or lead to a concept of realization that leaves us with a blank stare when it comes to providing a step by step account of just how auditory hallucinations are produced in cases of delusional schizophrenia.

More pessimistically (and even more speculatively), it is possible that the concept of realization satisfying one of these sets of desiderata must be different from the concept of realization that satisfies the other. While a certain kind of peaceful coexistence could persist were the preceding scenario to eventuate, to find that there was some kind of deep incompatibility between the desiderata of the Metaphysician of Mind and of the Cognitive Scientist would be a sort of intellectual disaster. So much so, we think, that one of the desiderata on the list of each should be that their view of realization should be at least consistent with (ideally, well-integrated with) that of their counterpart.

2. Realization and the Metaphysics of Mind

To get some sense of the role that appeals to realization have played in the philosophy of mind and cognitive science, and of how realization has come to serve two masters, let's look at how and why talk of realization was initially introduced.

Hilary Putnam brought the concept of realization into contemporary philosophy of mind in 1960 in his classic paper "Minds and Machines". In that paper Putnam described the relationship between the mental and the physical as one of realization. In doing so, Putnam drew an analogy between minds and machines. In particular, he argued that the relationship that holds between minds and brains is the same as that holding between abstract Turing machines and the physical arrangements of matter in which they are instantiated or implemented: realization.

Part of Putnam's point was to dissolve the mind-body problem. He claimed that the relationship between mind and body, or mind and brain, should be no more puzzling—indeed, no more interesting—than that between the abstract states of a given Turing machine and the structural states of the device realizing it.

Putnam's original introduction of "realization" also formed a part of the cognitive revolution that spawned the cognitive sciences as we know (and love) them. In doing so, it also provided a way of thinking about the relationship between the mental and the physical taken up within psychology, linguistics, and computer science themselves. At least that is part of philosophical lore. The idea was that even if we did not hear the term "realization" in the mouths of those working within these disciplines, their exploration of the mechanisms underlying psychological functions, capacities, and abilities could be adequately glossed in terms of more metaphysical-sounding notions, such as realization. As we said in the first section, talk of the neural correlates of, or of the neural mechanisms for, a given psychological capacity have been viewed as loose science-speak for something like the relation of realization.

Within a dozen years or so of its introduction, realization came to be seen as useful in articulating three of the closely related "isms" that had, by that time, gained currency: physicalism, functionalism, and computationalism.

First, as a metaphysical relation weaker than identity, realization was thought suitable for developing a brand of physicalism that made room for the autonomy of psychology, for genuine mental causation, and for psychological laws. And, unlike identity, as an asymmetrical relation, realization was also thought well suited to capturing the dependence of the mental on the physical, or the *determinative* nature of underlying, physical states.

Second, as functionalism came to replace type-type identity views as the theory of choice for materialists in the philosophy of mind, the physical was seen as realizing the mental, rather than being strictly identical to it. Functionalism is the view that psychological states are to be identified with the causal or functional role they play in the overall causal network of psychological states. We identify psychological states in terms of their functional roles much as one might identify a part of a machine in terms of the causal role that it plays in the operation of the machine, or as one might identify a job in an organization, such as a manager, in terms of what the person in that position contributes to the overall set of tasks that the organization undertakes. In all of these cases, there is a distinction to be drawn between the role itself and the occupant of that role. In the psychological case, brain states and mechanisms are taken as the occupants of the functional roles defining psychological states. The concept of realization was drawn on to tighten this intuitive notion of an occupant.

If we extend these analogies, then it is seemingly easy to infer that the functionalist view of the mind is committed to the idea that mental states are *multiply realizable* in physical states. For just as different persons can fill one and the same role within an organization, and parts of a machine can be replaced by other parts, sometimes parts quite different in many of their properties, so too in principle can psychological states be realized by very different physical states. We have multiple realization just when we have the *same kind* of psychological state realized by physical states that are *different in kind* from one another.

Reflecting the proclivity of philosophers for cutting edge technology, a favorite example used to illustrate this idea was the mousetrap. What mousetraps *are*, what they are to be strictly identified as, is a kind of device that is designed to catch mice. It has that function, and a mousetrap can be thought of, perhaps arcanely, as an input-output device that accepts live or unconfined mice as inputs and delivers dead or confined mice as its output. But as every cat knows, there are many ways to catch a mouse. (Correlatively, as every mouse knows, there are many ways to skin a cat.) Mousetraps are realizable by different kinds of physical devices: traditional neck-snappers, balance cages, and pit traps, for example.

Third, functionalism and the appeal to (multiple) realizability sat well with the rise of the computational metaphor and the nascent cognitive sciences to which that metaphor is central. As Putnam's original appeal to Turing machines suggested, minds are to brains as programs are to the hardware that runs them. Computationalism moves beyond this sort of appeal to the computational metaphor in providing a more precise way of characterizing the functional roles that define psychological states. Computationalism takes these roles to be specifiable as *algorithms*, i.e., as effective procedures that can be realized in principle, and often in practice, by machines.

Turing machine functionalism, the earliest and in many respects the most problematic form of computationalism, held that these algorithms were computable by a universal Turing machine— not an actual machine but a "theoretical machine" conceptualized by the mathematician Alan Turing in his well-known work on the foundations of logic. A Turing machine is a simply structured device that can perform several mechanical functions: it can scan a symbol, write a symbol to a specific location, or erase a symbol from that location, and move its scanner from one location to the next. Although it is common in the philosophy of mind and cognitive science to talk of mental states as being realizable by particular Turing machines, note how quickly this leads to confusion, given that Turing machines are themselves characterized purely functionally.

Computationalism has been central to classic, early work in artificial intelligence, to more recent work in the computational modeling developed by connectionists, and to those using Bayesian and statistical techniques to understand psychological functions and capacities. *Strong* artificial intelligence can be characterized as the view that appropriately designed computational programs would not simply be models of but realizations of certain psychological capacities, such as the capacity to understand (fragments of) natural language. Given that the computer hardware that realizes such programs are of a distinct physical kind from the brainware that realizes these very same programs in us, strong AI entails the multiple realizability of the mental in the physical.

In this section we have concentrated on the development of the concept of realization from the point of view of the Metaphysician of Mind, albeit one who also likes to think of himself as allied with the Cognitive Scientist. We now turn our attention to the view of realization implicit in the work of the Cognitive Scientist.

3. Realization and the Cognitive Sciences

The Cognitive Scientist, as noted above, rarely makes the notion of realization explicit. Instead, the idea of realization is implicit in the practice of explaining by decomposing and in a variety of experimental techniques used to test claims to the effect that some N realizes some P. We can learn something about this implicit notion of realization by looking at an exemplar case: the voltage-gated sodium channel. We can then turn our attention to a higher-level cognitive case.

Our first example suggests that the explanatory practices of science embody several different varieties of realization. The primary division is between the material realization of an entity and the explanatory realization of a property. A secondary division, holding among explanatory realization relations, turns on whether what is getting realized is a property or an activity and on the relevant kind of organization in the realizer. The example will help to make these distinctions transparent.

Figure 1 Near Here

Suppose that we want to understand how neurons generate action potentials. Action potentials are electrical waves that propagate along the axons of neurons and are generated near the cell body (in a structure known as the axon hillock) in response to summing excitation from neighboring cells. In their resting state, neurons have a voltage gradient across their membranes, and an action potential (as shown in Figure 1) is a fluctuation in that gradient that propagates from the hillock down the length of the axon. The explanation for the action potential involves breaking a neuron and its membrane into parts and showing how their activities are organized together to produce the action potential.

The action potential is realized by the cooperative activities of a variety of ion channels in the membrane. One of these channels, the one responsible for the initial rising phase (I) of Figure 1, is the voltage-gated sodium channel. (Detailed description of the sodium channel can be found in most neuroscience texts [e.g., Kandel and Schwartz 2001]). One hoping to understand how the sodium channel is responsible for the rising phase of the action potential will similarly appeal to its parts and their cooperative activities. But as we look more closely at this exemplar of explanation in neuroscience, it quickly becomes clear that there are several varieties of realization in play. These differences reflect the fact that different kinds of things are getting realized and that different kinds of things are doing the realizing.

One part of understanding the action potential will involve understanding the material composition of the sodium channel. This membrane-spanning channel is a protein, and so it is made of amino acids. If we ask what realizes the sodium channel, we might appeal to different parts— subunits of the channel protein, or stretches of the sequence within those subunits, or the atoms bound into amino acids— but in each case we are appealing to the material parts composing the sodium channel. And this is the case whenever we look for the realizer of an entity or entity kind.

Sometimes entity kinds have been the Metaphysician's primary examples. They have appealed to jade, water, or corkscrews as realized items (see Shapiro 2000, 2004). And when we talk of entity kinds being realized, we tend to point to the material realizers— the units of material (however those are to be individuated) located within the entity's spatial boundaries. Jade (or more specifically, nephrite) is realized by a silicate of calcium, magnesium, and iron. Corkscrews are realized by steel. Computers are realized by silicon. When we say such things, we are talking about the realization of an entity as an entity (a spatially bounded object) rather than as some particular kind of entity (i.e., under some description).

For example, we neglect the organization of components that gives the sodium channel its shape and unique activities, and pay attention instead only to the matter in its boundaries.

Learning how entities are materially realized is often an important descriptive stage of science. Learning the primary sequence of the sodium channel, for example, allows one to begin to sort its various orders of structure, from the sequence (order) of amino acids, beginning to end, to how these give rise to local spatial forms (e.g., helices and sheets) and subsequently to more complicated folding arrangements that constitute the whole protein. Cell fractionation and centrifugation were essential to the development of biochemistry and molecular biology precisely because they allowed investigators to discover the constituents of cells. Learning the material realizers of the sodium channel has been similarly important to learning how it works and why it has the properties that it does.

The task of specifying a material realizer for an entity is complete once one has listed (or specified) its constituents exhaustively. Such completeness is a descriptive success. When attention shifts from describing to explaining, however, entities and their material realizers are no longer the primary focus. Instead, attention shifts to properties and activities as realized kinds and, correlatively, to component parts, their properties, their activities and their organization. Explanatory forms of realization are not just exhaustive lists of material constituents, but selective descriptions of the relevant parts for some explanatory purpose. The explanation neglects some properties of the material realizers and accentuates others. For explanatory realization (as opposed to material realization), it frequently does matter which parts one attends to: arbitrary parts will not be explanatorily relevant. We can distinguish three varieties of this explanatory form of realization: realization of an aggregate, realization of a structural property, and realization of an activity. Consider these in turn.

The wave form of the action potential is not the product of a single sodium channel but of thousands of channels in the cell membrane. The axon hillock is especially dense in such channels, and when large numbers of channels activate at once, they give rise to the wave form for the cell as a whole. This shift from single channels to populations of channels is often simply presumed in an explanation (note that this step is typically left out of standard textbooks) because the total current produced by opening the sodium channels in the hillock is approximately a sum of the currents flowing through individual sodium channels. The total current is an example of what Wimsatt (1997) has called an *aggregative property*. The mass of a pile of sand is realized aggregatively by the masses of the individual grains. Other subunits, such as half-grains of sand, would work just as well. Likewise, the total sodium current is a sum of the currents through individual sodium channels. In contrast to material realizers, aggregate realizers are selective in that some properties of the realizers are important for the explanation and some are not. It is the mass of the individual grains (and not their color, melting point, or texture) that fully constitutes the mass of the whole pile. In purely aggregative cases of realization, it makes sense to speak of realized properties as being composed of realizing properties: the realized property and the realizing properties are different values of the same variable, and the value of the realized variable is exhausted by the sum of the values for the same variable for the individual realizers. Unlike material realization, however, cases of aggregate realization are selective: the realizers include only some of the properties of the constituents (just those involved in the sum). Explaining the mass of the pile of sand appeals only to the masses of the components, and not to their color, for example. And understanding the current through the population of sodium channels involves calculating only the values of the currents through the individual channels (or at least so it is often presumed in explanations).

Truly aggregative properties are difficult to come by. Indeed, the opening of different sodium channels is not, properly speaking, aggregative, since the influx of sodium gradually changes the charge of the cell and the relative concentrations of sodium inside and outside of the cell, each of which is relevant to the rate of ionic flow through the channel. Nonetheless, some explanations in the sciences of the mind do depend on properties that are closer to the aggregate end on a spectrum of organization, and they represent an important type of realization in these sciences. For example, population effects of neurotransmitters at the synapse and sums of excitation to a node in a connectionist network each approximate this ideal.

Sometimes it is a structural property that calls out for explanation. One may want to understand, for example, how the sodium receptor has its characteristic sequence of amino acids (an explanation fleshed out by appeal to translation and transcription), how it forms a channel through the membrane, or how the channel selectively allows (primarily) for the flow of sodium ions. Of particular explanatory interest in the effort to understand the action potential is the sequence of amino acids in each of the repeating subregions of the channel known as the S4 regions. The amino acids in this region are ordered such that every third amino acid residue is an arginine or a lysine. Given the charges on these residues, this linear order of components produces a helical structure with evenly spaced positive charges. Leaving aside for the moment the significance of this arrangement, one might want to understand why that type of amino acid sequence produces an alpha helix. For this, one will appeal not merely to the material constituents of the helix, but to the sequence of amino acids and their polar and nonpolar (hydrophilic and hydrophobic) properties. It is these that determine how the molecule folds and coils in water. Structural forms of realization thus lay out not only the (relevant) material components, but also the features that determine their overall shape and configuration: not just the matter but the spatial forms of organization.

The details of molecular folding have been worked out for relatively simple structures such as the alpha helix, but are only partially worked out for even moderately complex forms of tertiary structures. As in the case of aggregative realization, there is a straightforward sense in which the spatial properties of the parts are part of the structure of the whole. Unlike cases of aggregative realization, however, the lower-level properties are not summed, but may involve interaction and organization of the components (as the evenly spaced charges participate in S4's helical structure).

Now suppose that the realized kind is not merely an aggregate or structural property of the sodium channel but an *activity*: something that the sodium channel does. Suppose, for example, that we want to understand how the sodium channel is activated by the summing excitation of cell body. As with the above discussion of the alpha helix, the explanatory task selectively directs attention preferentially to some parts of the receptor and away from others. The helix is significant because of its putative role as an activating gate for the sodium channel. In cases of *mechanistic realization*, the burden of realization is borne by some constituents (working parts or components) more than others, and organization among the components figures increasingly in our description of the realizer. Let us continue with our example of the sodium channel.

Under the resting electrical conditions, a positive extracellular potential holds the alpha helix— lined with positive charges— in a stable position in the membrane. Weakening that positive potential, as occurs when the cell is depolarized from its resting state, allows the helix to rotate out toward the extracellular side. This rotation is thought to occur in each of the sodium channel's subunits which destabilizes the balance of forces holding the channel in its closed state and allows it to move to a new equilibrium state, one that has an open channel through the membrane. There is, as of yet, no complete story to tell about how the displacement of the helix, coupled with the attraction and repulsion among component polar and nonpolar amino acids, alters the conformation of the entire protein (the mathematics required to predict how such a change would ramify through the molecule is daunting). But we nonetheless know the component processes and can see how those properties and activities could, if organized appropriately, change the conformation of the channels. Mathematical models are often used to simulate the folding in such complex arrangements, thereby showing how complex structures can be realized by such simple component forces and sequential organization and how they can be altered by such factors as voltage changes in the cell.

Mechanistic realizers are composed of the working parts of the mechanism, such as amino acids, membranes and ions and their activities, the things that these entities do: their repelling, rotating, opening. The components are organized such that they exhibit the behavior of the mechanism as a whole (Machamer et al. 2000). There are many varieties of organization in mechanisms. Spatial properties of the parts (e.g., their size, shape, and orientation) or spatial relationships among the parts (e.g., their positions, compartmentalization, fit, motion) are as important for mechanistic realization as they are for structural realization. But mechanistic realizers also have a temporal component: the activities in the mechanism have characteristic orders, rates and durations. And, finally, the parts of this mechanism act and interact with one another such that they exhibit the behavior of the mechanism as a whole. Mechanisms, in other words, also have a causal component to their organization.

Descriptions of mechanistic realizers are selective. There are no mechanisms simpliciter; all mechanisms are mechanisms *of* something. It is by reference to this behavior of the mechanism as a whole that the relevance of components is established. In describing the activation of the sodium channel as we just have, we selectively attend to the alpha helix and selectively neglect those aspects of the receptor responsible for the ion selectivity, or the channel's inactivation, or the binding cites for neurotoxins. Material realization includes all of these parts; mechanistic realization includes only the relevant ones. The description of a mechanistic realizer is considered complete when all of the relevant components have been included such that it is possible to describe its working in terms of intelligible activities from beginning to end, without gaps or promissory notes. (See Wilson and Keil [1998] on the surprising shallowness of our grasp of everyday mechanistic explanations.)

The search for mechanistic realization is embodied in many of the techniques of cognitive neuroscience. Lesion and stimulation experiments, reaction time studies, PET and fMRI experiments are all designed to tease apart the mechanistic realizer for cognitive phenomena. In "bottom up" experiments, one intervenes to remove or stimulate a component and monitors for changes in the behavior of the mechanism as a whole. In "top down" experiments, one manipulates the behavior of the mechanism as a whole and monitors changes in the states or activities of putative components. The goal is to show that an item has a function within the mechanism and to characterize what, precisely, that function is.

Perhaps the most plausible way to develop the link between the search for mechanisms in neuroscience and the notion of realization at play in the metaphysics of mind is through attention to the notion of a *function*. We have already briefly discussed functionalism in the philosophy of mind, where "functional role" and "causal role" are often used interchangeably. But there is a more constrained notion of function that has been used by biologists, one that links functions to mechanisms, and it is this notion that is particularly relevant to linking mechanisms to realization. Roughly for now, the function of X is what X does or is supposed to do in the mechanism in which X operates. On this view, the function of X is not the total causal role that X plays but some more selective subset of the causes and effects of X: those relevant for the explanatory purposes at hand. One describes X's causal role when it is possible to show how X is organized (spatially, temporally and actively) into a higher level mechanism. Only a subset of its total causal role will be directly relevant to that higher level mechanism.

Moreover, functions are frequently *hierarchically realized* in that higher level functions are achieved through the performance of sub-functions. To take a stock example, one of the functions that the circulatory system performs is to move blood around the body, and one of the functions necessary for this is for there to be a source of force for moving the blood. The parts of the heart, their activities, and their organization together constitute the mechanism for this sub-function, even though there are many other inputs to and outputs from the heart (e.g., it makes noises, produces heat, etc.).

This appeal to biological systems, their functions, and the mechanisms that realize them has a natural application to cognition. Cognitive systems are composed of mechanisms that perform specific functions, and those mechanisms in turn are composed of further mechanisms with even more specific functions, together constituting a hierarchy of mechanisms. For example, the mammalian visual system is composed of a series of mechanisms—the retina, the superior colliculi, the lateral geniculate nuclei, primary visual cortex (V_1), and the extrastriate areas (such as $V_{2\text{-}4}$ and MT). Each of these has its particular components that perform specific functions, and in turn they are chunked together to form larger functional units.

As we noted at the start of this chapter, scientific notions of realization are implicit in the practice of explaining by decomposition. But that practice involves different kinds of realization. This difference in kind tends to track different criteria for assessing the success of the description of a realizer. Descriptions of material realizers are complete when they include all of the material constituents of the realized entity. For aggregative properties, it should be demonstrable that the realized property is a sum of lower level properties. For structural properties, determinants of shape and position are of primary importance. And finally in cases of mechanistic realization for activities, the goal is to provide a complete description, without gaps, from the beginning of the mechanism to the end, exhibiting the relevant entities, properties and activities, and showing how they are organized together within the realized activity.

4. A Working Account of Realization

One way to proceed in articulating the concept of realization further would be to seek to provide a traditional philosophical analysis of that concept, a set of individually necessary conditions that are collectively sufficient for anything to be a realization. Because we are deeply skeptical of the likelihood of success of such an approach (here, as elsewhere), we offer instead the following working account of realization. This working account serves both to articulate some commonalities that are shared by Metaphysicians of Mind and Cognitive Scientists and to sharpen some of the differences that separate them.

Let's begin with a simple canonical statement of the realization relationship:

(R_0) An object O's having property or activity A is realized by O's having property or activity B.

We have stated R_0 as a relationship between properties, since this is a common way of proceeding in the literature. For material realization, R_0 would be a relationship between objects, and for mechanistic realization, the relata would be an activity and its mechanism. R_0 describes an ontological relationship: it is a relation between O's having A and O's having B rather than between statements about O, P and Q, or between theories or explanations involving O, A, and B. O's having B is the realizer, and O's having A is what's getting realized. O is an object, in the minimal sense that it is a bearer of properties, an agent of activities, or a relatum in relations that constitute A and B.

R_0 is typically taken to be *asymmetric*: O's having B realizes O's having A and not the other way around. This restriction helps to distinguish realization from identity. This asymmetry can be understood as a dependency relation: O's having B is *sufficient but not necessary* for O's having A. To accommodate the possibility of multiple realization, both the metaphysician and scientist will want to allow that O's having B is unnecessary for its having A; A could be realized by different objects and different properties. On this much, it seems to us that the Scientist and the Metaphysician can agree. Let's consider some candidates for strengthening the realization relation, some of the motivations for so strengthening it, and some of the limitations of doing so.

4.1 Must Realization Be Decompositional?

The most common scientific conception of realization— discussed above in association with explanation by analysis into parts—tends to be *decompositional.* In cases of decompositional realization, a property of the whole is realized by the properties of its parts:

(R_1) O's having P is realized by parts of O having Q

For example, the mousetrap's behavior of catching mice (or its capacity to catch mice, P) is realized by the parts of the trap (a trigger, a latch, an impact bar) and their organized activities (Q). The realized property is a property of the whole, and the realizing properties are the intrinsic and relational properties of the parts. By way of contrast, if one were to say that the mouse trap's being red is realized by its being some determinate shade of red (e.g., scarlet), this would not be a decompositional realization relation. Similarly, David Marr's (1982) now-famous computational, algorithmic and implementation levels are not decompositional. Algorithms are not parts of computations but instead are different descriptions of one and the same thing. In these last two examples, it is not the parts of O that do the realizing, but rather a different property of O as a whole.

Metaphysicians, who associate realization with discussions of physicalism and functionalism, may not be concerned with understanding the relationship between a whole and its parts *simpliciter* but rather with the relationship between the whole and the parts *plus their organization.* They are asking: is there anything more to O and its properties than the parts of O, their properties, and their organization? Spooky varieties of emergence and vitalism are spooky *not* because they insist that the whole is greater than the sum of its parts— even the most ardent physicalist will grant that the organization of the parts realizes properties of the whole that would not be realized if the parts were organized differently—but because they insist that the whole is greater than the parts plus their organization together. Because many metaphysicians who are interested in mind-body relations are interested in whether the mind is anything "over and above" the organized activities of neurons (and molecules, brain regions, etc.), they need a notion of realization that frames that question.

Along these lines, Kim has suggested that realization is a relationship not between "micro-macro" levels but between orders. Kim thinks of levels as sorted by a mereological relation: things at lower levels are parts of the things at higher levels. He calls this hierarchy of levels the "micro-macro hierarchy." Orders, however, are sorted by a dependency relationship: things have their higher-order properties in virtue of having their lower-order properties. Kim stresses that levels and orders should not be confused:

> Notice the following important fact about this [realization] hierarchy: *this hierarchy does not parallel the micro-macro hierarchy*— to put it another way, *the realization relation does not track the micro-macro relation*. The reason is simple: *both second-order properties and their first-order realizers are properties of the same entities and systems*. (1999, 82)

Realization, on this view, is like Marr's levels or like the relationship between O's being red and O's being scarlet. This view can also be extended to the mousetrap. The mousetrap has the ability to catch mice because it has the property of being an organized collection of triggers, springs, and levers (and perhaps a mouse). Having the ability to trap mice and being an organized collection of triggers, springs and levers, on this view, are not properties at different levels but properties of different orders: the trap has the ability to catch mice in virtue of the organized relations among its parts. The description of the parts plus organization and the description of the mousetrap are descriptions of one and the same thing, and so the realization relation, on this view, is not decompositional. It is not an interlevel relationship.

Scientists, unlike metaphysicians, spend little time worrying about whether wholes are greater than their parts plus their organization. It is a working assumption that the wholes just are the parts organized together. When they say that "the whole is greater than the sum of the parts," they should be taken as literally talking about sums of parts. When the whole is greater than the sum, this is because the parts are organized in a particular way, they interact with one another, and they do things together that they could not do alone. But these assertions are entirely consistent with the idea that wholes just are the parts and their organization. Perhaps empirical investigation will shake the scientist from this working assumption, as the periodic revival of forms of holism and vitalism in the history of biology and neuroscience might attest. But such revivals are not taken especially seriously by those engaged in the task of seeking lower-level explanations.

The lack of any lasting success for vitalism or emergent properties in the past and the corresponding successes of sciences driven by the search for lower-level explanations are for many convincing reasons to ignore holism and vitalism and to continue to seek such explanations. Scientists spend more time worrying about whether they have identified the right parts and the right aspects of their organization to fully explain O's having P in terms of the parts of O having Q. Failure to discover such an explanation is only rarely accompanied by calls to abandon this physicalist working hypothesis and much more commonly accompanied by calls to do more experiments.

Be that as it may, this issue points to an important pragmatic difference in the uses to which scientists and metaphysicians are putting the concept of realization. The metaphysician is concerned with evaluating the status of physicalism while scientists largely presume some form of physicalism. The latter are concerned with questions like: "What sort of realizer could realize a property like that?;" "What parts are included in this realizer, and which parts are irrelevant?;" "How are these parts organized together such that the realize that property?" These latter questions are more likely to be addressed by attention to varieties of realization and criteria of adequacy for describing and evaluating hypothesized realizers (see section 3) than they are by attention to global metaphysical principles. On the issue of decompositionality, then, it may turn out that the metaphysician and the scientist do not (and perhaps need not) agree.

4.2 Must Realization be Constitutive or Intrinsic?

A second possible restriction on the realization relation would be to make it constitutive— that is, to require that the realizer of O's having P be wholly contained within O's spatial boundaries. This is the view that Wilson (2001, 2004) calls the *constitutivity thesis* about realization, a thesis that is widely shared amongst Metaphysicians of Mind. The problem with this constraint is that many of the activities in which objects engage cannot be realized by the parts of those objects alone in that these parts—along with their properties and how they are organized—are not metaphysically sufficient for many of the activities in which those entities engage. In such cases, it is plausible to argue that if realizers are to be metaphysically sufficient for the properties and activities they realize, then they must extend beyond the boundary of O and beyond the relationships that exist among the things inside O. The constitutivity thesis would rule out such appeals to contextual realizers, and might be formulated in the following putative constraint:

> (R_2) In C, O's having P is realized by the intrinsic physical properties Q of O and its parts.

To see why R_2 is controversial, consider some of the functions that might be explored in mechanistic realization. We might describe the trap as firing, or as catching mice, or as protecting the cookie jar. We might describe the heart as pumping blood, as circulating blood, or as delivering the poison. If we say that the trap is catching mice or protecting the cookie jar, then we are referring to the world beyond the trap. Its function is then world-involving in a way that suggests that the realization of that function (as sufficient for "catching mice" or "protecting the cookie jar") must also extend beyond the trap or the cookie jar. If the heart is taken to be pumping the blood, circulating the blood or delivering the poison, then explanations of these activities will necessarily include different aspects of the heart's context.

So if we are to embrace a view of realization that provides an account of the realization of these sorts of relational properties, activities, and functions, then it seems that we should reject the idea that realizations are constitutive or intrinsic. But perhaps we should go further. Once we attend to the role of context in realization in cases like these, even the relatively weak claim that there are *any* instances of truly intrinsic realization seems difficult to maintain. Suppose that we consider not the heart's pumping of blood but its beating, an activity that we might think of as being purely intrinsic to it. Yet even the beating of the heart, truth be told, cannot be explained by reference only to intrinsic properties of hearts. Sympathetic and parasympathetic inputs to the heart, both from neural input and from circulating hormones, all contribute to different aspects of the heart's beating (its rate, regularity, etc.). It seems likely to us that mechanistic realization, as a species of explanatory realization, will typically (if not always) involve contextual in addition to constitutive factors. And so there seem to be good reasons for rejecting R_2 as a constraint on realization.

4.3 Must Realization be Synchronic?

It might be suggested, and most seem to assume, that realization is a synchronic relationship. We might build this into our canonical realization statement as follows:

(R_3) In C, O's being P over the interval $[t_o, t_f]$is realized by O's being Q over the interval $[t_o, t_f]$.

Features of O at $t_n < t_o$ or at $t_m > t_f$ cannot be part of P's realizer. One reason that the metaphysician might want to make realization synchronic is that P's realizer is supposed to explain the "causal powers" individuative of O's being P— that is, the set of physically possible causes and effects of O's being P. Since it seems reasonable to suppose that O's causal powers depend exclusively on the properties of O (along with its parts and its context) at a given time— and not to features of O's past— one might wish to restrict the realization relation to the occurent properties of O.

The problem with this putative constraint is that there are well-known properties that are individuated by their histories. Consider first objects of which this is true. The Mona Lisa would not be the Mona Lisa if Brad Pitt painted it instead of DaVinci. Dollar bills would not be dollar bills if Mugsy made them rather than the treasury. And a mouse would not be a mouse if it did not descend from parents that were mice. For those who have called attention to such cases, this is not an empirical fact: we could not discover that classic paintings and dollar bills could be made by anyone or that mice could in fact be made in test-tubes (perhaps they can, but that is very much beside the point). At best we could discover that things very much like the Mona Lisa, dollar bills and mice could be fabricated down to the last detail— but they would still not be the Mona Lisa, dollar bills or mice. In order to accommodate such cases, we may wish to recognize *historical realization* as a sub-species of realization. Historical realizers, in all of the cases with which we are familiar have contextual components. They include relationships to objects outside of O, such as DaVinci, the U.S. Treasury, and parents. But in addition, they include non-occurent, historical properties of O.

A number of philosophers have appealed to historical realization to, as Elliott Sober has said, "put the function back in functionalism" (Sober 1985; see also Neander 1991, Millikan 1993). According to one standard account of biological functions, some item S has the function of T-ing in O (as opposed to merely functioning as a T-er in O) if and only if O has S because S Ts. "Because S Ts" is then fleshed out with an historical causal story to the effect that Os with Ss T-ed better than those that did not and so were preserved in the population of Os. Many have thought that appeal to such a biologically-inspired notion of proper functions can distinguish how an item ought to behave from the way that it in fact behaves, that it can vouchsafe the existence of higher level kinds in spite of the fact that there is no common occurent realizer that all of them share, and that it can provide a criterion for individuating biological kinds (what makes a heart a heart is that it was selected for its pumping).

Again, this is an important starting point for many scientists and philosophers, and so we see no reason to build into the very notion of realization a commitment that metaphysically excludes historical realizers. Thus, we do not think that R_3 should be viewed as a constraint on realization. It seems best to treat the acceptability of historical realization (in discussions of function, or information, or computing) on a case by case basis. For understanding higher level causation, this might be a problem, since it certainly seems that only occurent properties contribute to causal powers: the same causal powers could be achieved by any number of histories. But issues of this sort require more discussion than space affords at this time. To the extent that different scientific and philosophical discussions turn on allowing the possibility of historical realization, we think it would be imprudent to rule out historical realization from the start.

5. Autonomy, Reduction, and Multiple Realization

We have already noted that, from the get-go, philosophers of mind have happily endorsed the idea that mental states are multiply realizable in different physical states. Fictitious Martians, fancied artificial intelligences, and presumed cross-species physical variation have all been invoked in support of this idea. To many, the multiple realizability of the mental provided the basis for thinking that strict mind-brain identity theories were mistaken, and that non-reductionist, functionalist views of the mind constituted a more compelling account of the relationship between the mental and the physical. Along with this metaphysics came the methodological moral that psychology and the cognitive sciences more generally are *autonomous* of the nitty-gritty details of basic neuroscience, such that we could do much (if not all) of our psychological theory construction without being hostage to developments within neuroscience.

This sort of view, which predominated throughout much of the 1970s and 1980s, came under attack toward the end of that period. There were two prongs to the critique, which together have suggested to many that the complacency about multiple realization, and its putative implications, cannot be justified.

The first was delivered by Jaegwon Kim in his attack on “the myth of non-reductive materialism”. In a series of influential papers (1989, 1992, 1993), Kim argued that there were metaphysical reasons for thinking that non-reductive forms of materialism, including functionalism, were unstable hybrids. By appealing to causal powers, the nature of scientific explanation, natural kinds, and laws, Kim argued for the reductive materialism that he had begun his career defending, and against the autonomy of psychology.

To give the flavor of Kim’s discussion and to see how it manifests the metaphysical strand to views of realization, consider one principle that he appeals to, what he calls the *principle of causal inheritance*: “if M is instantiated on a given occasion by being realized by P, then the causal powers of *this instance of* M are identical with (perhaps, a subset of) the causal powers of P” (Kim 1993: 355). The intuitive idea behind the principle is that instances of higher-order kinds, such as psychological kinds, inherit the causal powers of instances of the lower-order kinds that realize them. The plausibility of this intuition turns on the instances of each of these kinds being the very same glob of matter in the world, or close enough. And if the causal powers of instances are inherited from the physical to the mental, then it seems that all of the real causal action is captured by the lower order description in terms of physical properties and powers. Thus, it is confused to think of mental causation as somehow distinct from or autonomous of physical causation.

The second prong to the attack on appeals to multiple realization derived from advances in the neurosciences, particularly from the emergence of *cognitive neuroscience* as a distinct subfield within the cognitive sciences. Developing as an area bridging cognitive psychology and neuroscience over the last twenty years, cognitive neuroscience has come to occupy a prominent place within the cognitive sciences. Perhaps this place will become more prominent still. In concluding their introduction to the hundred or so articles on neuroscience in *The MIT Encyclopedia of the Cognitive Sciences* (1999), Tom Albright and Helen Neville make a telling (even if tongue-in-cheek) prediction. They say that "if cognitive neuroscience fulfills its grand promise, later editions of this volume may contain a section on history, into which all of the nonneuro cognitive sciences discussion will be swept" (1999:lxix).

Hyperbole and rhetorical flourish aside, the rise of cognitive neuroscience has provided reason to rethink the phenomenon of multiple realization and what has been claimed about cognition in its name. As Bechtel and Mundale (1999) have argued, appeals to function are an intrinsic part of neurotaxonomy, and such appeals are often justified by cross-species investigations. Although the taxonomy of the brain into 47 regions by Korbinian Brodmann early in the 20^{th}-century had an anatomical basis, as Brodmann says, its "ultimate goal was the advancement of a theory of function and its pathological deviations" (1909/1994: 243, as quoted by Bechtel and Mundale 1999:180).

Likewise, deficit studies in both human and non-human animals, a major source of information about psychological functioning, presupposes that the same part of the brain plays much the same physical role across different individuals. Hence the contrast between "functional", psychological levels and "physical", neural levels, and the idea that cross-species investigations and generalizations provide the basis for viewing the neural level as too fine-grained to capture genuinely psychological commonalities, both presuppose a view of the neurosciences that is sadly out of touch with the taxonomic and explanatory practices in those sciences.

Moreover, a closer look at neurotaxonomic practices raises doubts about the very terms in which the multiple realizability thesis has been stated. Philosophers typically talk of mental states as being realized in brain states. As Bechtel and Mundale also note, "the notion of a brain state is a philosopher's fiction." (1999: 177). Cognitive scientists themselves are often interested in activities, functions, and mechanisms, not states *per se*. While brain states, whatever those are, may be multiply realizable, this will not establish that the functions and activities of the brain are. What is needed is a notion of realization, such as that discussed at the end of Section 3, that will accommodate cases of mechanistic realization.

These points—the mischaracterization of neurotaxonomy, the false dichotomy between functional and physical levels, and departure that talk of states makes from empirically-grounded practice--together suggest several ways in which traditional debates and positions which invoke the notion of multiple realization need to be rethought in light of attention to work in the neurosciences.

Larry Shapiro (2000, 2004) has also argued that both traditional and more recent putative examples of multiple realization fail because they do not constitute examples where *the same* cognitive function is realized by *relevantly different* physical mechanisms. For example, consider common appeals to the idea that silicon chips could replace neurons, one by one, to produce a silicon brain, without changing the psychological functions realized. Such examples provide support for the multiple realizability of cognition only if the *ways* in which neurons and silicon chips realize psychological functions are relevantly different. As Shapiro says, if "each neuron's contribution to psychological capacities is solely its transmission of an electrical signal, and if silicon chips contribute to psychological capacities in precisely the same way, then the silicon brain and the neural brain are not distinct realizations of the mind" (2000: 645). At best, this would be an example of trivial multiple realization.

Shapiro's general point can be applied to more detailed examples that draw on recent work in cognitive neuroscience on neural plasticity (see Shapiro 2004: ch.2). Ned Block and Jerry Fodor had appealed early on to the lability of the brain, as evidenced in the differential realization of the capacity for language in the left and right hemispheres in some cases, to defend the multiple realizability of psychological functions (Block and Fodor 1972). The flexibility and adaptability of both organism-level psychological functions, such as vision or memory and the determinate forms they take have long been a subject of investigation within psychology, and much of this work has been viewed more recently under the rubric of "neural plasticity" (Gilbert 1999). There has been a surge of work on varieties of neural plasticity, including experience-dependent neural reprogramming (Elbert et al. 1995), neural adaptation following cortical lesions (Kaas 2000), and cross-modal neural rewiring (von Melchner et al. 2000). Consider this last bit of work in more detail.

Mriganka Sur and his colleagues have explored lesion-induced neural rewiring in the auditory system of young ferrets. In this paradigm, a lesion was induced in a part of the auditory system upstream from the auditory cortex. This involved severing the normal connection between the inferior and superior colliculus in the midbrain and the medial geniculate nucleus (MGN) in the left hemisphere of the ferrets. The effect of the procedure was that, over time, the MGN then comes to accept projections from the retina, passing on this information in turn to the auditory cortex. This allows the rewired ferrets to "see with their auditory systems". Could this be a case where cognitive neuroscience has shown that psychological functions are multiply realizable?

Like the imagined cases that are the stock and trade of traditional philosophical analysis, such cases must also jointly satisfy two conditions: they must describe the same psychological capacity and it must be realized in a relevantly different manner. Sur's work on cross-modal neural rewiring, interesting as it is, appears to satisfy neither of these conditions. Although rewired ferrets have some visual capacity, they clearly do not have the same visual capacity as normal ferrets. But perhaps we need to focus not on the coarse-grained capacity of vision but on specific visual abilities, such as the capacity to detect light and to distinguish grates of light and dark. While even here there remain differences between normal and rewired ferrets, they are interestingly similar in their abilities, and perhaps we can grant that they do share some basic visual capacities.

This brings us to the second condition: does the auditory cortex of the rewired ferrets, for example, process visual information in a relevantly different manner than does the visual cortex of normal ferrets? We have a case of multiple realization only if the answer to this question is "yes". As Sur and his colleagues point out, however, although this question remains somewhat open, the organization of the auditory cortex in rewired ferrets is strikingly similar to that of the visual cortex in normal ferrets, and strikingly unlike that of the auditory cortex in normal ferrets. Sameness of structure does not, of course, imply sameness of function, but this at least casts some doubt on whether this example would satisfy Shapiro's second condition even if it were to satisfy the first. As Shapiro says, it "seems not unreasonable to suppose that a rewired ferret sees as well as it does only to the extent that its auditory cortex resembles a visual cortex" (2004:ch.2, p.64). If this supposition is not just reasonable but correct, then there is little support for the idea of multiple realizability to be garnered from at least this example of cross-modal neural rewiring.

Our view is that further consideration of the varieties of neural plasticity, and of the details of psychological and neural taxonomies, is very unlikely to support the actual multiple realizability of the psychological in the neural. Both Shapiro and Bechtel and Mundale have provided reasons that we think are the basis for skepticism here.

The two basic constraints that Shapiro identifies—same psychological function but (relevantly) different neural realization—pull in opposite directions. This is in large part because "the psychological" and "the neural" should not be viewed simply as two levels. Instead, as Bill Lycan (1987:ch.4) and Carl Craver (2001) have suggested, there are many levels interstitial between lowest and highest. Further, the relationship between them is complex enough that we should be wary of untutored level-talk and might even question the very utility of the levels metaphor

Bechtel and Mundale (1999, section 5) have argued that one reason why multiple realizability seems so obviously to hold of the relation between the mental and the physical is the tacit adoption of a sort of double standard in thinking about the two. In thinking about psychological capacities, it is common to describe them coarsely—as the capacity for vision, or for short-term memory. By contrast, realizing neural structures and their "immediate" functions receive comparatively fine-grained description. For example, it is common to consider cases in which pain is the psychological state under consideration, while the physical realizers are described in terms of the different kinds of brains that, say, human beings and octopi have. Given that a coarse grain of description facilitates the view that two cases are instances of the same (psychological) kind, and that a fine grain of description does the same for the view that two cases are instances of different (neurological) kinds, the bias that this double-standard introduces is one that creates the impression that such cases satisfy both of Shapiro's constraints.

Finally, our discussion of neural plasticity, brief as it is, raises the issue of what has been called *emergent realization* (Wilson 2001), i.e., of a physical realizer for a given psychological capacity that could realize some other capacity were the world different in various ways. The idea of emergent realization is sometimes introduced in contrast to that of multiple realization, as a case where rather than there being a one-many relation between the psychological and the physical, there is a many-one relation. Metaphysicians of mind, in at least tacitly accepting the idea that realizers are metaphysically sufficient for the properties, states, and capacities that they realize, typically reject the possibility of emergent realization as incompatible with versions of physicalism that hold that the physical facts must, in some sense, determine all the facts there are.

The examples of neural plasticity that we have mentioned might be thought to give us pause here. For these appear to be cases in which once the world is changed in various ways—certain lesions are made, neural rewiring develops over time—one part of the brain comes to realize psychological capacities that it previously did not realize. But our preceding discussion should make us wary of any quick inferences here, and remind us to consider the empirical details more fully. What happens in these cases, after all, is that parts of the brain come to be not only differently configured in terms of how they connect to other parts of the brain, but also differently *structured* themselves. In the ferret brain, for example, the neural rewiring introduced by severing the auditory input channels to the MGN makes the physical structure of the auditory cortex, downstream, much more like the visual cortex of normal ferrets, bearing a columnar structure that is orientation-sensitive as well as a two-dimensional map of the surrounding visual space.

6. Conclusion: Getting Our Heads Together

We began with the idea that the concept of realization was the servant of two masters. The idea of realization is at the heart of a number of debates in contemporary metaphysics and the philosophy of mind. It is also (at least implicitly) central to the explanatory and investigative practices of cognitive scientists. Here we have tried to review possible lines of rapprochement. We think that a notion of realization will continue to have an important role to play in systematically thinking about the nature of cognition, and we have tried to accommodate the demands placed on the concept by both philosophers and cognitive scientists. While we have tended to adopt an ecumenical view of how we should think about realization, we have also tried to identify a series of issues that remain contentious and that invite further discussion. Some of these concern what properties the relation of realization has, or what constraints it should be subject to—is it decompositional, intrinsic, or synchronic in nature? (section 4)—while others focus more directly on the interplay between abstract philosophy concerns (say, about the autonomy of psychology) and ongoing explanatory practice in cognitive neuroscience (section 5). Both, we reckon, require that metaphysicians and scientists continue to get their heads together.

See also: mechanism, reduction

References

Albright, T., and H. Neville, 1999, "Neurosciences", in R.A. Wilson and F.C. Keil (editors), *The MIT Encyclopedia of the Cognitive Sciences*. Cambridge, MA: MIT Press, pp.li-lxxii.

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

Block, N., and J.A., Fodor, 1972, "What Psychological States are Not". Reprinted in J.A. Fodor, *Representations*. Sussex, England: Harvester, 1981.

Brodmann, K., 1909, *Vergleichende Lokalisationslehre der Grosshirnrinde*. Leipzig: J.A. Barth. English translation by L.J. Garey, *Brodmann's 'Localisation in the Cerebral Cortex'*. Cambridge, MA: MIT Press, 1994.

Craver, C.F., 2001, "Role Functions, Mechanisms and Hierarchy" *Philosophy of Science* 68: 31-55.

Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., Taub, E., 1995, "Increased use of the left hand in string players associated with increased cortical representation of the fingers" *Science* 270: 305-307.

Gilbert, C., 1999, "Neural Plasticity", in R.A. Wilson and F.C. Keil (editors), *The MIT Encyclopedia of the Cognitive Sciences*. Cambridge, MA: MIT Press, pp.598-601.

Gillett, C., 2002, "The dimensions view of realization: a critique of the Standard view" *Analysis* 64: 316-23.

Kaas, J., 2000, "The Reorganization of Sensory and Motor Maps after Injury in Adult Mammals", in M. Gazzaniga (editor), *The New Cognitive Neurosciences*, 2nd edition. Cambridge, MA: MIT Press, pp.223-236.

Kandel, E. R., and J. H. Schwartz (editors), 2000, *Principles of Neural Science*. 4th Edition. New York: McGraw Hill.

Kim, J., 1989, "The Myth of Nonreductive Materialism". Reprinted in his *Supervenience and Mind*. New York: Cambridge University Press, 1993.

Kim, J., 1992, "Multiple Realization and the Metaphysics of Reduction". Reprinted in his *Supervenience and Mind*. New York: Cambridge University Press, 1993.

Kim, J., 1993, *Supervenience and Mind*. New York: Cambridge University Press, 1993.

Kim, J., 1996, *Philosophy of Mind*. Cambridge, MA: MIT Press.

Kim, J., 1998, *Mind in a Physical World*. Cambridge, MA: MIT Press.

Lycan, W. G., 1990, “The Continuity of Levels of Nature,” in W. G. Lycan (editor), 1990, *Mind and Cognition: A Reader*. Oxford: Bradford.

Machamer, P.K., L. Darden and C. F. Craver, 2000, "Thinking about Mechanisms" *Philosophy of Science* 67: 1-25.

Marr, D., 1982, *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco, CA: Freeman.

Millikan, R.G., 1993, *White Queen Psychology and Other Essays for Alice*. Cambridge, MA: MIT Press.

Neander, K., 1991, “Functions as Selected Effects”, *Philosophy of Science* 58: 168-184.

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

Polger, T., 2004, “Neural Machinery and Realization”, *Philosophy of Science* 71: 997-1006.

Putnam, H., 1960, “Minds and Machines”. Reprinted in his *Mind, Language, and Reality: Philosophical Papers, Volume 2*. New York: Cambridge University Press, 1975.

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

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

Smart, J.J.C., 1959, “Sensations and Brain Processes”. Reprinted in D. Rosenthal (editor), *The Nature of Mind*. New York, Oxford University Press, 1991.

Sober, E., 1985, “Putting the Function Back into Functionalism”. Reprinted in W.G. Lycan (editor), *Mind and Cognition: An Anthology*. Oxford: Blackwell, 1998, 2nd edition.

von Melchner, L., S. Pallas, and M. Sur, 2000, “Visual Behaviour Mediated by Auditory Cortex Directed to the Auditory Pathway”, *Nature* 404: 871-876.

Wilson, R.A., 2001, “Two Views of Realization”, *Philosophical Studies* 104: 1-31.

Wilson, R.A., 2004, *Boundaries of the Mind: The Individual in the Fragile Sciences: Cognition*. New York: Cambridge University Press.

Wilson, R.A. and F.C. Keil, 1998, “The Shadows and Shallows of Explanation”, modified version reprinted in F.C. Keil and R.A.

Wilson (editors), *Explanation and Cognition*. Cambridge, MA: MIT Press, 2000.

Wilson, R.A. and F.C. Keil (editors), 1999, *The MIT Encyclopedia of the Cognitive Sciences*. Cambridge, MA: MIT Press.

Wimsatt, W., 1997, "Aggregativity: Reductive Heuristics for Finding Emergence" *Philosophy of Science* 64: S372-S384.