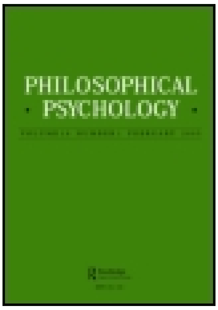

Philosophical Psychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/cphp20

Mechanisms and natural kinds

Carl F. Craver

Available online: 06 Oct 2009

To cite this article: Carl F. Craver (2009): Mechanisms and natural kinds, Philosophical Psychology, 22:5, 575-594

To link to this article: http://dx.doi.org/10.1080/09515080903238930

Philosophical Psychology
Vol. 22, No. 5, October 2009, 575–594

Routledge
Taylor & Francis Group

Mechanisms and natural kinds

Carl F. Craver

It is common to defend the Homeostatic Property Cluster (HPC) view as a third way between conventionalism and essentialism about natural kinds (Boyd, 1989, 1991, 1997, 1999; Griffiths, 1997, 1999; Keil, 2003; Kornblith, 1993; Wilson, 1999, 2005; Wilson, Barker, & Brigandt, forthcoming). According to the HPC view, property clusters are not merely conventionally clustered together; the co-occurrence of properties in the cluster is sustained by a similarity generating (or homeostatic) mechanism. I argue that conventional elements are involved partly but ineliminably in deciding which mechanisms define kinds, for deciding when two mechanisms are mechanisms of the same type, and for deciding where one particular mechanism ends and another begins. This intrusion of conventional perspective into the idea of a mechanism raises doubts as to whether the HPC view is sufficiently free of conventional elements to serve as an objective arbiter in scientific disputes about what the kinds of the special sciences should be.

Keywords: Mechanism; Natural Kinds; Taxonomy

1. Introduction

The homeostatic property cluster (or HPC) view of natural kinds is the view that natural kinds are clusters of properties that co-occur because of the operation of a homeostatic mechanism. One normative constraint often taken to follow from the HPC view is that taxonomies of natural kinds are adequate to the extent—and only to the extent—that the kinds in that taxonomy track mechanisms, that is, that the recognized kinds correspond to the mechanisms that constitute the causal structure of the world. Kind concepts cut nature at its joints, and according to this normative constraint, nature's joints are located at the boundaries of mechanisms.

Can the mechanistic structure of the world serve as an objective foundation for an adequate taxonomy of kinds? If the normative constraint that kinds should track

Carl F. Craver is Associate Professor of Philosophy at the Washington University in St. Louis.
Correspondence to: Carl F. Craver, Washington University St. Louis, St. Louis, MO 63130, United States. Email: ccraver@artsci.wustl.edu

ISSN 0951-5089 (print)/ISSN 1465-394X (online)/09/050575-20 © 2009 Taylor & Francis
DOI: 10.1080/09515080903238930

mechanisms is to guide the effort to build scientific taxonomies, we must say what mechanisms are and how they can be distinguished from one another. Recent philosophical work dedicated to understanding mechanisms can be used to make the notion of "mechanism" central to the HPC account more precise (Bechtel & Abrahamsen, 2005; Bechtel & Richardson, 1993; Craver, 2007; Machamer, Darden, & Craver, 2000). However, that work provides scant basis for saying when two mechanisms are mechanisms of different types and for saying where one mechanism ends and another begins. In that literature, answers to such questions appeal to pragmatic considerations about the intended use of one's knowledge about the mechanism. Perhaps this is a limitation of these accounts of mechanisms. Perhaps, however, this is a reason to question whether the mechanistic structure of the world can arbitrate objectively among competing taxonomies of scientific kinds. After describing and motivating the HPC view and the related normative constraint, I review three limitations to the ability of the mechanistic structure of the world to serve as such an objective arbiter. While these limitations do not render the HPC view hopeless as a touchstone in the effort to regiment our taxonomic practices, they show that there is much work to be done to show how some descendent of the HPC view might assist in the effort to say what the natural kinds in the special sciences are.

2. HPC Kinds

Hillary Kornblith (1993) presents the HPC view as a third way between conventionalism and essentialism about natural kinds (see also Wilson et al., forthcoming). His discussion highlights the many attractions of the HPC view and sets the stage for the problems I consider below.

Kornblith's *conventionalists* hold that there is no objective distinction between natural kinds and any other cluster of properties. Any given property cluster lies on a continuum between useful and useless, and the term "natural kind" picks out nothing more than a cluster of properties falling closer to the useful end of the spectrum. The difference between natural kinds and merely conventional property clusters lies not in the phenomena themselves but in the psychological or social factors that determine how human beings find it useful to chunk the world for a given purpose. Ian Hacking aptly describes conventional kinds as kinds "constructed along lines of family resemblance .... What puts things into a family is not nature but people in concert" (quoted in Boyd, 1989, p. 28). Conventional kinds are defined by our judgments about how useful a particular family of properties is thought to be for the purposes of prediction, explanation, and control, by our aesthetic yen for desert landscapes or rainforest ontologies, and by psychological and cultural constraints on what we take to be acceptable kinds. Advocates of the HPC view argue, in contrast, that a cluster of properties constitutes a natural kind when the co-occurrence of the properties in the cluster is explained by a homeostatic mechanism. My focus below is on the apparent contrast between the HPC view and conventionalism.

Kornblith's *essentialists* hold that natural kinds are defined by necessary and sufficient conditions, that is, the set of properties shared by all and only members of the kind. Most advocates of the HPC view are philosophers of biology and psychology. They object to essentialism on the grounds that the natural variability among biological and psychological individuals is too great for kinds in such domains to be identified with necessary and sufficient conditions (see especially Wilson 2005; Wilson et al., forthcoming). Indeed, Boyd (1989, 1991, 1997) advances the HPC view as a way to preserve the idea that there are natural kinds in the domains of the special sciences while acknowledging the shortcomings of essentialism for describing them. On his view, special science kind terms are often vague in the sense that there is no fact of the matter whether a given particular is a member of the kind. Instances of a special science kind can and do differ from one another both in the number of phenomenal properties they share and in the components and organizational features of the mechanisms that explain those phenomenal properties. In other words, HPC kind concepts have a prototype structure, and HPC kinds are multiply realizable.

The limitations I discuss below suggest that the HPC view will have difficulty accommodating variability and multiple realizability in special science kinds without sliding into a promiscuous form of conventionalism.[1] Promiscuity is not a serious problem for the view of natural kinds as such; there might be many more kinds than we expected. However, absent further refinement, the HPC view countenances a level of promiscuity that weakens the force of the normative demand that one should regiment scientific taxonomies (such as species taxa, DSM diagnostic categories, and mental modules) by conforming them to the mechanistic structure of the world.

Advocates of the HPC view charge that conventionalists cannot explain why some taxonomies of kinds are more useful for prediction, explanation, and control than are conventional kinds. If kinds are purely conventional, they argue, nothing in nature makes it the case that one member the kind is relevantly similar to other members. Predictions warranted for one member of a kind may not be warranted for another because conventional kinds might gloss over predictively heterogeneous clusters of properties and mechanisms. Likewise, if kind membership is a matter of convention then kind membership arguably cannot explain why an item has the properties it does. Finally, if kinds are conventional, there is no reason to expect that like interventions will have like effects on two members of the kind. For the conventionalist, there is no further fact of the matter: some kinds are more useful or interesting to us than others, but there is no (or need be no) further fact about the world that explains why this is so. The natural kinds are merely the useful ones.[2]

Defenders of the HPC view, in contrast to conventionalists, hold that natural kinds are real divisions in the structure of the world. Useful kinds are useful for prediction, explanation, and control, they claim, precisely because the kinds are sustained by mechanisms. Instances of the kind are similar, not merely in their phenomenal properties, but also because the co-occurrence of the phenomenal properties is explained by a similarity-generating (that is, homeostatic) mechanism.

Boyd (1989, pp. 16–17; see also 1991, pp. 140–143) summarizes his view in ten points. Four features of HPC kinds are especially relevant below:

B1) Property Cluster: a cluster of properties that regularly occur together;
B2) Homeostatic Mechanism: a mechanism that explains why the clustered properties occur together;
B3) Causal Import: that the property cluster figures in important causal generalizations; and
B4) Accommodation: that any refinement of the definition of the kind either introduces causally and inductively irrelevant distinctions or glosses over causally and inductively relevant similarities.

B1 is self-explanatory. B2 is introduced to distinguish natural kinds from conventionally defined property clusters and to account for the predictive, explanatory, and instrumental value of natural kind concepts. The term "homeostatic" is potentially misleading; it is much weaker than Bernard's idea of maintaining the constancy of an internal milieu in the face of environmental changes. A mechanism is homeostatic in Boyd's sense if the mechanism explains the regular co-occurrence of phenomenal properties in the cluster. Boyd (1989) refers specifically to *underlying* mechanisms, but he also allows that a homeostatic mechanism might be composed of causal relations among the phenomenal properties themselves (in which case, the mechanism is not underlying or "lower-level"), and it might involve the antecedent (etiological) causes of the property cluster (see also Keil, 2003; Wilson et al., forthcoming). The central point is that the mechanism explains the co-occurrence of the properties in the cluster. Griffiths points out that the homeostatic mechanisms in this account play the role of essences in the traditional account (1999, p. 218). To recognize that an item is a member of an HPC kind licenses the prediction that the item likely has the properties in the cluster because there is a mechanism that explains why things of that kind tend to have the properties they do. To recognize that an item is a member of an HPC kind licenses explanations because natural kind terms identify features of the causal structure of the world, and to explain something, one might plausibly argue, just is to show how it fits into the causal structure of the world (Craver, 2007; Salmon, 1984). Furthermore, to recognize that an item is a member of an HPC kind affords control because the mechanisms underlying the kind are presumably such that they respond in similar ways to interventions. In short, B2 distinguishes the HPC account from conventionalism by providing a reason for thinking that natural kinds will be useful for prediction, explanation, and control in ways that mere conventional kinds are not.

B3 rules out property clusters that are sustained by homeostatic mechanisms but are of little or no theoretical or practical value. For example, any regularly correlated effects of a common cause (such as storms, falling barometers, and joint pain in the elderly) constitute a cluster of properties sustained by a mechanism as required by B2. B3 rules out many such cases. On the most natural reading of "import," B3 clearly introduces perspectival elements into HPC kinds, given that import is a matter of whether the kind in question appears in our theories or is otherwise important for our aims and objectives. Yet one might understand "import" in more objective terms as requiring merely that the clustering of properties be causally relevant to something (perhaps in the sense that the cluster could make difference to whether or how something happens). The gain in objectivity is balanced by a corresponding increase in promiscuity: on this reading, there are many more natural kinds than science will ever find it useful or interesting to recognize, let alone study.

B4 is also susceptible to a weak and a strong reading. On the strong reading, it demands as a normative principle that the correct taxonomy of natural kinds must be immune to revision in light of future discoveries about the mechanistic structure of the world. For example, if there are causally relevant differences between individual cases of schizophrenia (such as that some respond impressively to treatment with antipsychotic medications and that some do not), then B4 reasonably enjoins psychiatry to recognize this difference in its taxonomy (the DSM). Similarly, if there are no causally relevant differences between individual samples of jade, then the taxonomy should not distinguish them. This is how I read Boyd when he says the naturalness of a natural kind "consists in a certain accommodation between the relevant conceptual and classificatory practices and independently existing causal structures" (1997, p. 55) and that the taxonomy of kinds should neither (i) gloss over causally relevant differences between members of a kind, nor (ii) recognize differences among members of a kind that are causally irrelevant. To paraphrase the strong formulation, kinds should track mechanisms. On the weaker reading of B4, one simply wants the taxonomy of kinds to pick out sets of things that are similar enough for our explanatory or instrumental projects. For biologists, onions might well be lilies; for gardners, onions are not lilies. It all depends on what you are trying to do, and none of this has to be very precise. Accommodation in this case is to the instrumental objectives of the researcher, first and foremost, and only secondarily (if at all) to the causal structure of the world.

One can weaken this view of natural kinds by dropping or weakening one or more of Boyd's commitments. It should be noted, however, that each commitment clearly hangs with the others in his defense of realism about natural kinds. It is possible, for example, to reject B2 and to keep the rest as a *simple causal theory* of natural kinds. According to this view, natural kinds are the kinds appearing in generalizations that correctly describe the causal structure of the world regardless of whether a mechanism explains the clustering of properties definitive of the kind. This is clearly not Boyd's view. Boyd introduces B2 to defend natural kinds against conventionalism; the mechanism explains why the kind is so useful for prediction, explanation, and control in a way that conventionalism cannot. Wilson distinguishes the HPC view from the weak Wittgenstinian notion of a cluster concept (or "family resemblance concept") by insisting that "the mechanisms that maintain any given HPC are a part of the natural world, not simply our way of thinking about or intervening in the world" (2005, p. 118). The predictive and instrumental value of building a taxonomy that corresponds to these causal structures, as discussed above, relies on the fact that the kinds pick out real mechanisms. The predictive and instrumentally valuable kinds are those that get the causal structure right. In considering the existence of taxa at higher levels than species, for example, Wilson argues that unless one can identify a "common mechanism" to explain why the properties in a cluster occur together, one would simply have a "disjunction" of two or more clusters, not a homeostatic cluster of properties (Wilson, 2005, p. 110). Finally, Griffiths notes: "the essential property that makes particular instances members of the kind is their relation to that causal mechanism [that explains the clustering of properties], whatever it may be" (1999, p. 218). HPC kinds would not be HPC kinds without B2.

The HPC view stripped of requirement B3 yields a simple metaphysical view of kinds, according to which there are kinds wherever there are similarity-generating mechanisms. Consider the property cluster composed of falling mercury, storms, and joint pain in the elderly. This cluster is a natural kind because there is a common barometric mechanism that sustains their regular (though imperfect) co-occurrence. The cluster is not causally or theoretically important for anything that we wish to do at the moment, but it is nonetheless a portion of the causal structure of the world. (It satisfies the objective but promiscuous reading of B3; but an imaginative thinker can make any property cluster do so). I assume that a view of natural kinds that includes such kinds as this is too weak to arbitrate between scientific taxonomies. This *weak metaphysical HPC*, while not conventionalism full-blown, is closer to the conventionalist than the realist end of the spectrum (Boyd, 1999, embraces such a weak view; Wilson, 2005, defends a stronger view). Below I raise objections to the normative, *scientific construal* of the HPC (henceforth, the scientific HPC), not this weak metaphysical thesis.

Finally, one can reject B4 and allow that a scientific taxonomy picks out natural kinds even when it would be predictively and explanatorily fruitful to further refine the taxonomy on the basis of what one learns about the causal structure of the world. To abandon B4, however, is to weaken the normative force of the HPC for arbitrating among scientific taxonomies. The core normative edict is that the distinct natural kinds ought to correspond to distinct mechanisms. The scientific HPC is a normative principle for deciding what the kinds are. If major depression and schizophrenia are natural kinds, there should be mechanisms underlying each, and they should be different. If schizophrenia turns out to be produced by two distinct underlying mechanisms in different patients, then the correct taxonomy should recognize that difference. Peter Carruthers (2004), Churchland and Sejnowski (1992), Paul Griffiths (1997), and Rob Wilson (2005, p. 110) endorse something like this scientific construal, as apparently do many scientists. One central (perhaps the central) thrust of the present essay is that a weak accommodation thesis is too weak to avoid conventionalism and that a strong accommodation thesis is untenable. One way forward, then, would be to articulate B4 in a way that makes clear what the taxonomy of kinds is being accommodated to (is it the causal structure of the world, our instrumental needs, or our inferential practices?) and how B4 can be made consistent with pluralism in our special science taxonomies without sliding all the way into conventionalism (that is, how do we accommodate our taxonomies so as to include all the kinds we want and none of the kinds we don't want?).

3. Taxonomic Reform: Splitting and Lumping

Two common strategies for accommodating a scientific taxonomy to the mechanistic structure of the world embody central commitments of the scientific HPC. The geneticist Victor McKusick (1969) calls them splitting and lumping. One intuitive appeal of the HPC view among scientifically minded philosophers is that it provides a natural and obvious explanation of why these strategies are so important in the special sciences (see Wilson, 2005, p. 58).

The splitting strategy can be expressed informally in the directive: if you find that a single cluster of properties is explained by more than one mechanism, split the cluster into subset clusters, each of which is explained by a single mechanism. Paul Griffiths (1997) uses this strategy to argue the putative kind "emotion" should be replaced by three kinds—affect programs, higher cognitive emotions, and social emotions. Psychiatric geneticists (such as McKusick) argue that nosological categories (such as "schizophrenia" and Rett syndrome) should be reformulated on the basis of the genetic mechanisms underlying these disorders. Many of those who claim that the mind is massively modular (e.g., Carruthers, 2004; Pinker, 1997) argue for their thesis on the basis of splitter arguments. Many neuroscientists and psychologists argue that memory is not a natural kind, and should be split into, for example, echoic, episodic, procedural, and working memory because there are demonstrably different mechanisms for each (e.g., Schacter & Tulving, 1994). The logic of dissociation is splitter logic.

Arguments for splitting come in many forms. It is claimed that the mechanisms are distinct because they develop independently, because they evolved separately, and because they can be manipulated or damaged independently of one another. If so, the argument continues, the putative kind is heterogeneous, and one is forced by B4 to refine the taxonomy further. The refined taxonomy will be more useful for prediction, explanation, and control than is the taxonomy that includes the heterogeneous kind. The refined taxonomy, in other words, is better accommodated to the causal structures that sustain the inductive practices of a science.

The lumping strategy is expressed in the informal directive: if you find that two or more putatively distinct kinds are explained by the same mechanism, lump the putative kinds into one. Burning, rusting, and breathing are all oxidation. For Freud, neurotic symptoms, jokes, and slips of the tongue are all distinct expressions of repressed ideas and primary process transformations. Pascal Boyer (2003) argues that religious rituals and symptoms of obsessive-compulsive disorder are both produced by changes in a single mechanism in the basal ganglia. Again, a taxonomy that lumps these into a single kind will not introduce differences without distinctions and so will better accommodate findings about the causal structure of the world.

Splitting and lumping might reasonably be described as strategies for producing scientific taxonomies that track the mechanistic structure of the world. The HPC view makes some of the reasoning in these strategies explicit; but not all of it. The splitting strategy, as read through the lens of the HPC view, applies when one discovers that a putative kind is explained by more than one kind of mechanism. Defenders of the HPC view have not yet shown that there is an objective basis for saying when two kinds of mechanism are different. The lumping strategy read through the same lens applies when one discovers that multiple putative kinds are explained by one underlying mechanism. Advocates of the HPC account have not shown that there is an objective basis for saying when two mechanisms are the same. The challenge I explore below is that one must appeal to conventional factors (or factors beyond what the mechanistic structure of the world affords) to show how mechanisms are typed and individuated. If so, then the scientific HPC slides in the direction of the overly weak metaphysical HPC, which provides little guidance in arbitrating among taxonomies of scientific kinds.

4. Mechanisms

To see why, it is necessary to say what mechanisms are. A consensus view among mechanists holds that mechanisms are entities and activities organized together such that they do something (Bechtel & Abrahamsen, 2005; Bechtel & Richardson, 1993; Craver, 2007; Darden, 2006; Glennan, 1996; 2002; Machamer et al., 2000). The entities are things like brain regions, cells, proteins, and organisms. Activities are the causal components of mechanisms: they are the things that entities do and the ways that the entities act and interact with one another. The entities and activities are organized spatially, temporally, and causally with one another such that they do something that the components could not do on their own. A homeostatic mechanism, again, is a mechanism that explains why the properties in the cluster regularly co-occur.

5. Which Mechanism?

The first challenge to the scientific HPC is to say which of the many homeostatic mechanisms for a given kind matters for the purposes of refining one's scientific taxonomy. A given putative kind typically participates in many different mechanisms. In figure 1, Frank Keil (2003) illustrates the many ways that a property cluster can participate in mechanisms.[3]

Take any given behavior exhibited by some part of an organism (such as the generation of an action potential). This behavior, or function, can be described in terms of a cluster of regularly co-occurring properties.[4] That behavior is explained in part by an underlying mechanism (B in Keil's diagram). However, the different properties in the cluster might all cause one another (as shown in A and D). The rising phase of the action potential, for example, is a cause of its subsequent declining phase. The cluster of properties might be maintained by some external regulatory mechanism, such as the genetic machinery that produces ion channels.

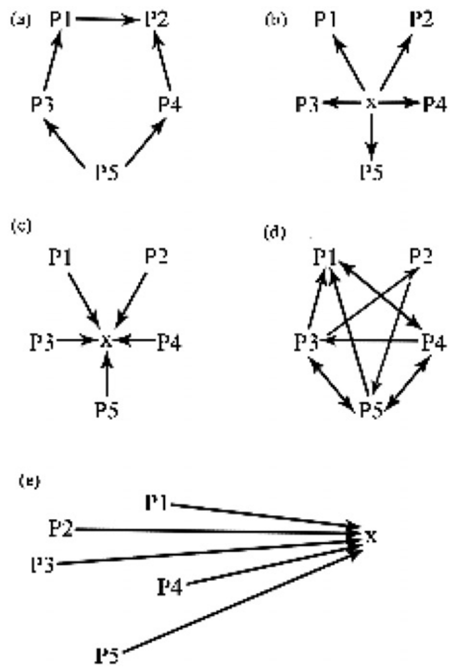

Figure 1 Relations between a property cluster and a mechanism (redrawn from Keil, 2003).

The behavior will also be explained by proximal etiological causes (such as input from pre-synaptic cells), by the less proximal developmental mechanisms that brought the behavior into being (again involving genetics), and by distal evolutionary mechanisms that brought the organ or mechanism to the species in the first place. Keil also considers cases in which the properties in the cluster cause some common mechanism or effect, which expands the range of possible mechanistic entanglements for any given property cluster considerably.

If the HPC view identifies kinds with property clusters that participate somehow in mechanisms (as Keil suggests), then the HPC view recognizes any cluster of properties caused by, sustained by, or productive of a mechanism. If one agrees with Reichenbach (1956) that every correlation among causally independent mechanisms is sustained by a common cause, and one embraces the weak reading of causal importance in B3, then one will find HPC kinds wherever one finds correlations. There is a kind of promiscuity in all of this, but surely the HPC is still less promiscuous than conventionalism if only because it requires that the putative kind participate in some way in the causal structure of the world (see Wilson 1999, 2005).

The problem is rather that one can be led to lump or split the same putative kind in different ways depending on which mechanism one consults in accommodating the taxonomy to the mechanistic structure of the world (to satisfy B4). It is commonplace that multiple etiological pathways can terminate in the same property cluster (think, for example, of the many causes of HIV infection and clinical depression). Focusing on the etiological mechanism, one should split. Focusing on the constitutive mechanism, one should lump. There are also cases in which the same etiological mechanism gives rise to different constitutive mechanisms in different contexts (for example, tertiary syphilis and AIDS each have common etiological pathways that result in strikingly diverse symptom complexes in different patients). Mechanisms that produce the same effects with different components should be lumped on the basis of their effects and split on the basis of their constitutive mechanisms.

This sort of problem has plagued the effort to set psychiatric classification on an objective causal foundation. The same psychiatric disorders (from the perspective of the current taxonomy) often have multiple distinct etiological mechanisms, multiple underlying mechanisms, and variable effects. Often groups with different etiologies and constitutive mechanisms nonetheless suffer similar symptoms and so, for example, can benefit from participation in the same patient groups (Zachar & Kendler, 2007). The simple point is that most putative kinds are entangled with myriad mechanisms. By attending to some of these, one is led to lump. By attending to others, one is led to split.

One is tempted to allow that each resulting taxonomy is equally legitimate, and so to allow that the kind should be lumped for some purposes and split for others (Boyd, 1999, embraces this view). But this allowance seems to run afoul of the accommodation requirement (B4) that there must be no further finding about the causal structure of the world that would lead one to further refine one's taxonomy of kinds. If different mechanistic entanglements carry different implications for the taxonomy of kinds, then further revision is possible in each case, and this arguably disqualifies each of them as an HPC kind.

One might also be tempted to privilege one kind of homeostatic mechanism over the others (constitutive mechanisms and etiological mechanisms, such as evolution by natural selection, are the most common choices). However, nothing in the causal structure of the world justifies the privilege. The rather obvious thing to suggest is that, in such cases of conflict, whether one should lump or split depends what one is trying to understand or to do. Pluralism about natural kinds would seem the most prudent option: there are as many kinds as there are distinct and dissociable mechanistic entanglements (Boyd, 1999, and Wilson, 2005, both embrace this kind of idea). What counts as a maximally refined scientific kind depends on which mechanistic entanglement is most relevant to the practical or scientific project in which one is engaged. (B4) should make this claim explicit.

This minor amendment introduces a perspectival element into the scientific HPC, but this need not threaten the realist motivations for the HPC. In each case, the objective causal structure of the world is called upon as a foundation for the taxonomy of natural kinds. However, if these considerations are right, the scientific HPC will be unable to settle disputes among those who disagree about the taxonomy of natural kinds because they attend to different mechanistic entanglements of the putative kinds in question. Psychiatrists interested in prevention might pick out different kinds than those interested in treatment, for example. The point of this section is that they needn't argue. The HPC view allows that they can each be right for their own purposes, so long as we include perspectival and pragmatic considerations in (B4). The same, of course, could be said for taxonomies of memory, emotions, and other mental modules.

6. Kinds of Mechanism

I argue above that human perspective enters into decisions about which mechanisms matter for splitting or lumping a putative kind. In this section, I raise the question of when two mechanisms are mechanisms of the same kind. The splitting strategy (read through the HPC lens) applies when two distinct kinds of mechanism explain different properties in a property cluster. The lumping strategy (read through the HPC lens) applies when the same kind of mechanism explains two putatively different property clusters. Both strategies rely on facts about types or kinds of mechanism, not particular mechanisms.

Consider the hippocampus, a brain region thought to house a mechanism involved in encoding declarative and spatial memories. Ramon y Cajal and Lorente de Nò made crisp and delicate drawings of hippocampal specimens, detailing the precise locations, shapes, and orientations of the neurons in this elegant structure.

If one focuses on the particular cells and structures in these specimens, their particular locations and shapes, their particular activities, their exact numbers, and so on, then no two hippocampi are identical. Even those in the same head or in the same location moment to moment are different in many respects. If the HPC account were to suggest that we should split kinds when the mechanisms differ in any of the myriad ways that any two mechanisms might differ—in the precise features of their

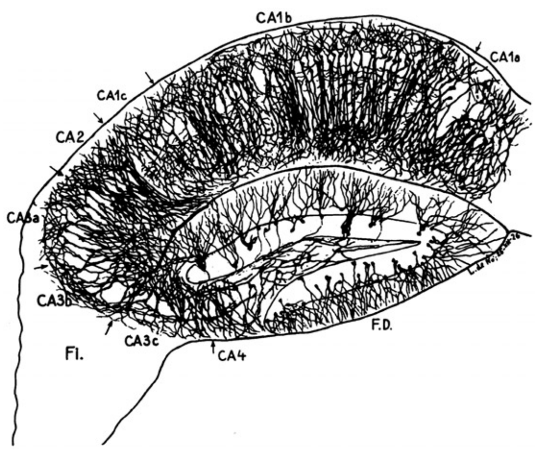

Figure 2 Sagittal section of the hippocampus of a mouse stained with the Golgi-Cox method (from Lorente de Nó, 1934).

component entities, activities, and organization—then there would be as many kinds of declarative memory encoding mechanism as there are individual hippocampi. Boyd explicitly wants to avoid this result. The scientific HPC must allow for variability among instances of the kind and for variability in the mechanisms that realize it.[6] To solve this problem, it will not suffice to demand that the differences in underlying mechanisms must make a causal difference to (or be otherwise explanatorily relevant to) the behavior of a mechanism as a whole because any *detectable* difference in the underlying mechanism must make some such causal difference. Likewise, one cannot object that the difference made is too small or insignificant, because such judgments depend on our assessment of which differences are too small to be relevant for our interests, not on the objective features of the causal structure of the world alone. Similarly, one cannot object that the causal differences are irrelevant for the property cluster, for this presupposes a method of lumping properties into clusters that is independent of discoveries concerning the mechanism that explains their clustering.

The dependence of the HPC account on the notion of kinds of mechanism raises the philosopher's worry that the metaphysical HPC view regresses. Property clusters are united in a kind because their clustering is explained by a single kind of mechanism. When are mechanisms mechanisms of the same kind? If one responds that mechanisms are mechanisms of the same kind when they are explained by a single kind of mechanism, the regress is transparent. If the answer is that mechanisms of the same kind are composed of the same kinds of entities, activities, and organizational features, then we need some way to unite entities and activities into natural kinds. Either way, we only stave off our ignorance of natural kinds a little longer.

With regard to the *scientific* HPC, the problem is not a regress but rather an apparent tension between the need to recognize the variability and multiple realizability of kinds in the special sciences, on the one hand, and the accommodation thesis on the other. Boyd intends HPC kinds to be multiply realizable; this is one of the crucial contrasts between the HPC view and essentialism. The same kind of mechanism (as described in an abstract model or schema) can be realized with variable components and variable organizational features. Presumably the variability in underlying mechanisms (or environmental constraints) accounts for the variability in the properties of the cluster instantiated in any member of the kind and in the specific causal relations that the member of the kind has with its environment. If the kind is variable in its properties, it is bound to be variable in at least some of its causal relations. If the underlying mechanism is different, then there are bound to be differences in the ways that the mechanism behaves or, at least, in the ways it responds to interventions. (B4) would seem to demand splitting the cluster so that membership in the kind underwrites the practices of prediction, explanation, and control across members of the kind. If kinds are to be accommodated to the objective mechanistic structure of the world, then one would appear to be forced to split kinds on the basis of even minor (even trivial) differences in property clusters or mechanisms.

What, then, is the appropriate degree of abstraction to use in characterizing a kind of mechanism? Characterizing the mechanism very abstractly potentially glosses over sub-kinds of mechanism. Characterizing the mechanism in maximal detail threatens to make each particular mechanism a kind unto itself. Between these is a continuum of degrees of abstraction at very many points along which one could identify a mechanism and a corresponding natural kind.

Consider some other schemata of the hippocampus.

The standard diagram of the hippocampal trisynaptic circuit illustrated in every introductory neuroscience textbook is more abstract than Ramon y Cajal's drawings. It captures the rough spatial organization of excitatory synapses in the hippocampus. It leaves out most of the excitatory neurons, all of the inhibitory neurons, all of the glial and support cells, and many of the details about their specific spatial organization. Understood so abstractly, such diagrams might as easily represent any trisynaptic circuit in the central nervous system. It is not detailed enough to characterize the specific workings of the memory encoding mechanism.

The flow diagrams in computational models of the hippocampus abstract away from neurons and their connections almost entirely, and instead represent abstract functional relationships among subregions of the hippocampus (such as the dentate gyrus, the CA1, and the CA3 regions). Different regions of the hippocampus are represented as computing distinct functions, such as rotation, shifting, filtering, and synchronization. As it stands, this schema need not be applied to neurons, to human hippocampi, or even to biological organisms. It could apply to any system that shares its abstract control structure. If one were to follow a lumping strategy with respect to this description of the mechanism, one would have to say that anything exhibiting this structure is part of the same natural kind. One could abstract from this diagram even further, perhaps describing the mechanism in the hippocampus as a sequential Hopfield net, or as a filter in an emulator, or in terms whatever information-processing function one thinks it computes (just as Marr, 1969, described part of the visual system as computing the Laplacian of Gaussian kernel).

These three schemata (pictorial images of specimens, tri-synaptic flow diagrams, and computational models) have different, though overlapping, extensions. Two mechanisms count as mechanisms of the same kind according to one schema and as

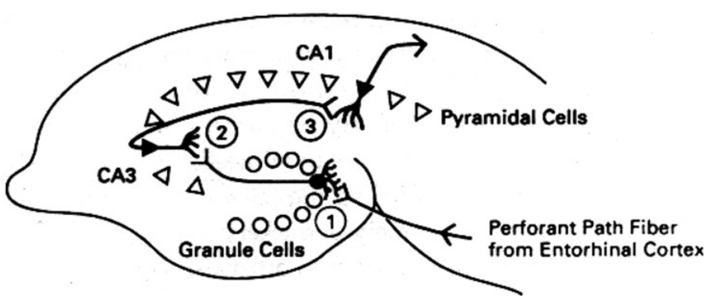

Figure 3 Schematic hippocampus.

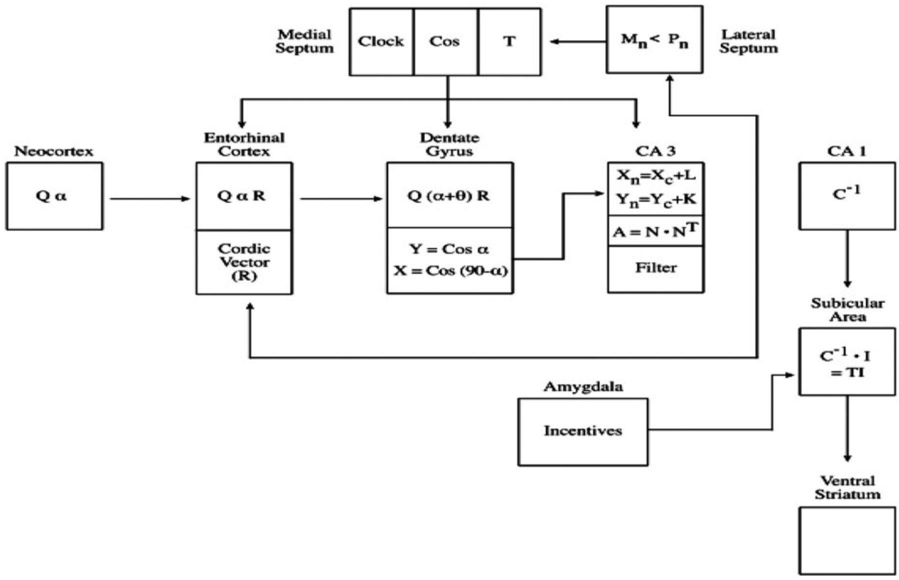

Figure 4 Computation in the Hippocampus.

mechanisms of different kinds according to another. To which of these mechanistic structures should the taxonomy of natural kinds be accommodated (as (B4) recommends)?

The defender of the scientific HPC view could argue that one of these schemata is objectively at the appropriate degree of abstraction for the purposes of refining our scientific taxonomy. I have no argument that this response must fail, but it seems unlikely to succeed for the simple reason that each of these schemata makes a nonredundant contribution to our ability to predict, explain, and control what the hippocampus does. Where more precise schemata capture relevant distinctions among the mechanisms classed together by the abstract schemata, the abstract schemata reveal regularities in the behavior of the hippocampus that are invisible in the more precise schemata. Choosing any one schema as maximally refined would thus involve defining an appropriate balance between the schema that generalizes over large numbers of mechanisms and the schema that acknowledges the potentially significant differences in the precise details of the different mechanisms. What makes an objective solution seem implausible is that the appropriate balance between generality and precision is just the kind of thing that depends on what *we* want to *do* with the schema. For some purposes (surgery, for example) precision is of the utmost importance. For other purposes (such as building an abstract computational model) generality is more important.

A second, more promiscuous, option is to allow that these schemata pick different, yet partially overlapping kinds of mechanism. While I find this inclusive approach the most appealing, it is difficult to square with (B4). Where there is multiple realization, there is always by definition further basis for splitting. Given Boyd's requirement and the consideration of this section, there will be a multitude of incompatible answers to the question "is X a natural kind?" depending on the appropriate degree of abstraction for the problem at hand. This poses no problem for conventionalists who allow that for some purposes, X should be treated as a natural kind, and for others, it should not. But this strategy is not available to the defender of the scientific HPC view. In the dialectic with which I opened this paper, the defender of the HPC view appeals to mechanisms in order to explain why a given property cluster is useful for the purposes of prediction, explanation, and control. But this promiscuous strategy reverses the direction of fit. It accepts mechanism schemata of different degrees of abstraction because they are useful for the purposes of prediction, explanation, and control. At this point what work is the appeal to mechanisms in (B2) doing for the scientific HPC? A simple causal view of natural kinds could fill this role just as well.

To summarize this second claim, scientists describe mechanisms at different degrees of abstraction. Depending on which degree of abstraction one chooses, the lumping and splitting strategies deliver different results. This is because whether two mechanisms are mechanisms of the same kind depends upon which grain of abstraction one chooses to describe them. If there is no objectively appropriate degree of abstraction for typing mechanisms, then judgments about whether two mechanisms are mechanisms of the same kind rely ineliminably on judgments by people (in concert) about the appropriate degree of abstraction required for the problem at hand. By extension, judgments about how to use lumper and splitter strategies to refine a scientific taxonomy rely ineliminably on judgments by people (in concert) about the appropriate scope and precision required to predict, explain, and/or control for a particular purpose. For biologists, onions are lilies. For flower gardeners, they are not.

7. Perspective and the Boundaries of Mechanisms

The third challenge to the HPC concerns the boundaries of mechanisms. Which entities, activities, and organizational features are part of a mechanism (or kind of mechanism) and which are not? Descartes and other Cartesian physiologists noted that a mechanical world has no resources internal to itself to identify the boundaries of machines, organs, and organisms (see Des Chene, 2000). In the absence of any purposes, goals, and forms in nature, there is no principle for dividing the organism into working parts. Contemporary mechanists face a related challenge.

Consider an action potential at a given location along the axon of a neuron. The mechanism for this process, it is well known, involves membranes, ion channels, and ions. With an initial depolarization of the membrane, voltage-sensitive sodium (Na^+) channels open, allowing Na^+ to diffuse into the cell, thereby raising the membrane voltage. The Na^+ channels inactivate as voltage-dependent potassium (K^+) channels open, allowing K^+ to diffuse rapidly out of the cell and thereby dropping the membrane voltage. Then K^+ channels inactivate. By virtue of what objective feature might one say that these are all components in the same mechanism? How does one identify where one mechanism ends and another begins?

One might, for example, equate the boundaries of mechanisms with compartmental boundaries. Some mechanisms are entirely contained within physical compartments, such as a nucleus, a cell membrane, or skin. Transcription happens within the nucleus (typically), and translation occurs in the cytoplasm (typically). However, mechanisms frequently transgress compartmental boundaries. The mechanism of the action potential, for example, relies crucially on the fact that some components of the mechanism are inside the membrane and some are outside. Some flow across that boundary.

Some Cartesian mechanists argue that mechanisms are bounded by contact among the parts or by the fact that the parts move together. Few contemporary biologists or neuroscientists endorse the idea that causation requires contact action (think for example of the attractive and repulsive forces involved in the action potential mechanism), and most parts of a mechanism interact with entities outside of the mechanism (as when a Na^+ ion collides with a passing protein, mitochondrion, or membrane). Parts of mechanisms often move in separate directions (as do the Na^+ and K^+ ions in the action potential). Some mechanisms are more ephemeral than others. For example, in many biochemical cascades, the relevant reactions could happen anywhere in the cytoplasm. Such mechanisms lack stable spatial relations, and they could not be picked up and carried from one place to the next. Stable spatial relations, I conclude, are neither necessary nor sufficient to define the boundaries of mechanisms.

Herbert Simon (1969) and Bill Wimsatt (1974) define the boundaries of mechanisms in terms of the *intensity of interaction* among components. The idea is that interactions are stronger within the boundaries of a mechanism than they are between the mechanism and its environment. Even supposing that we have a workable notion of the strength of a causal interaction, this proposal does not recover our scientifically informed intuitions about the boundaries of many mechanisms. The action potential depends strongly on background conditions such as the maintenance of the resting potential, the operation of the Na^+/K^+ pump, the presence of ATP, and other factors that are typically treated as outside boundaries of the action potential mechanism. The generation of action potentials also has long-term effects on blood oxygenation, glucose metabolism, membrane turnover, the packaging of neurotransmitters, and protein synthesis. Again, these activities are not typically included in the mechanism; they either produce no changes in the other components of the mechanism, or the changes they produce make no difference to action potentials. Still, their causal ties with the mechanism are as strong as any of the causal ties among the mechanism's components.

The boundaries of mechanisms therefore cannot be delimited by spatial propinquity, compartmentalization, and causal interaction alone. This is because the spatial and causal boundaries of mechanisms depend on the epistemologically prior delineation of *relevance boundaries*. But relevance to what? The answer is: relevance to the phenomena that we seek to predict, explain, and control. Within the boundaries of a mechanism are all and only the entities, activities, and organizational features relevant to the phenomenon selected as our explanatory, predictive, or instrumental focus.[7] Consequently, the boundaries of a mechanism depend on our choice of the explanandum phenomenon (that is, our choice of the property cluster) and on the way that we choose to describe that phenomenon or property cluster. If so, then the mechanistic structure of the world depends in part upon our explanatory interests and our descriptive choices at the level of property clusters. Stuart Kauffman (1971) claims there are many ways to break an organism into parts and that one will break it into different parts depending on what one takes the organism to be doing. We can generalize the point by saying that there are many ways to break the world into mechanisms depending upon what one thinks needs to be explained and on how one characterizes what needs to be explained.

This is a challenge not only because of the intrusion of perspective into the boundaries of mechanisms, but more fundamentally because on the HPC account, it is the explanandum phenomena—that is the property clusters—that mechanisms are supposed to vouchsafe as kinds (see Wilson, 2005, pp. 56 & 118). If what I have said is correct, components are clustered into mechanisms in part because of the property clusters they explain. The conventionalist will wonder again what objectivity the search for mechanisms adds over and above the usefulness of the property cluster for the purposes of prediction, explanation, and control.

8. Conclusion

The scientific HPC account is supposed to provide an alternative to conventionalism by grounding taxonomic judgments in objective facts about the mechanistic structure of the world. I argue that human perspectives and conventions enter into judgments about how mechanisms should be typed and individuated. This raises a challenge to defenders of the HPC account either to find an objective basis for taxonomizing the mechanistic structure of the world, or to argue that these perspectival intrusions into the accommodation process do not threaten the realist objectives that motivate the belief in natural kinds in the first place.

Clearly something is right about the scientific HPC view. If the idea of a natural kind is to have any application in the special sciences, something like the HPC view is the most promising option available at present. The above challenges are designed to highlight how one might explicitly develop a more pluralistic HPC view. The cost of such liberalization, however, is that there are many disputes about the taxonomies of the special sciences that will not be resolved by the amended HPC. It will not tell us in any definitive way how many kinds of memory there are, whether the emotions and concepts are natural kinds, or how to refine the DSM. Instead, there will be multiple incompatible answers to these questions depending on which mechanism one attends to, on how one describes the phenomenon, and on where one draws the boundaries of the mechanism. Perhaps this is as it should be. On a more pluralistic conception of special science kinds, such taxonomic problems are not so much solved as dissolved. Perhaps such conflicts arise from the flawed assumption that there must be a uniquely correct and orderly taxonomy of the mechanistic structure of the world. If the above considerations are correct, our interests and objectives contribute partly, but ineliminably, to the kinds of mechanisms we find in our world.

Acknowledgments

I thank Ken Aizawa, Matt Barker, Jim Bogen, Peter Carruthers, Lindley Darden, Carl Gillett, Don Goodman, Todd Grantham, David Kaplan, Ken Kendler, Max Kistler, Peter Langland-Hassan, Tom Polger, Samuli Pöyhönen, Georges Rey, Sarah Robins, Eric Seidel, Dan Weiskopf, Rob Wilson and Petri Ylikoski for helpful discussion and feedback on early drafts of this paper. Any mistakes that remain are mine alone. Thanks also to Pamela Speh for redrawing the figures, to Ben Graham for editorial assistance, and to Tamara Casanova, Mindy Danner, and Kimberly Mount for invaluable administrative support.

Notes

[1] Duprè (1993) argues for ‘‘promiscuous realism,’’ the view that there are many inconsistent yet legitimate ways of classifying things into ‘‘natural’’ kinds. Duprè’s promiscuity thesis outruns anything considered here. Here, I grant that natural kinds are in some way restricted to those that figure in ‘‘important’’ causal generalizations and that are underwritten by mechanisms. Boyd (1999, p. 159–162) largely embraces Dupre’s promiscuous realism, holding, for example that ‘‘lily’’ is a legitimate horticultural/landscaping kind, even if it is not a kind recognized by contemporary biological taxonomy. (See Wilson, 1999, for an effort to defend a stronger form of realism about HPC kinds). The present challenge is that this restriction to causally important kinds (accommodation to the instrumental demands of some disciplinary matrix or another) is consistent with a level of promiscuity incommensurate with the normative ambitions for the HPC view as an arbitrator among taxonomies of kinds in the special sciences.

[2] Jaegwon Kim (1992) gives the example of jade, a conventional kind that includes both jadeite and nephrite. Membership in the kind ‘jade’ does not guarantee that a particular sample will exhibit the same physical and chemical properties, respond in like ways to similar interventions, or figure in the same explanatorily and causally significant generalizations.

[3] Keil’s project is to develop a psychological theory of how organisms chunk the world into kinds, not (primarily) to develop a normative account of how they ought to, which is one of the objectives of philosophical defenders of the HPC.

[4] This manner of speaking is admittedly awkward. Boyd developed the HPC view in order to argue that species are natural kinds. He does not consider the kinds of physiology and psychology, such as the parts of animals and their characteristic behaviors, though it is clear that he intends the account to have such application. Other defenders of the HPC view have made this extension explicit (Wilson et al., forthcoming).

[5] Thanks to Ken Kendler for this example.

[6] This problem cannot be blocked by requiring, in addition, that there must be a difference in the memory phenomenon itself. Part of characterizing the phenomenon is to characterize the conditions of breakdown and intervention. This is why such research strategies are useful for shaping the space of possible mechanisms (see Craver, 2007, chapter 4). As noted above,

different mechanisms, if detectably different in their entities, activities, and organizational features, will respond differently to some interventions.

[7] For an account of this constitutive variety of explanatory relevance, see Craver, 2007, chapter 4.

References

Bechtel, W., & Abrahamsen, A. (2005). Explanation: A mechanistic alternative. *Studies in History of Philosophy of the Biological and Biomedical Sciences*, *36*, 421–441.

Bechtel, W., & Richardson, R.C. (1993). *Discovering complexity: Decomposition and localization as strategies in scientific research*. Princeton, NJ: Princeton University Press.

Boyd, R. (1989). What realism implies and what it does not. *Dialectica*, *43*, 5–29.

Boyd, R. (1991). Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philosophical Studies*, *61*, 127–148.

Boyd, R. (1997). Kinds as the 'workmanship of men': Realism, constructivism, and natural kinds. In J. Nida-Rumelin (Ed.), *Rationality, realism, revision: Proceedings of the 3rd international congress of the Society for Analytical Philosophy* (pp. 52–89). New York: Walter de Gruyter.

Boyd, R. (1999). Homeostasis, species, and higher taxa. In R.A. Wilson (Ed.), *Species*. MIT Press: Cambridge.

Boyer, P. (2003). Religious thought and behaviour as by-products of brain function. *Trends in Cognitive Sciences*, *7*, 119–124.

Carruthers, P. (2004). The mind is a system of modules shaped by natural selection. In C. Hitchcock (Ed.), *Contemporary Debates in Philosophy of Science* (pp. 293–311). Malden, MA: Blackwell.

Cartwright, N. (2001). Modularity: It can and generally does fail. In D. Costantini, M.C. Galavotti, & P. Suppes (Eds.), *Stochastic dependence and causality* (pp. 65–84). Stanford, CA: CSLI Publications.

Churchland, P. S., & Sejnowski, T. J. (1992). *The computational brain*. Cambridge, MA: MIT Press.

Craver, C. F. (2007). *Explaining the brain: Mechanisms and the mosaic unity of neuroscience*. Oxford: Oxford University Press.

Darden, L. (2006). *Reasoning in biological discoveries*. Cambridge, MA: Cambridge University Press.

Des Chene, D. (2000). *Spirits and clocks*. Ithaca, NY: Cornell University Press.

Dupré, J. (1993). *The disorder of things: Metaphysical foundations of the disunity of science*. Cambridge, MA: Harvard University Press.

Fodor, J. (1983). *Modularity of mind*. Cambridge, MA: MIT Press.

Glennan, S. (1996). Mechanisms and the nature of causation. *Erkenntnis*, *44*, 49–71.

Glennan, S. (2002). Rethinking mechanistic explanation. *Philosophy of Science*, *69*, 342–353.

Griffiths, P. E. (1997). *What emotions really are: the problem of psychological categories*. Chicago: University of Chicago Press.

Griffiths, P. E. (1999). Squaring the circle: Natural kinds with historical essences. In R. A. Wilson (Ed.), *Species*. MIT Press: Cambridge.

Lorente de Nó (1934) from figure 2.

Kauffman, S. A. (1971). Articulation of parts explanation in biology and the rational search for them. In R.C. Buck & R.S. Cohen (Eds.), *PSA 1970*. Dordrecht: Reidel.

Keil, F. C. (2003). Categories, cognitive development and cognitive science. In D. H. Rakison & L. M. Oakes (Eds.), *Early category and concept development: Making sense of the blooming buzzing confusion*. New York: Oxford University Press.

Kornblith, H. (1993). *Inductive inference and its natural ground*. Cambridge, MA: MIT Press.

Machamer, P. K., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science*, *57*, 1–25.

McKusick, V. A. (1969). On lumpers and splitters, or the nosology of genetic disease. *Perspectives on Biology and Medicine*, *12*, 298–312.

Marr, D. (1969). A theory of cerebellar cortex. *Journal of Physiology*, *202*, 437–470.

Pinker, S. (1997). *How the mind works*. New York: Norton.

Simon, H. (1969). *The sciences of the artificial*. Cambridge: MIT Press.

Ramon y Cajal, S. (1911). *Histologie du système nerveux de l'homme et des vertébrés*. In. N. & L.W. Swanson (Trans.). Oxford: Oxford University Press. 1994.

Reichenbach, H. (1956). *The direction of time*. Berkeley, CA: University of California Press.

Salmon, W. C. (1984). *Scientific explanation and the causal structure of the world*. Princeton, NJ: Princeton University Press.

Schacter, D. L., & Tulving, E. (1994). What are the memory systems of 1994? In D. L. Schacter & E. Tulving (Eds.), *Memory systems 1994* (pp. 1–38). Cambridge, MA: MIT Press.

Skipper, R. (1999). Selection and the extent of explanatory unification. *Philosophy of Science (Proceedings)*, *66*, 196–209.

Wilson, R. A. (1999). Realism, essence, and kind: Resuscitating species essentialism? In R. A. Wilson (Ed.), *Species*. Cambridge, MA: MIT Press.

Wilson, R. A. (2005). *Genes and the agents of life*. Cambridge: Cambridge University Press.

Wilson, R. A., Barker, M., & Brigandt, I. (forthcoming). When traditional essentialism fails: Biological natural kinds.

Wimsatt, W. (1974). Complexity and organization. In K. F. Schaffner & R. S. Cohen (Eds.), *PSA 1972: Boston Studies in the Philosophy of Science* (Vol. 2, pp. 67–86). Dordrecht: Reidel.

Woodward, J. (2003). *Making things happen*. Oxford: Oxford University Press.

Woodward, J., & Hausman, D. (2004). Modularity and the causal markov condition: A restatement. *British Journal for the Philosophy of Science*, *55*, 147–161.

Zachar, P., & Kendler, K. (2007). Psychiatric disorders: A conceptual taxonomy. *American Journal of Psychiatry*, *164*, 557–565.