Mechanisms and Explanatory Realization Relations*

Thomas W. Polger
University of Cincinnati
thomas.polger@uc.edu
http://homepages.uc.edu/~polgertw

1. The Times, They Are a Changin'

My topic is the confluence of two recently active philosophical research programs. One research program concerns the metaphysics of *realization*. Within the literature on realization there is substantial disagreement about even the general outlines of a theory. Occasionally it seems that the only common ground is that realization is a dependence relation that sometimes or always relates entities that figure in different explanatory schema, such as those of the special sciences and those of more fundamental sciences. The other research program concerns scientific explanation in terms of *mechanisms*. According to the advocates of the "new mechanism" there is a distinctive and important kind of scientific explanation that is generated by discovering and describing causal mechanisms. There is some dispute about whether this kind of explanation is genuinely distinctive. And there is much dispute about how exactly to characterize mechanisms. But even most critics of the new mechanism do not wish to replace it with old fashioned deductive-nomological explanation.

* Portions of this work were presented at the 2009 Kazimierz Naturalism Workshop in Kazimierz-Dolney Poland, and at the 2010 Philosophy of Neuroscience and Multiple Realization Workshop in Birmingham, Alabama. The research and those trips were made possible in part by the support of the Charles P. Taft Research Center at the University of Cincinnati. I am indebted to the organizers and audiences at those events, especially Konrad Talmont-Kaminski and Jacqueline Sullivan. And, for numerous conversations about realization, thanks to Ken Aizawa, John Bickle, Carl Craver, Ron Endicott, Carl Gillett, Doug Keaton, Robert Richardson, Larry Shapiro, and Rob Skipper.

In the philosophical literature these two research programs come together most explicitly in the discussion of the so-called special sciences, their relation to one another, and their relation to more basic sciences. One need only notice the similarities in the examples discussed and diagrams drawn to be attracted to the idea that realization and mechanism go together, providing an attractive metaphysical and explanatory package for understanding the special sciences and perhaps much else. The ontology of realization seems well suited to the epistemology and methodology of mechanistic explanation, and vice versa.

It is possible to doubt each piece in the story told above: to doubt that there is any realization relation, to doubt that mechanistic explanation is distinctive or useful, to doubt that realization and mechanism are complementary. Elsewhere I have argued that metaphysical functionalism, a realization-based ontology of the special sciences, gets illicit support from its conflation with the explanatory theses of mechanism (2004). But I will not be addressing those questions here.

Rather, I will address some questions that arise if we agree that realization and mechanism are at least interesting or useful, and if we think that they can be brought together. It is enough that some philosophers have taken this approach, for what I will argue is that once one takes this approach it introduces some constraints on theories of realization and mechanism. If they are connected, they are mutually constraining. One cannot just hook up any account of realization with any account of mechanistic explanation. In particular, I will argue that a well-known account of realization, due to Carl Gillett, is incompatible with a well-known account of mechanistic explanation, due to Peter Machamer, Lindley Darden, and Carl Craver (2000, hereafter MDC). This is

surprising, not least of which because Gillett has cited MDC's work as evidence that his account of realization is the right way to think about realization in the sciences.

In the next section I will lay out the issues surrounding realization and mechanism in somewhat more detail. This will allow me to be more precise about the question of their interrelations, which are taken up in the subsequent section. Then we will sketch several approaches to realization, and I will argue that some of them complement the new mechanistic philosophy better than others.

2. Realization, Mechanism, and the Metaphysics of Science

Let us now say, in somewhat more detail, what is common ground among various theories of realization. According to the received view in contemporary philosophy of science and mind, special science entities, such as the entities that figure in psychological explanations, are realized by but not identical to (more) basic physical entities.[1] This view, often called realization physicalism, is sometimes restricted to one class of entities (e.g., just psychological entities, or just economic entities); but is sometimes entirely unrestricted, applying to all entities. In the latter case, realization physicalism is the comprehensive doctrine that all entities (or all actual entities) are either basic physical entities or are realized by basic physical entities (e.g., Lewis 1994, Poland 1994, Melnyk 2003).

Whether in the restricted or unrestricted form, the thesis is meant to be ontological: one entity or set of entities ontologically depends on another. The dependence is non-causal and, to use Amie Thomasson's terms (1999), existential and

[1] Throughout I will use "entity" as a generic for the class that potentially includes objects, properties, events, processes, and so on. Some theories of realization are discriminate as to which kinds of entities are related by realization, and others are not.

constant. That is, one set of entities depends on another for its very existence, and the dependence continues throughout the existence. An example of this kind of dependence is the way that the picture on a television screen depends constantly and existentially on the illuminated picture elements that make it up. As the elements are removed or changed, so the picture is destroyed or modified. Moreover, as the example illustrates, ontological dependence relations of this sort are normally asymmetric and irreflexive. Finally, the relation is supposed to be invariant and counterfactual supporting. Indeed, realization is often (but not always) thought to hold with a modal strength stronger than nomological necessity.

Realization is also often, though not universally, associated with functionalism in philosophy of mind, the theory that mental entities are realized by but not identical to brain entities. And realization is most often appealed to in order to account the multiple realization of mental states—the purported fact that the same mental entity kinds can be realized by various physical entity kinds. But these last characteristics of realization are, perhaps surprisingly, contentious.

The truth is that, for much of its life in philosophical discourse, realization has been an enigma. Until recently philosophers were generally satisfied to note that there would be some such relation as realization, without saying anything very detailed about the relation itself. In the next section we will return to the question of what the realization relation could be. That is, what is the invariant, asymmetric, ontological dependence relation that we're looking for?

More important, for the moment, is to notice the similarity between the features associated with realization as a dependence relation and the characteristics that are typical

of mechanisms as described by the new mechanists in philosophy of science. The new mechanists, including Machamer, Darden, and Craver (2000), Stuart Glennan (1996, 2002) and others, trace their recent lineage to the work of Robert Richardson and William Bechtel in their book *Discovering Complexity* (1993); and from there back to earlier work of William Wimsatt (1974, 1976). I'll take the MDC characterization to be representative. According to MDC, "mechanisms are entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination conditions" (2000: 3). The details of mechanism are at least as contentious as the details of various accounts of realization, and there will be no need to rehearse them here. The important thing at the moment is that mechanisms are usually conceived of as having a nested hierarchical structure, according to which "higher level" mechanisms depend operationally and existentially, on simultaneous (constant) activity of lower level mechanisms (Figure 1). An example, drawn from MDC and discussed by Gillett, it the claim that the mechanisms by which a mouse completes a spatial memory task involve a spatial map mechanism in the mouse's hippocampus. The hippocampal mechanism involves synaptic mechanisms of long-term potentiation; and those in turn involve the molecular mechanisms of an ion channel being opened in the synaptic membrane (Figure 2). According to this story, the dependence relation between the mechanisms is ontological: mechanisms are conceived of as entities. Despite the fact that mechanistic explanations may refer to both higher-level, same-level, and lower-level mechanisms, the relations between the mechanisms as entities are usually thought to be asymmetric and irreflexive. And, crucially, both the "vertical" and "horizontal" relations between and constitutive of mechanisms are invariant and counterfactual supporting.

The relationship between mechanisms appears to be at least very similar to the realization relation. And it is natural to wonder if that relation is, in fact, realization. If the relation is the same then it is natural to wonder whether realization can be explained in terms of mechanisms, or mechanisms in terms of realization.

One of these possibilities seems initially more plausible than the other. For, taken at face value, the theory of mechanisms and mechanistic explanations seems to presuppose rather than explain that mechanisms can stand in one or more potentially explanatory dependence relations to one another. It is not strictly speaking presupposed that those relations are asymmetric or irreflexive, although all of the familiar examples have that feature. (The mouse's spatial memory mechanism depends on long-term potentiation in hippocampal neurons, and not vice versa.) But it is presupposed that these relations are ontological and relatively invariant. Indeed most advocates of the new mechanism tout the theory's compatibility with interventionist accounts of causal explanation, as well similarly counterfactual-based models of the "vertical" relationship among mechanisms. For example, experiments using gene knock-out mice help establish the nested dependence of the spatial memory mechanism on the ion channel mechanism: Had the ion channels not opened, the long-term potentiation would not have occurred, and the memory would not have been stored.

If it is correct that the new mechanism presupposes rather than explains the ontological nature of the "vertical" inter-mechanism dependence relations, then it seems unlikely that we would be able to formulate a mechanistic explanation of realization. It will be much more promising, then, to hope that a correct account of realization can be used to explain these unanalyzed inter-mechanistic dependence relations than vice versa.

And this is exactly what Carl Gillett and Ken Aizawa purport to provide (2003, 2007; Aizawa and Gillett 2009).

Gillett and Aizawa are engaged in the project of understanding the metaphysics of science, which they characterize as "the careful, abstract investigation of ontological issues as they arise within the sciences and their findings, models, explanations, and so on" (Gillett 2007: 194-195). With specific reference to the examples of mechanistic explanation sketched above, Gillett writes:

> Over the last century, scientific investigations have given us detailed, integrated accounts of many natural phenomena through mechanistic explanations that explain one kind of entity using the very different kinds of entity that the sciences take to compose it. The neurosciences provide a particularly useful example because they quickly show how comprehensive such compositional relations are. (2007: 196)

By "comprehensive" here Gillett is pointing out the ubiquity of mechanistic explanations in the sciences, and particularly in the neurosciences. This is to emphasize the importance of having a good understanding of the dependence relations presupposed by mechanistic explanations, and thereby to highlight the kind of achievement that it would be to secure such an account.

Thus far we have seen how the ontological dependence relations presupposed by mechanistic explanations look quite similar to the realization relation. The question posed is whether the realization relation can be used to explain or ground those mechanistic relations. In particular, might it be the case that the relations among mechanisms just are realization relations? Gillett and Aizawa think that the answer is,

yes. And they also think that Gillett's "dimensioned" account of realization is up to the task. In the next section I will sketch some approaches to understanding realization relations, focusing on three distinctions that partially organize theories of realization. This will allow be to characterize, if only in broad brush strokes, some kinds of realization theories. In the subsequent section I will argue that if we think that the new mechanistic movement in philosophy of science is correct, and we think that the inter-mechanism relation is realization, then mechanisms have some features that provide constraints on realization relations. In particular, it will turn out that Aizawa and Gillett's account of realization is not well-suited to be the realization relation for mechanisms as characterized by MDC.

3. Approaches to Realization

Over the past decade there has been a flurry of philosophical work on realization relations. It goes without saying that this flurry of work has not produced any consensus about the correct theory, or even about the constraints on such a theory. For example, I have argued that a correct account of realization should at least make it possible that the relationship between physical systems and probabilistic automata—the relationship that Putnam called "realization" in the 1950s and 1960s—is a realization relation (Polger 2004, 2007). But some theorists flatly reject that requirement (Gillett YYYY). And other theorists posit an ambiguity between various kinds of realization relations, or uses of the term 'realization' (Endicott 2005; Wilson and Craver 2007; Piccinini 2009, forthcoming). Yet there is some agreement about the methodology for studying realization. Almost all of the theorists of realization agree with Aizawa and Gillett that

the context for our investigations is the metaphysics of science, in something like the way they characterize it. Because of that commitment, and a general naturalistic skepticism in this group about a priori philosophical analyses, there is general agreement that theories of realization will be evaluated for their fruitfulness. This is why Aizawa and Gillett aim to link their theory of realization to contemporary work on mechanistic explanation. Explaining the mechanistic dependence relation would surely be a win for their team. But what team is that?

Theories of realization, nearly absent from the philosophical literature not long ago, are now abundant. For present purposes it will be useful to make three sets of distinctions. Only the third distinction will play a role in the current argument, but it is important to get a sense of the terrain and also to see that the operative distinction is indeed orthogonal to some more familiar distinctions among theories of realization.

Perhaps the most familiar distinction among views of realization, going back to functionalist theories in philosophy of mind but not essentially tied to them, is between reductive and non-reductive approaches. In the philosophy of mind literature, the distinction between reduction and non-reductive views of realization was tied to "occupant" and "role" versions of functionalism (e.g., Block 1978.) Unfortunately, familiarity and clarity are not the same. As everyone knows, the topic of "reduction" in philosophy of science is vexed. And the distinction between so-called role and occupant theories is none too clear itself.[2] Furthermore, it will be of no use to say that theories of realization are reductive if there can be autonomous sciences that range over (non-

[2] One trouble is that the role/occupant distinction seems to apply only to what Shoemaker (1981) would call the core realizers, which are not metaphysically sufficient for what they realize. Once we elaborate the accounts so that the realizer is metaphysically sufficient for the realized, the role and occupant accounts may be coextensive. For serious discussion, see Keaton 2010.

fundamental) mechanisms, because the germane notion of autonomy is often already understood in terms of reduction. All of this confusion is tolerable and even appropriate in the context of trying to understand mechanism. For whether mechanistic dependence or mechanistic explanation are themselves reductive or non-reductive, either in ontological or explanatory senses, is itself a matter of great dispute (see, e.g., Craver 2007).

A distinction that is less familiar but more useful for our present purposes is between inter-level and intra-level accounts of realization. This distinction is the one that had driven Gillett's theorizing. Intra-level theories hold that realized and realizer properties are instantiated in the same individual; Gillett calls this the "flat" view (2002, 2003).[3] Inter-level theories hold that realization relates the properties of different individuals, canonically the properties of wholes with the properties of their parts; Gillett calls this the "dimensioned" view, and it is the one that he favors (2002, 2003). I favor a flat view of realization, but I will not rehearse my dispute with Gillett on that issue (Polger 2004, 2007; Polger and Shapiro 2008). In fact, it is worth noting that the dimensioned view is at least apparently more plausible as an account of "vertical" mechanistic explanations. For it is natural to think that mechanistic levels are compositional levels, and thus that the relation among mechanisms is a part-whole relationship. As illustrated in Figures 1 and 2, mechanism may appear to require an inter-level or dimensioned account of realization. On first pass, the dimensioned approach appears to have a big advantage.

[3] One distinction among theories of realization that I will not be pursuing here concerns the entities that are the relata to the relation. Here I speak in terms of properties; but Gillett and others usually formulate their views in terms of property instances. And I am inclined to view realization as crossing entity types, so that individuals can be realizers of kinds. What follows does not depend on these details, so I will talk in terms of properties out of convenience.

These first two distinctions give us a flavor for the variety of theories of realization.[4] But now I want to introduce a new distinction whose import has not been sufficiently appreciated. Because this distinction is new I will have to explain it in more detail than the preceding two distinctions. And the rationale behind my selection of labels for the views will come out only later. This third distinction between approaches to realization distinguishes what I will call, somewhat prejudicially, *descriptive* and *explanatory* accounts of realization.[5]

Let me begin by introducing what I call the descriptive approach. According to descriptive theories:

> (descriptive) Realization is the only metaphysically basic non-causal ontological dependence relation among properties [at different mereological levels].

The salient features of descriptive accounts are two: First, descriptive accounts are generally dimensioned (inter-level) theories. I hesitate to build that into the account, so I have bracketed the clause; but all the salient examples are inter-level cases. Second, descriptive accounts understand realization to be a single relation that accounts for all of the inter-level dependence relations among entities of a certain type (usually, as in the formulation above, among properties or property instances) when they occur at different mereological levels. This approach appears to be favored by Jeffry Poland (1994), and Sydney Shoemaker (2001, 2007). Both Jaegwon Kim (1998) and Andrew Melnyk (2003)

[4] Needless to say there are many more distinctions that could be drawn. See Polger 2004, 2007; Wilson and Craver 2007; Endicott 2005; and Keaton forthcoming-a, forthcoming-b, 2010.

[5] This distinction is similar to (or a flavor of) Keaton's distinction between exoneration and Q-realization (2010).

sometimes seem to endorse the descriptive approach, but it is unclear whether their apparent endorsements represent their considered views.

The most prominent and articulated version of the descriptive approach is Carl Gillett's account realization (2002, 2003, 2007; Aizawa and Gillett 2009). According to Gillett, "Given their interconnections, the various 'packages' of powers, properties, individuals, and mechanisms studied by lower-level sciences together compose the qualitatively different powers, property instances, individuals, and mechanisms studied by higher-level sciences" (2007: 201). On this view there is a distinctive "making up" relation for different entity types: comprising for powers, realization for properties, composition for individuals, and implementation for mechanisms, respectively. Realization is the basic property-relating making up (non-causal dependence) relation. Specifically:

> (Realization) Property instances F1-Fn in individuals s1-sn (or individual s*), realize a property instance G, in individual s* under background conditions $, if and only if the powers contributed by F1-Fn to s1-sn (or s*), which are constituents/parts of s*, together comprise the powers individuative of G, in s* under $, but not vice versa. (Gillett 2007: 202)

So realization is defined in terms of the comprising of powers when the individuals that instantiate the properties are parts (not necessarily proper parts) of the individual that instantiates the realized property.[6] The key thing to notice is that, for Gillett, realization

[6] *Comprising* is in turn defined in terms of the implementation of mechanisms: "Powers C1-Cn had by individuals s1-sn (or individual S*), comprise the power C*, had by individual s* under background conditions $, if and only if the mechanisms grounded by the triggering and manifestation of powers C1-Cn, under triggering conditions $t1-$tn and background condition $, would together implement the mechanisms grounded by the triggering and manifestation of C*, under triggering conditions $t1-$tn, and background condition $, but not vice versa" (Gillett 2007: 202).

is the making-up relation for properties when they occur in individuals that are related mereologically (by which I mean as part and whole, generically), as well as when the individuals are not related mereologically. Realization is the making up relation for properties regardless of what kinds of dependence relations hold between the individuals in which those properties are instantiated. This idea is inherent in the approach, and is explicitly guaranteed by Gillett's specification that the realizing properties may occur in s1-sn "which are constituents/parts of s*" or in s* itself.

The descriptive approach may seem uncontroversial until one considers the alternative, explanatory approach:

(explanatory) Realization is one among several ontological dependence relations that properties might stand in to one another.

According to the explanatory approach, some properties are realized by other properties, and some are not. Although it is not essential to the present discussion, we can note that advocates of explanatory views usually think of realization as being a relation among properties that stand in functional or role-playing relations, such as the property of having a certain inertia and the property of being a doorstop. Among those properties that are not realized are those that are composed (or that stand in whatever property relation goes along with material composition for particulars); those that are identified (in limiting cases, of course); and those that are constituted in the special sense of the term involving spatio-temporal overlap (e.g., Baker 2000), as the properties of the statue are constituted by but not mereologically composed from the properties of the statue-shaped clay.

The key difference between descriptive and explanatory approaches to realization is that explanatory approaches can be discriminate with respect to whether the objects

that instantiate the properties are in mereological relations, constitution relations, identity relations, and so forth. Only in some cases is it correct to say that the properties of one thing are realized by the properties of another. (For the moment we need not say which cases, only that it will not be all of them.) In contrast, the descriptive approaches are indiscriminate among those possibilities—all are examples of realization for the properties involved.[7]

I call the discriminate views explanatory because they can play an informative role in scientific explanations. If we have two properties, A and B, and we want to know what relation the properties stand in to one another, it can be informative and explanatory to learn that they stand in a realization relation. And it is precisely the distinctiveness of the relation that make it informative and potentially explanatory to learn this. In contrast, if the descriptive approach is correct, then it is almost always the case that properties are related by realization, so it is not informative or explanatory to learn that property B is realized by property A.

I say that it is only almost always the case that properties are related by realization on the descriptive view, because there are two notable exceptions: some properties may be eliminated, and some may be emergent. So in a scientific context in which elimination and emergence are live options, learning that property B is realized by property A is informative and explanatory. But it seems that those contexts are the exception not the rule. Most scientific contexts are those in which elimination and emergence are not live options, and what we want to know is what kinds for relations

[7] No doubt there are some "explanatory" (i.e., not descriptive, in the sense defined) accounts that do not draw exactly this distinction. The concern here is that descriptive accounts draw no distinctions, so are non-starters; whereas at least some explanatory accounts can make the distinction, and are therefore candidates for the "vertical" relation needed by mechanism.

hold among the properties in question. In that case, learning that they are realized (in the descriptive use) will only prompt a further question about *how* they are realized. And that is a question for which the theory of descriptive realization has no answer, and in which it plays not part. This is why I think of indiscriminate accounts of realization as merely descriptive. Their utility is exhausted when they affirm that there really are two properties (not elimination), and the realized property is not emergent. But that is only to state or describe the phenomena in need of explanation.

Even if it is wrong that most scientific contexts are not ones in which elimination and emergence are live options, it is still the case that most contexts *in which mechanistic explanations are offered* are not contexts in which elimination and emergence are live options. Mechanists are realists (or close enough) about multiple levels of mechanisms.[8] So if we want an account of realization that is going to help explain and ground mechanistic explanations, it will be little use to have a merely descriptive account of realization. For the descriptive account of realization only affirms what the new mechanist already assumes, viz., that there is some ontological dependence relation between the properties of mechanisms and the properties of their constitutive (in the generic rather than special sense) mechanisms. It is true, on this account, that the properties of mechanisms are realized by the properties of other things. But it is not informative or explanatory.

In this section I have distinguished a number of ways of understanding realization, and I focused on the distinction between descriptive and explanatory approaches. I have suggested that descriptive approaches are not good candidates for

[8] Some mechanists are not scientific realists in general; but general anti-realism and particular eliminativism are different theses. So it's fair to say that mechanisms are as real as anything else, on the new mechanist view.

understanding realization in the sciences, and for explaining and grounding mechanistic explanation in particular. At this point it is worthwhile to consider some reasons for resisting my arguments.

One might try to turn my complaint on its head. I argued that theories of realization that come from the indiscriminate approach, which I called descriptive, can not be used answer certain questions about why or how one set of properties ontologically depends on another. (With the exception of cases in which elimination or emergence are live options, in which case some information is given.) But this is no weakness in the account, one might claim. Rather this is just as it should be: philosophical theories of the metaphysics of science tell us only that there is some dependence relation, and describe it very abstractly. Scientific theories tell us, in detail, why and how those ontological relations hold in particular cases. So not only does the descriptive account survive my complaint, on second look my so-called explanatory approach is exposed as badly overreaching—of substituting metaphysics where science should tread.

I don't know of anyone who has advanced this kind of argument, yet it strikes me as tempting. But there are three reasons that this table-turning maneuver does not work. First, the descriptive accounts do try to explain some of the kinds of things that the explanatory accounts address, namely by distinguishing cases of realization from those of elimination and emergence. I think they should also distinguish other cases, in addition. But at any rate the descriptive accounts do not imply that these kinds of explanations should be left to science, or are somehow out of bounds for the metaphysics of science. Second, and more importantly, the kind of explanation that would be required to account

for the specific dependencies is precisely the kind of mechanistic explanation for which we are trying to find metaphysical grounding. We noted earlier that mechanistic explanation presupposes rather than explains the existence of "vertical" constituting ontological dependence relations. So we cannot look to mechanistic explanations to answer the questions that I claim a theory of realization should help to answer.

A third way of resisting the criticisms that I have so far made against the descriptive approach is to point out that even if it does not answer certain kinds of questions, the descriptive approach is nevertheless compatible with many answers to those questions. In particular it is compatible with their being some additional metaphysical theory (call it a theory of "realization*" if you must) that is explanatory in the way outlined above. This kind of response is akin to Robert Wilson and Carl Craver's idea that realization is a "servant to two masters" (2007), having distinct metaphysical and explanatory duties in different discourses. If so, maybe there are two different kinds of relation at play, and one unfortunate homonym to keep us all confused. What matters is that the descriptive approach captures at least part of one of those senses of "realization" that is useful in understanding or grounding mechanistic explanations.

This sort of response is surprisingly popular. It is surprising because philosophers generally prefer not to solve disagreements by postulating terminological ambiguities; and even more so when the term in question is universally agreed to be a technical term. But beyond surprise, there are two reasons to reject this move. First, there are independent reasons for supposing that some current and historical theories that invoke realization relations intend for one realization to do all the "jobs" of realization. The historico-philosophical story is complicated and has been only partially told; so I will rely

on it herein.[9] Second, the ambiguity-postulating response severely undercuts the interest of descriptive accounts of realization. At best, descriptive realization turns out to be one of several sorts; at worst, it is not even one of the interesting sorts that helps us to understand the scientifically kinds of ontological dependence among mechanisms. If descriptive realization is defended in this way then its victory will be Pyrrhic at best.

Now we come to the crux of the matter, the argument that the descriptive approach to realization isn't even one of the interesting sorts. More accurately: descriptive realization can't do the job that its proponents are interested in having it do. I will argue that the descriptive approach to realization cannot be the making-up relation for properties in the new mechanism.[10] The argument for this conclusion is easy to state but requires some explaining, and that is the task of the next section. Specifically, I argue that descriptive realization is insensitive to exactly the distinctions to which mechanistic explanations are sensitive. This is a serious problem for descriptive realization because its main virtue is supposed to be that it succeeds in doing this exactly this grounding work—the job of the metaphysics of science.

4. Mechanisms and Aggregates

Machamer, Darden, and Craver are not the only advocates of the new mechanistic philosophy of science. Nevertheless I will follow the formulations of mechanism and mechanistic explanation that they offer jointly and in their subsequent developments of the view. One reason for doing so is that their version of the view has been the most

[9] For a start, see my 2004 and 2007.

[10] I'm ignoring MDC's quirky ontology of entities and activities, with no mention of properties. I will not be criticizing descriptive accounts for assuming that mechanists will need to recognize properties or something like properties.

prominent and widely discussed. Another reason is that their characterization are explicitly cited by Aizawa and Gillett as supporting their favored account of realization, which I have grouped among the descriptive accounts. Finally, using the MDC account allows me to make my criticism quite plain. This last fact also raises questions about whether I am stacking the deck against the descriptive realization theorists by measuring their approach against a particular view of mechanisms that they might reject. This is a fair concern, but one that can be met; I will return to that worry at the end of the section.

Earlier I drew attention to the fact that descriptive approaches such as Gillett's are insensitive to whether the individuals that instantiate the realized and realizing properties are themselves related mereologically or in some other way. According to Gillett, realization is the fundamental property-relating making up relation. Gillett also holds that there is only one individual-relating making up relation, *composition*. So all cases of some individuals making up others are composition; and all cases of the properties of some thing or things making up the properties of another thing or things are *realization*. If there are different ways that individuals can make up one another Gillett's view is indifferent to them. And, consequently, if there are different ways that properties can make up one another then Gillett's view will be indifferent to those as well; for realization of properties is just part of the package that goes along with the making up of individuals.

Recall that, according to MDC, "mechanisms are entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination conditions" (2000: 3). One criticism that has been made of the new mechanistic philosophy is that it is vacuous because just about everything could count as

a mechanism (e.g., Psillos 2010). But even if *just about anything* could be a mechanism, it is not the case that *literally anything* could be a mechanism. Mechanisms may be abundant, but they are not ubiquitous. The reason, simply put, is that not every arrangement of entities and activities is "organized" and moreover organized so as to "produce regular changes" in a causal system. At the very least the new mechanists distinguish between mechanisms and mere aggregates (Craver 2001, 2002, 2007). But Gillett's account, like any descriptive account that takes realization to be the singular property-relating making up relation, is indifferent to whether the individual that instantiates the realized property is a mechanism or an aggregate. Consequently descriptive accounts cannot ground the new mechanistic philosophy of science according to which the ways that mechanisms are made-up is distinct from the ways that mere aggregates are made up. And this is why descriptive mechanism cannot be adequate to ground and explain the ontology of the new mechanism. One way to put the point is to say that mechanism is more fine grained than descriptive realization: mechanists distinguish "making up" relations that descriptive realization does not.

At this point we might well want to know more about the purported distinction between mechanisms and aggregates. Carl Craver (2001), following William Wimsatt (1976, 1986, 1997), draws the distinction by appealing to the idea that mechanisms are organized whereas aggregates are not. The conception of organization draws on a distinction between systems whose parts are interchangeable with one another and those that are not (Craver 2001: 58-59, emphasis original):

Suppose that a property Ψ of the whole S is a function of the properties $\{\Phi_1, \Phi_2, \ldots, \Phi_3\}$ of the parts $\{X_1, X_2, \ldots X_2\}$. Then a Ψ property of S is an aggregate of the Φ properties of Xs when:

(W1) Ψ is invariant under the *rearrangement and intersubstitution* of Xs;

(W2) Ψ remains *qualitatively similar* (if quantitative, differing only in value) with the addition or subtraction of Xs;

(W3) Ψ remains invariant under the *disaggregation and reaggregation* of Xs; and

(W4) There are no *cooperative or inhibitory* interactions among the Xs that are relevant to Ψ.

This is a conception of mere aggregates that lack organization, whose parts can be rearranged arbitrarily with no consequence for the aggregate property. For example, the mass of a pile of sand is a merely aggregate property. By extension, we can say that the pile of sand itself is a mere aggregate, not a mechanism because its pile-constituting properties are merely aggregate properties. (Put another way, the identity conditions for being a pile involve only mere aggregate properties.)

Mechanisms are not mere aggregates. The characteristic properties of mechanisms are not merely aggregate properties. Mechanisms are organized wholes, not mere wholes. What is interesting about mechanisms is precisely that they have properties other than those that an aggregate of their parts would have. An assembled lawn mower has the same aggregate properties as a pile of unassembled lawn mower pieces. But it also has many properties that the pile of parts does not; and those in fact are the

interesting properties that are practically useful and that stand in need of explanation or grounding.[11] What is perplexing about mechanisms is how we can explain what they do if their activity is not a mere sum of the activities of their parts. Of course we are not in practice stumped by the mechanistic properties of lawn mowers. But the metaphysics of science is not about providing novel mechanistic explanations, which in this case are already understood. Rather, it is about understanding or grounding those explanations. Moreover there are plenty of cases in which the first order explanatory question is not yet settled, and there it is more clear that we are seeking an informative answer to the question of how the properties of the whole are related to the properties of the constituents. As I understand it the mechanisms that allow ice skates to glide across ice are not well understood, for example (Chang 2006).

Gillett's example of the hardness of a diamond is useful here. The hardness of the diamond, in contrast to its mass, is not a mere aggregate property. This is because the hardness of the diamond depends not only on the presence of the carbon atoms but also on their arrangement. So at least (W1) fails for hardness, and quite likely (W4) as well. In contrast (W1)-(W4) all hold for the mass of the diamond. Yet for Gillett both hardness and mass are realized properties of the diamond. Gillett's descriptive account of realization tells us only *that* the properties of the diamond ontologically depend on their properties of its constituents. But it is indifferent to *how* the different kinds of diamond properties are made up, and therefore useless for explaining those variations.

The trouble for the descriptive approach is that it appears to be indifferent to exactly the distinction to which the new mechanistic philosophy of science intends to be

[11] Craver (2001: 59) compares the organization of the circulatory system to the (relative) disorganization of the molecules in a glass of gin.

sensitive. I conclude that descriptive accounts of realization like Gillett's cannot be the correct metaphysics of mechanistic science if the new mechanism is the correct explanatory framework.

One might wonder whether Gillett's account of realization is representative of the descriptive approach to realization, and whether Craver's distinction between mechanism and aggregates is representative of the new mechanistic philosophy of science. If not, then my critique is of limited application: it only shows that Gillett's account does not support Craver's, and no more.[12] And if those theories are representative of their kinds but the distinction between mechanisms and aggregates cannot be maintained, then this might seem to undercut my critique. Finally, if the accounts are representative and the distinction can be maintained but realization is supposed to apply to aggregates *rather than* mechanisms, then perhaps my critique could be deflected. Let us consider each of these objections in turn.

First, we can question whether Gillett's account of realization is truly representative of views that fall under what I am calling the descriptive approach. Because I introduced the descriptive approach precisely in terms of the indiscriminate feature that I then pointed out in Gillett's account of realization, it's a safe bet that Gillett's account is representative in this respect. A more serious concern is that Gillett's account could be the only descriptive account—that no other philosopher advocates a view according to which realization is indifferent to whether the realizer properties occur in the parts of the entity that instantiates the realized property, and to whether those are organized parts or aggregate parts. Earlier I indicated that I think this is not the case, that

[12] Though it should be said that if a prominent mechanistic theory lends no support to a prominent account of realization, that would be interesting enough in its own right.

at least Poland (1994) and Shoemaker (2001, 2007) also hold views according to which realization is a single and fundamental property-relating making-up relation. But there is not space to make that case herein, so suppose instead that I am wrong and Gillett is the lone descriptive theorist, in my sense. In that case the scope of my critique is more limited; but because Gillett's account is the best known in the current literature, the result is still significant. Moreover, I cannot see that it would turn out to be a good thing for Gillett if his account is iconoclastic in this way. This would mean that his approach to the metaphysics of science is distinct from everyone else's. This kind of disconnect would make it hard to assess Gillett's arguments in favor of his view and against his competitors' views, for it appear that they are simply engaged in different projects and consequently talking past one another. So I think it would be more charitable to suppose that Gillett's view is representative of a class of accounts of realization.

I will bundle the second and third objections together. These concern whether Craver's account of mechanism specifically (and the MDC account more generally) is representative of the new mechanistic philosophy of science, and whether the key distinction between mechanisms and aggregates (if it is characteristic of the mechanistic philosophy of science) can be maintained. The gist of these concerns is that if the mechanism/aggregate distinction is not central to the new mechanistic philosophy of science then the fact that descriptive accounts are insensitive to it is no mark against them. And, indeed, within the mechanisms literature a there is plenty of room for dispute over whether the MDC account should be preferred. Moreover, one of the hotly contested issues concerns whether mechanisms are distinct or distinctive phenomena that can be theoretically distinguished from aggregates, collections, and the like. One might

think that mechanisms are just the aggregates that happen to interest scientists, or the complicated aggregates, or some such. Or, one might think that all things (including aggregates) are mechanisms for something or other, so talk of mechanisms is just highfalutin talk about causal explanation.

If mechanisms can not in general be ontologically distinguished from aggregates or if mechanistic philosophy of science does not require any such hard and fast distinction, then this seems like bad news for my critique of the descriptive approach to realization. But in fact the situation is not so clear cut. First, mechanisms appear in Gillett's basic ontology; so it could not be a good thing for him if there are no mechanisms or if mechanisms cannot be consistently distinguished from other entity types. But maybe that is incidental. More directly, if there are no mechanisms or mechanisms cannot be distinguished from aggregates then it cannot be a virtue of descriptive approaches to realization that they jive with mechanistic philosophy of science. Of course consonance with mechanism won't be a virtue for any other approaches, either; but that is just to say that there is no advantage to be gained in that vicinity.

Perhaps the more important point is that even if the mechanist neither needs or has any general distinction between mechanisms and aggregates, it will nevertheless be the case that there are different ways in which the properties of some things can "make up" the properties of others. One doesn't need the principled distinction between aggregates and mechanisms to recognize that the unassembled pieces in a pile of lawnmower parts contribute to the properties of the pile in a different way than those same pieces contribute to the properties of the lawn mower when it is assembled.

Whether the right characterization of the difference is as between aggregates and mechanisms, between unorganized parts and organized parts, or something else is (as it is said) academic. If it is only a particularist distinction that is made on a case by case basis, it will still be true that the descriptive approach to realization is insensitive to the distinction and is therefore unhelpful in explaining it, or in explaining the different ways that the macro-properties are "made up" in each case. Of course someone could always try denying that there is even a particular and case-by-case distinction to be made. One could try claiming that the assembled lawnmower is just like the pile, after all. Or one could try arguing that only one of them is a case of genuine making up at all, and therein lies the apparent difference. One could argue, following Peter van Inwagen (1990), that neither is a genuine case of making up, and so neither is a candidate for having realized properties. These moves are desperate if their only goal is to save an account of realization; and, more importantly, none of them would be welcomed by advocates of descriptive realization such as Gillett.

Finally, one might worry, assuming that the mechanism/aggregate distinction holds up, that I have located realization on the wrong side of the divide. You might think that realization is exactly the relation by which aggregates get their properties, and that descriptive realization is the correct approach to understanding that relation. A mechanism, if there are any, is an entity whose properties (or some of them) are "greater than the sum of its parts" (Craver 2001: 58). But this is just to say that the properties of mechanisms are not realized but are instead *emergent*. If so, the descriptive approach only applies to "realized" properties, the properties of aggregates—it need not have anything to say about the properties of mechanisms precisely because they are not merely

made up by the properties of their parts. So far from being indifferent to those properties, descriptive realization counts them as emergent. Moreover, if the new mechanists are nevertheless correct about the scientific importance of mechanisms, and if the descriptive approach is the right way to think about realization, then contrary to my earlier claims the question of whether there are emergent mechanistic properties is a live one and ubiquitous. In that case, again contrary to my early claims and the connotations of the labels I chose, it will normally be informative and explanatory to learn that a property is descriptively realized. For then mechanistic emergence is a perpetual and live alternative.

This response turns my critique on its head. If the properties of mechanisms are emergent rather than realized, then my critique of Gillett and the descriptive theorists misfires. But this only happens if we understand mechanisms in terms of emergence, and that is very controversial to say the least. Certainly, and maybe this is the most basic thing to say, that is not the line of reasoning that Gillett or any other realization theorist has advanced. Far from thinking of mechanisms as emergent, Gillett considers mechanistic explanation to be a kind of reduction. The mechanisms-as-emergent move would be quite odd in the present context. The question we set out to investigate, after all, was whether realization could be the ontological dependence relation that is presupposed by the new mechanists. According to the objection being considered, the answer is that it cannot because realization only applies to the properties of aggregates. That would be bad news for my critique of the descriptive approach to realization. But, again, it's hard to see how it could be good news for the advocates of the descriptive view, who claim that realization applies to the properties of mechanisms.

So the objections we have considered fall short. Whether mechanists must distinguish between aggregates and mechanisms is undecided. But they certainly can and do. And scientists and philosophers of science certainly can and do recognize and distinguish different ways that the properties of macro-objects are made-up, be they piles, lawn mowers, or brains. And that is enough for my critique to gain traction: Here is an explanatorily useful distinction to which descriptive approaches to realization are insensitive. The best that can be said for those approaches is that they do not tell the whole story, that some relations that are (after all) sensitive to the distinction will need to be marshaled in order to complete the explanations. But lacking the discriminatory resources, the descriptive approach to realization will not figure in those explanations.

5. Explanatory Realization

The alternatives to descriptive accounts of realization are explanatory accounts. I call these counts explanatory because they do or can figure in explanations of how the properties of macro-objects are made up: Some are realized. But some are not, because they are identical to, aggregates of, or constituted by (in the special sense) other entities.

The explanatory approaches to realization that I favor are those that construe realization in terms of functions, so that for one entity to realize another is for it to have a function or play a functional role (2004, 2007; Polger and Shapiro 2008). Because there are myriad notions of function on the table, this functional approach to explanatory realization is quite malleable; and much remains to be said about how best to flesh it out. I have said that the approach denies that there is any singular realization relation, replacing that idea with a family of functionally characterized notions of

realization—different realization relations for different kinds of functions (2004, 2007). What is crucial is that, at least for many notions of function, functions are not universal, and so many functional notions of realization have the capacity to discriminate between aggregates and organized mechanisms. One might even suppose that having a function *just is* a way of being organized that mere aggregates lack.[13]

Functional accounts of realization may prove to be problematic. I will not be defending them herein. The present conclusion is only that if we want an account of realization that will help us to ground and understand scientific explanation in terms of mechanisms, and if the new mechanists are right about the nature of mechanisms, then we will need some explanatory account of realization. That is, we will need an account of realization that discriminates between the ways that properties of mechanisms are made up and the ways that properties of aggregates are made up. This is what accounts like Gillett's cannot manage.

Here I have focused on the ways that the requirements of mechanism constrain accounts of realization. But once we see what accounts of realization are candidates for mechanistic "making up" relations, we can examine what consequences they have for the theory of mechanisms. It remains to be seen whether the total package of mechanism and explanatory realization is a useful one, and where it leaves us on traditional issues such as physicalism, multiple realization, reductionism, the causal efficacy of mechanisms qua mechanisms, or the autonomy of the so-called special sciences that explain mechanisms.

[13] Undeniably there are some notions of function that apply to everything, and which would therefore fail to discriminate between the assembled and unassembled parts of a lawn mower. So be it. What matters is that some notions of function are discriminate, and so can be explanatory. (See also note. 7.)

6. References

Aizawa, K. and C. Gillett. 2009. The (Multiple) Realization of Psychological and other Properties in the Sciences. *Mind & Language*, *24*: 181-208.

Baker, L. 2000. *Persons and Bodies: A Constitution View*. Cambridge: Cambridge University Press.

Bechtel, W. and R. Richardson. 1993. *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. Princeton: Princeton University Press.

Block, N. 1978. Troubles with functionalism. C. W. Savage (ed.), *Minnesota Studies in the Philosophy of Science, Vol. IX* (Minneapolis, MN: University of Minnesota Press)

Chang, K. 2006. Explaining Ice: The Answers Are Slippery. *The New York Times*, 21 Feb. 2006: D1.

Craver, C. 2001. Role functions, mechanisms and hierarchy. *Philosophy of Science* 68: 31-55.

Craver, C. 2002. Interlevel Experiments and Multilevel Mechanisms in the Neuroscience of Memory. Philosophy of Science, Vol. 69, No. 3, Supplement, vol. 2: S83 -S97

Craver, C. 2007. *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*. Oxford: Oxford University Press.

Craver, C. and L. Darden. 2001. Discovering Mechanisms in Neurobiology: The Case of Spatial Memory", in P. K. Machamer, R. Grush, and P. McLaughlin (eds.), *Theory and Method in the Neurosciences*. Pittsburgh: Pittsburgh University Press, pp. 112-137.

Endicott, R. 2005. Multiple Realizability". In *The Encyclopedia of Philosophy*. 2nd Edition. New York: Macmillan.

Gillett, C. 2002. The Dimensions of Realization: A Critique of the Standard View, *Analysis* 64, 4: 316-323.

Gillett, C. 2003. The Metaphysics of Realization, Multiple Realizability, and the Special Sciences. *The Journal of Philosophy,* Nov. 2003: 591-603.

Gillett, C. 2007. Understanding the New Reductionism: The Metaphysics of Science and Compositional Reduction. *The Journal of Philosophy*, CIV (4): 193-216.

Glennan, S. 1996. Mechanisms and the Nature of Causation. *Erkenntnis*, vol. 44: 49-71.

Glennan, S. 2002. Rethinking Mechanistic Explanation. *Philosophy of Science, Supplement* 69: S342-S353.

Keaton, D. forthcoming-a. Kim's Supervenience Argument and the Nature of Total Realizers. *European Journal of Philosophy*.

Keaton, D. forthcoming-b. Two Kinds of Role Property. *Philosophia*.

Keaton, D. 2010. Realization and Causal Role-Playing: An Essay on the Mind/Body Problem. Ph.D. Dissertation, University of Cincinnati.

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

Lewis, D. 1994. Lewis, David: Reduction of Mind. In Guttenplan (1994).

Lewis, D. 1999. *Papers in Metaphysics and Epistemology*. New York: Cambridge University Press.

Machamer, P., L. Darden and C. Craver. 2000. Thinking about Mechanisms. *Philosophy of Science* 57: 1-25.

Melnyk, A. 2003. *A Physicalist Manifesto: Thoroughly Modern Materialism*. Cambridge: Cambridge University Press.

Piccinini, G. 2009 . Computationalism in the Philosophy of Mind. *Philosophy Compass*, 4.3: 515-532.

Piccinini, G. forthcoming . The Mind as Neural Software? Understanding Functionalism, Computationalism, and Computational Functionalism. *Philosophy and Phenomenological Research*.

Poland, Jeff. 1994. *Physicalism*. New York: Oxford University Press.

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

Polger, T. 2007. Realization and the Metaphysics of Mind. *Australasian Journal of Philosophy*, 85 (2): 233-259.

Polger, T. and L. Shapiro. 2008. Understanding the Dimensions of Realization. *Journal of Philosophy*, CV (4): 213-222.

Psillos, S. forthcoming. The Idea of Mechanism. In P. McKay, F. Russo & J. Williamson (eds.) *Causation in the Sciences* (Oxford: Oxford University Press).

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

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

Shoemaker, S. 1981. Some Varieties of Functionalism. *Philosophical Topics* 12, 1: 83-118.

Shoemaker, S. 2001. Realization and Mental Causation, in C. Gillett and B. Loewer (eds.), *Physicalism and Its Discontents*. (Cambridge, UK: Cambridge University Press, 2001).

Shoemaker, S. 2007. *Physical Realization*. Oxford: Oxford University Press.

Thomasson, A. 1999. *Fiction and Metaphysics*. Cambridge: Cambridge University Press.

van Inwagen, P. 1990. *Material Beings*. Ithica: Cornell University Press.

Wilson, R. and C. Craver. 2007. Realization. In P. Thagard (ed.), *Handbook of the Philosophy of Science, Volume 12, Philosophy of Psychology and Cognitive Science* (The Netherlands: Elsevier, 2007).

Wimsatt, W. 1974. Reductive Explanation: A Functional Account. In R. Cohen et. al. (eds.) *PSA Proceedings 1974*. Dordrecht: Reidel.

Wimsatt, W. 1976. Reductionism, Levels of Organization and the Mind-Body Problem". In Globus et al (ed.) *Consciousness and the Brain*.

Wimsatt,W. 1986. Forms of Aggregativity. In A. Donagan, A . N. Perovich, and M. V. Wedin (eds.), *Human Nature and Natural Knowledge* (Boston : D. Reidel): 259-291.

Wimsatt, W. 1997. Aggregativity: Reductive heuristics for finding emergence. *Philosophy of Science*, Supplement Vol. 2, 1997: S372-384.

Figure 1. Hierarchy of mechanisms (adapted from Craver 2001: 66)

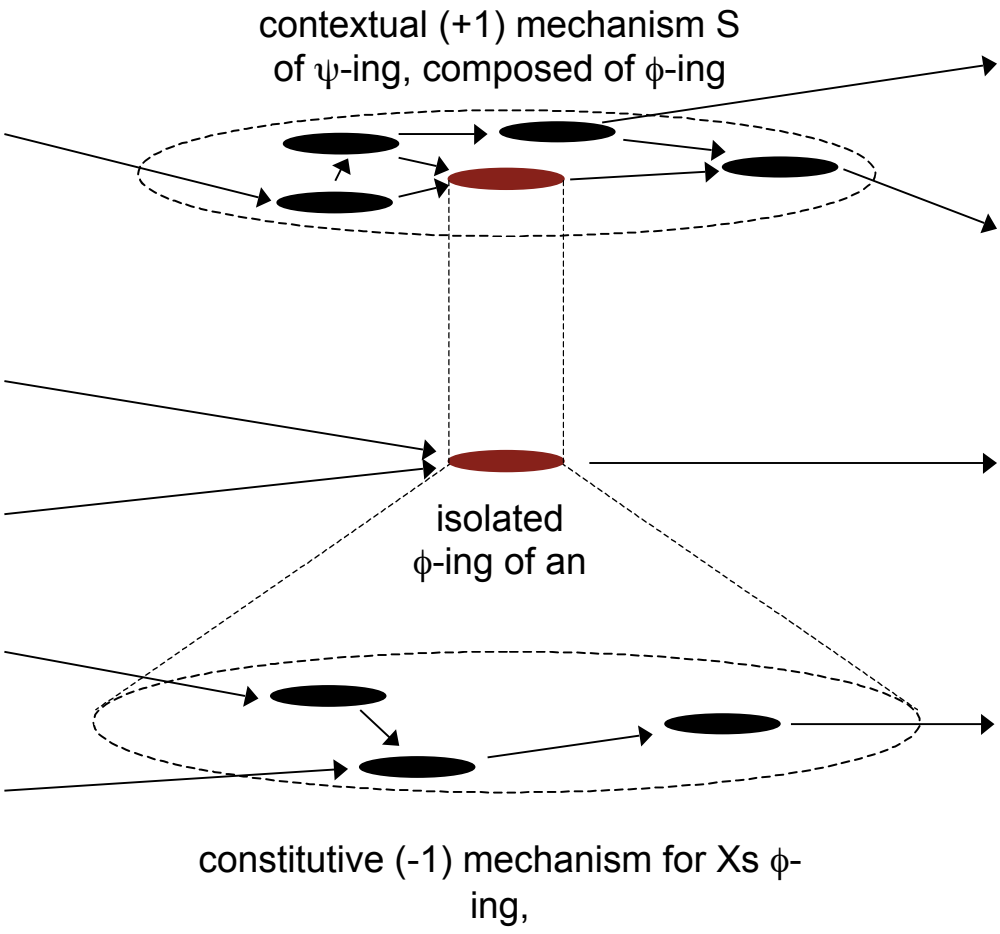

Figure 2. Mechanisms of spatial memory (from Craver 2002: S90)

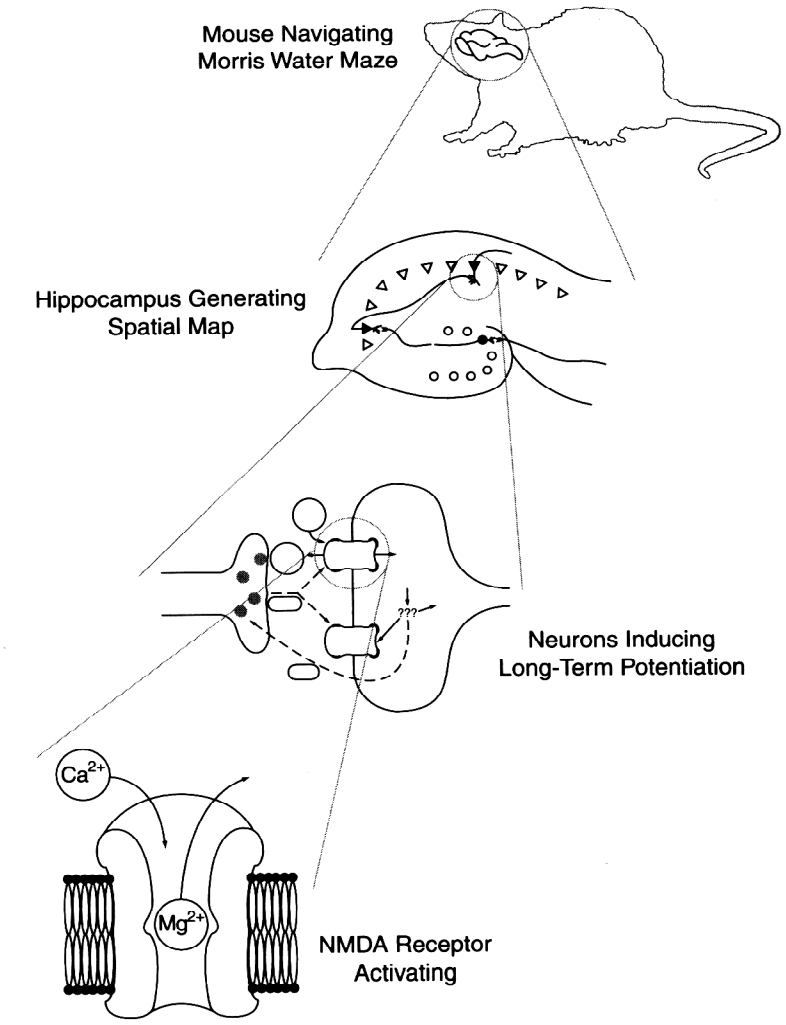

Figure 6.4 Levels in the hierarchical organization of the mechanism of spatial memory.