Synthese (2016) 193:3843–3863
DOI 10.1007/s11229-016-1256-0

S.I.: NEUROSCIENCE AND ITS PHILOSOPHY

The unity of neuroscience: a flat view

Arnon Levy[1]

Received: 8 February 2016 / Accepted: 20 October 2016 / Published online: 12 November 2016

Abstract This paper offers a novel view of unity in neuroscience. I set out by discussing problems with the classical account of unity-by-reduction, due to Oppenheim and Putnam. That view relies on a strong notion of levels, which has substantial problems. A more recent alternative, the mechanistic "mosaic" view due to Craver, does not have such problems. But I argue that the mosaic ideal of unity is too minimal, and we should, if possible, aspire for more. Relying on a number of recent works in theoretical neuroscience—network motifs, canonical neural computations (CNCs) and design-principles—I then present my alternative: a "flat" view of unity, i.e. one that is not based on levels. Instead, it treats unity as attained via the identification of recurrent explanatory patterns, under which a range of neuroscientific phenomena are subsumed. I develop this view by recourse to a causal conception of explanation, and distinguish it from Kitcher's view of explanatory unification and related ideas. Such a view of unity is suitably ambitious, I suggest, and has empirical plausibility. It is fit to serve as an appropriate working hypothesis for 21st century neuroscience.

Keywords Unity of science · Levels · Oppenheim and Putnam · Mechanistic mosaic · CNC · Network motifs · Sparsify · Explanation in neuroscience

1 Introduction

Many have seen the attainment of unity as a worthwhile, indeed a preeminent goal of science in general and neuroscience in particular (Bechtel and Hamilton 2007; Cat 2013). A classic and highly influential view of unification—the process of attaining

---

✉ Arnon Levy
arnon.levy@mail.huji.ac.il

[1] Philosophy Department, The Hebrew University of Jerusalem, Mt. Scopus, 91905 Jerusalem, Israel

unity—holds that it proceeds via stepwise reduction of hierarchically ordered levels (Oppenheim and Putnam 1958). Much of the debate surrounding this classic view has revolved around problems and prospects for reduction. This paper will discuss unity in neuroscience, but rather than focusing on reduction I will be concerned with the prior issue of levels. The key claim is that we can and should seek unity sans levels. On the one hand I argue, negatively, that the notion of levels required for the classic, Oppenheim–Putnam view, faces daunting challenges. A prominent recent alternative, the mechanist "mosaic" picture, avoids these challenges only at the price of a relatively minimal, less-than-fully-satisfying view of unity. On the positive side, I suggest that we can reasonably aspire for more: several recent strands of work in neuroscience point the way to a "flat" picture, on which unification is achieved via the identification of recurrent explanatory patterns. Such a picture offers a stronger notion of unity than mechanists do. And it is not grounded in levels, thereby avoiding problems facing the classic view. I put it forward in a programmatic spirit, as a kind of working hypothesis for theoretical unity in neuroscience.

In line with this overall message, the discussion is structured as follows. I start by outlining conditions that must be met by levels, if they are to serve as a basis for unity á la Oppenheim and Putnam. This serves as a backdrop for a discussion of problems facing such a notion of levels. Here, the discussion is primarily synthetic, as many of the arguments I will rely on have been made, over the years, by others. In Sect. 3 I explain why the mechanist alternative, while internally coherent and useful in some contexts, is unappealing as a foundation for scientific unity. The rest of the paper is primarily positive, and more novel, outlining my alternative proposal. Section 4 discusses three strands of work in recent theoretical neuroscience: network motifs, canonical computations, and principles of efficient neural design. These examples are suggestive of a "flat" view of unity, one that is not grounded in a hierarchy of levels. Section 5 develops this suggestion, embedding it in the causal view of explanation. I also distinguish the "flat" account from other views of unity (especially Kitcher's view of explanatory unification) and discuss its prospects within current theoretical neuroscience.

Let me highlight some assumptions before delving in. Firstly, I take it that both the shape of a potential unification and its prospects of success are in no small part an empirical matter. They depend on the nature of the phenomena in question and on the kind of theorizing that is likely to succeed. Therefore the account I give is geared at neuroscience and I do not presuppose that it applies to other areas or generalizes in other ways.

Second, I will not offer a justification for seeking unity. This is partly because I follow Oppenheim and Putnam in thinking that "a concern with the unity of science hardly needs justification" (Ibid, p. 3). But it is partly a dialectical matter: the authors I will engage with—primarily Oppenheim and Putnam, Carl Craver and others with a mechanistic viewpoint, and several neuroscientists whose work I discuss in Sect. 4—all agree that unity is desirable. That said, I alert readers who do not share such convictions that they are assumed at the outset and not argued for.

Finally, the idea of levels figures in various philosophical discussions. It is central, for instance, in debates over top-down causation. But my discussion (especially the critique) of levels is geared at unity, and I do not presume that it has relevance, and if so what that relevance might be, vis-à-vis these other issues.

2 The classical picture of levels and its shortcomings

Perhaps the most influential modern view of the unity of science occurs in Oppenheim and Putnam's 1958 essay *The unity of Science as a Working Hypothesis.* Oppenheim and Putnam viewed unification in terms of reduction: to unify different theories is to reduce them, ideally to a single fundamental theory. This unifying process is envisioned as akin to climbing up (or, rather, down) a ladder: each theory is reduced to a theory at a lower level than it, with the most basic theory residing at the lowest rung of the ladder. A lot of the discussion of this classical picture has revolved around the concept of reduction. I will focus, instead, on the prior idea of divvying up the brain into levels. I believe that it faces serious problems irrespective of reduction. The key problem is that there does not seem to be a satisfactory criterion (or set of criteria) for sorting elements of the brain into levels. Arguments to this effect have appeared in the literature, but in a fairly scattered way, with their connection to unity not always made salient. So, while I do make some original points, the bulk of the discussion in this section summarizes and organizes existing arguments and places them in the context of the unity of neuroscience. When put together, I believe we obtain a compelling case against the classical notion of levels.

2.1 Levels for unity

What must levels be like, if they are to serve as a framework for the unity of neuroscience on the classical Oppenheim and Putnam model? We can spell out four requirements; or more specifically, two pairs of requirements. First of all, since unity is by its very nature a matter of the structure of an entire domain it seems that levels ought to be such as to not leave out too much of the relevant domain (Oppenheim & Putnam raise a concern about "things that do not belong to any level" in Sect. 3.2 of their paper). By "not leave out too much" I do not intend to imply that levels are exceptionless, i.e. that every single element of the brain must belong to some level or other. But I do think levels must be comprehensive *enough* or else they cannot serve as basis for a view of unity. Next, a division into levels must not be ad hoc. Indeed, it ought to rely on criteria that are "justifiable, from the standpoint of present-day empirical science" as Oppenheim and Putnam put it (1958, p. 9). That is, grouping elements into a level should be done on the basis of criteria that capture their neuroscientifically significant properties. These two desiderata—which we may call *comprehensiveness* and *neuroscientific significance*—pertain to the horizontal dimension of the levels picture. They concern elements *at* a level.

Two further desiderata pertain to the vertical dimension of the levels picture, i.e. to the relationship *among* levels. First, since the process of unification to which the levels picture is subservient is supposed to be stepwise, it appears that levels ought to be well-demarcated (Oppenheim and Putnam presuppose this when they speak of "…the 'proper' level of that thing." (Ibid, p. 10). Ideally, levels ought to form disjoint sets, which do not share members. This way the envisioned unification-by-reduction will indeed proceed step-by-step. Second, there should be a natural ordering of levels, such that the stepwise progression can be said to move us towards the lowest level. After all, the motivation for this picture is the thought that the unity of neuroscience consists in exhibiting how all of its phenomena can be comprehended on a single general basis. (Oppenheim and Putnam formulate a criterion for ordering, discussed below, and explicitly require that there "be a unique lowest level"—Ibid.) We may call these vertical conditions *distinctness* and *ordering*. They concern relations *between* levels. Together with comprehensiveness and neuroscientific significance, they form desiderata that a notion of levels suitable for the classical picture must meet.

2.2 Problems for composition-based levels

Having anchored the previous discussion in Oppenheim and Putnam, let me begin my critical survey by looking at problems pertaining to their own suggested criterion for level-hood. My discussion here largely tracks Kim (2002/2010).

The idea underlying Oppenheim and Putnam's criterion is that things at higher levels are composed of things at lower levels. The most basic level consists of simples—say subatomic particles. Any non-simple object is a composite of simples, or a composite of composites of simples, and so on. Levels are defined in terms of the number of such compositional steps, with the stipulation that "the highest level to which a thing belongs will be considered the "proper" level of that thing." (Ibid.). This entails that levels are discrete and well-ordered, meeting two of the aforementioned desiderata.

But significant problems arise elsewhere. For one thing, on the compositional view every composite ought to be decomposable into elements at *each and every* level below it, and ultimately into simples. But as Kim notes: "surely most living organisms have as part of their functional/organizational structure "free" molecules that are not part of cells (e.g., molecules of food and other particles in bodily fluids). These free molecules may play indispensable causal roles in the biological functioning of such organisms." (2002/2010, p. 57). In terms of the desiderata set out in Sect. 2, it is not clear that Oppenheim and Putnam's view allows for a comprehensive levels picture, one that encompasses cases such as Kim's "free" molecules.[1]

The compositional view also faces serious problems with respect to the neuroscientific significance desideratum. As Kim notes, some properties and structures recur at different rungs in a compositional hierarchy. For instance, a neural network may exhibit close functional and behavioral similarities to other networks, such as gene regulatory networks. Yet on a view that grounds levels in relations of composition, they are likely to be placed at different levels, rather than grouped together. A related problem, not mentioned by Kim, is that some theoretically important composite entities, such as synapses, are not uniformly decomposable into entities at the next level below.

Now, it should be noted that Oppenheim and Putnam's view of constitution is relatively simple: it assumes a basic set of constituents and treats all subsequent levels

---

[1] Oppenheim and Putnam do address the worry that "some things do not belong at any level" (Ibid, p. 11). But the case they discuss ("a man in a phone booth") is, by their own lights, uninteresting from a scientific view-point and they thus conclude that "the problem posed by such [cases] is not serious…" (Ibid). The problems Kim alludes to are, however, interesting and serious.

as straightforward constructs thereof. Recently several authors have suggested new and more sophisticated criteria for composition. In particular, Harbecke (2015) suggests that a constituent can be understood in terms of a regularity account, analogous in several ways to Mackie's well-known INUS analysis of causation: a constituent of a system S is a member of a set of conditions minimally sufficient for S, where the disjunction of all such minimally sufficient sets is a necessary condition for S. Baumgartner and Casini (forthcoming) propose an analysis of constitution in terms of "unbreakable" correlations, i.e. where the part cannot be manipulated independently of the whole and vice versa.

Both views are given in formal-technical terms which I do not have space to reproduce and discuss, but I would like to briefly explain why these more sophisticated views, while they mark improvements in terms of our understanding of constitution relations, do not salvage Oppenheim and Putnam's picture. The key reason is that they allow for a division into levels at the *local context*—within a particular system or mechanism—but not across systems. In this respect, the situation they lead to is rather similar to the mechanistic account, to be discussed below. (Indeed, both are motivated as improvements over mechanistic accounts of constitutive relevance.) So while these accounts may very well do a good job of accounting for relations between levels of *a particular system*, they do not, and are not intended to, account for relations among levels of the more *global sort* required for a view underlying unity, as seen in Oppenheim and Putnam.

Thus, Oppenheim and Putnam's compositional criterion does not seem to meet the comprehensiveness desideratum, as seen by Kim's "free" molecules case. And it doesn't match up with neuroscientifically significant groupings, as illustrated by neural networks, on the one hand, and synapses on the other hand. More recent accounts of constitution can handle some of these issues (such as "free" molecules) but do so in a local, one-system-at-a-time, way and therefore do not yield a sorting into levels that is fit for the purposes of an account of unity.

2.3 Size-based levels?

If levels cannot be derived from relations of composition, what other options are there? In principle, there are many. I will discuss two important ones in this subsection and the next, and briefly comment on others afterwards. I begin with the idea of demarcating levels according to size, i.e. the physical dimensions of their occupants. Here I draw on a critique of levels in ecology by Potochnick and McGill (2012), where some analogous problems arise, and on related points made by Craver (2007, pp. 180–184).

On the size criterion, lower levels consist of smaller objects like atoms and small molecules, and as we go up the ladder we find larger molecules, cellular organelles, cells and so on, up through the brain as a whole.[2] Levels of size are comprehensive and can be simply and naturally ordered. It is less clear, however, that they are discrete, since size is a continuous property. But the more significant problems concern the

---

[2] For a vivid example of this way of thinking about levels—see the image in Churchland and Sejnowski (1988, p. 16).

other desideratum—neuroscientific significance. Does size mark, or correlate with, theoretically important properties in neuroscience? Consider, for instance, axons. They vary in size considerably: the giant axon of the squid is up to three orders of magnitude bigger than many other axons. So sorting axons according to size, we would have to break up this theoretically unified category—axons—into objects at different levels. The converse kind of situation is also common: some ions, for instance, act as freely diffusing elements within the cell (e.g. in action potentials). But other ions serve as components of larger, multi-molecular structures, not subject to diffusion, with entirely different roles. So while ions have similar sizes, they differ substantially in what they do in the brain—they do not seem to share neuroscientifically significant properties. Examples such as this can be multiplied. They attest to the fact that size need not correspond to theoretical significance, and so cannot serve as a good basis for demarcating levels.

2.4 Structure-based levels?

A second potential criterion for level-hood focuses on structural features. The difference between, for instance, atoms and molecules, or cells and tissues isn't only, or even primarily, a matter of size. They also have distinctive structures. Can we sort levels on a structural basis?

To fully answer this question, the structural criterion needs to be fleshed out more fully. But even absent that, we can see that it is unlikely to do the trick. Again, a central issue has to do with whether levels of structure correspond to neuroscientifically significant groupings. Consider cells: Some cells, such as astrocytes or retinal ganglion cells, are parts of organisms. Other cells, like bacteria and yeast, are organisms in their own right. It is hard to see the theoretical motivation for grouping unicellular organisms with cells in a multicellular organism.[3] Another problem for a structural demarcation of levels has to do with "combinational" elements—elements whose structure consists of parts of other structures—such as synapses. Synapses are surely a theoretically important category in neuroscience. Ought we to recognize a synaptic level? Perhaps, but on a structure-based demarcation this level would substantially overlap with other levels, such as the cellular level, contravening the distinctness desideratum. Thus, it seems that structure is not a good criterion for level-hood, at least not if levels are to serve as a basis for neuroscientific unity.

There are other possible criteria for level-hood, such as complexity and the types of physical forces operating on an entity (Wimsatt 1976 mentions these). But they seem less like likely than the options we have already looked at. Complexity is hard to define in a plausible way, and anyway does not seem to correlate with scale or theoretical significance, in neuroscience and in general (McShea 1991). Typical forces may play an important role in some physical contexts, including the brain, but it seems unlikely that one can construct the entire levels hierarchy on such a basis.

Given all this, it seems fair to conclude that we do not have available a satisfactory way of demarcating levels, classically understood. And absent such a notion of levels,

---

[3] A related argument is made by Potochnick and McGill (2012, p. 131) with respect to pheromones.

the attendant conception of unity collapses. If we are still to regard unity as a central goal, we must look for an alternative.

3 The mechanist view: local levels and unity-as-synergy

In this section I want to discuss what is perhaps the most significant contemporary alternative to the classical approach—the mechanist view. It does not lead to problems of the sort we reviewed. But that is because mechanistic levels are not intended to play the role of classical levels, at least not with relation to unity. Instead, the mechanistic approach suggests a view of unification as the synergistic study of mechanisms by different disciplines. This view is internally coherent and consonant with some episodes in the history of neuroscience. But I'll suggest that it presents a version of unity that is overly modest and for that reason not very attractive—in a sense, what mechanists offer is not quite a view of unity.

3.1 Mechanisms and mechanistic levels

Writers on mechanism differ on various issues, but all seem to accept the following basic picture. A mechanism is an organized set of causally interacting elements, which together give rise to a phenomenon. A mechanistic explanation is an account that specifies a system's components, their organization in space and time, and their mutual interactions, and shows how these give rise to some overall behavior of the system. An important aspect of mechanisms is that they can be nested: a component in one mechanism can itself be constituted by an organized set of elements that give rise to its behavior. A mechanistic explanation can be provided for the operation of a system that is itself a component in a more encompassing system.

Several authors have suggested a view of levels that is grounded in this picture (Bechtel and Craver 2006; Craver 2007, Ch. 5; Glennan 2010). The basic idea is that levels are derivative from mechanistic part-whole relations, such that the ordering of levels falls out of said relations. These are not to be confused with part-whole relations as discussed by Oppenheim and Putnam. A mechanistic component isn't any old part of a system. It is a part that plays a causal-explanatory role in generating some system-level behavior or property. Ion channels are components of the action potential. But many parts of the axon, those that do not play a role in the action potential, are not. Carl Craver offers a concise formulation of this idea: "The interlevel relationship is as follows: X's ϕ-ing is at a lower mechanistic level than S's ψ-ing if and only if X's ϕ-ing is a component in the mechanism for S's ψ-ing" (2007, p. 189). To this we may add that the intralevel relationship is derivative from causal interactions: two elements of a mechanism are at the same level just in case they interact causally.

This view of levels can be used as the basis for an account of top-down causation (Bechtel and Craver 2006) or of the status of abstract (higher-level) properties (Glennan 2010). But it is not intended to, and indeed cannot, substitute for the classical notion of levels in the context of unification. Let me clarify, and let me note that in doing so, I draw on Eronen (2013, 2014).

The mechanistic account derives levels from part-whole relations between a phenomenon and components in its underlying mechanism. But most elements in the brain do not stand in a mechanistic constitutive relation to each other, because they do not belong to a common mechanism. In fact, even elements of the *same mechanism* are often not level-ordered, according to the mechanistic conception. Consider, for instance, a population of sodium ions and a potassium channel. Both are components of the cellular circuitry underlying the action potential, but neither is a component of the other. Nor do they interact with each other in any direct way within the mechanism.[4] So, on the mechanistic account, they are neither at different levels nor at the same one.[5] Thus, mechanistic levels do not satisfy the desideratum of comprehensiveness, in that in many, perhaps most, cases it is not possible to assign elements of the brain to any (mechanistic) level.

The mechanistic view is not designed to meet the neuroscientific significance desideratum either, i.e. it does not group elements of the brain into sets whose members share theoretically interesting properties. For, as noted, the mechanistic criterion does not apply to entities and processes that belong to different mechanisms. A sodium channel and a calcium channel may have important shared properties. But since they do not belong to a common mechanism, there is no fact of the matter as to whether they are at the same level or not.

Thus in the mechanistic outlook, levels are local and partial: things belong to a level only within the context of a particular mechanism, with no overarching set of levels. Therefore levels of mechanisms cannot serve to buttress a view of unity. But this is at it should be: levels of mechanisms are not meant to buttress a view of unity. Instead, proponents of the mechanistic outlook have developed views of unity that are not grounded in levels, and which differ considerably from the classical one in both aim and content. I discuss these points next.

3.2 Unity as (mechanistic) synergy

The most developed account of unity within the mechanistic framework has been provided by Craver (2007, Ch. 7; Craver and Darden 2013, Ch. 10). He vividly labels it the "mosaic unity" account.[6] At one point he summarizes it thus: "[T]he unity of neuroscience is achieved as different fields integrate their research by adding constraints on multilevel mechanistic explanations." (2007, p. 228). Soon afterwards Craver clarifies that this view of unity differs sharply from the classical view in that it is not grounded in a process of reduction (stepwise or otherwise), and takes a cue from the idea that

---

[4] I add the qualifier 'direct' for in a sufficiently broad and indirect sense, almost anything interacts with anything else. But that would not make for a useful notion of levels.

[5] For more on action potentials and the mechanistic outlook see (Levy 2014).

[6] Several authors with mechanist leaning have developed views of the relation between different parts of science in terms of theoretical integration. See for instance Darden and Maull (1977), Bechtel (1984), and Craver and Darden (2013, Ch. 10). These views are kindred spirits to Craver's, as he himself notes. But as my discussion is geared towards understanding unity within a domain, namely neuroscience, and as I do not presuppose that what holds for neuroscience holds in other cases, or vice versa, I will not discuss inter-field integration here.

explanations in neuroscience are mechanistic: "The central idea is that neuroscience is unified not by the reduction of all phenomena to a fundamental level, but rather by using results from different fields to constrain a multilevel mechanistic explanation." (2007, p. 231).

In essence, the mosaic view equates unity with synergy: neuroscience is unified to the extent that results from different studies, often obtained via different methods and employing different underlying assumptions, are brought to bear on a similar problem. Typically, the problem is the elucidation of a mechanism for some phenomenon. On this view, while work at/on different levels of a mechanism usually proceeds with different techniques and is often couched within different conceptual frameworks, it is nevertheless mutually supportive because work on one level constrains work on other, adjacent, levels. Craver places most emphasis, it appears, on cross-level *evidential* constraints of different sorts—pertaining to time-scales, componency relations, spatial organization and other features. For example, knowledge of what goes on during an action potential at the whole-cell (mechanistic) level—where the action potential occurs, what size its components may have, the speed and rate at which they operate and so on—can tell us about the spatial organization of the underlying molecular level—what possible structures of the underlying molecules may (or may not) have, what kinds of changes they might undergo and so on. Conversely, too: knowledge about the molecules involved can constrain, principally by way of serving as evidence for, hypotheses about the kind of whole-cell process at work.[7]

Craver, along with others who offer related views, argues for this account in large part by showing that it fits well with developments in neuroscience. He discusses at some length the history of studies into LTP and memory, arguing that it is well described by the mosaic view. Specifically, it fits that case better than the classical picture of unification. Researchers did not proceed by reducing the phenomena of memory to lower-level principles of a cellular and molecular nature such as LTP. Rather, they integrated results from studies in different fields to understand the relevant mechanism—both at a (mechanistic) level and across (mechanistic) levels.

Now, I do not contest the claim that the mosaic picture captures the history of LTP research, and perhaps other important episodes, better than the classical picture. And there's no doubt that it fits nicely with the mechanistic outlook. In these respects, the notion of mosaic unity is appropriate and well-motivated. But I do think that, from the point of view of the unity of science, neuroscience in particular, the mosaic view presents a less than satisfactory ideal. Let me explain what I have in mind.

Unity, in any domain and in neuroscience in particular, is of course a graded matter. On one end we have the possibility of an entirely disunified domain, where theories, assumptions and methods are unconnected, perhaps even partly conflicting. On the other end we have a situation in which all knowledge in the domain is, in some sense, a single, homogeneous whole. And then there are various intermediate options.

---

[7] The mosaic can, in principle, encompass synergies of a non-evidential sort—for instance by mutually supportive explanations (where one field explains one aspect of a process or system, another field a different aspect), synergistic experimental techniques (e.g. recording from a cell while a subject performs a behavioral task) and so on. These kinds of aspects are emphasized in (Darden and Maull 1977) and in (Craver and Darden 2013)

There are philosophers who view the prospects for unification very dimly and envision science as forever at the "dappled" end of the spectrum (Cartwright 1983, 1999; Dupré 1993). On such a view, different problems, arising in different contexts and studied by scientists with different agendas tend to lead to different, locally appropriate but largely disconnected solutions. The classical picture, in its strong formulations, resides squarely on the other end of the unity spectrum: it envisioned reduction as resulting in a single "all-comprehensive explanatory system" (Oppenheim and Putnam 1958, p. 4).

I will assume that the low, "dappled" end of the spectrum is to be avoided if possible. And of course I accept, along with Craver and others with mechanist leanings, that the classical ideal of unity via reduction is unattainable. This is due, in part, to the problems discussed above, concerning the underlying notion of levels (and in part due to issues surrounding reduction, which I set aside here). The mosaic account of the unity of neuroscience can be seen as seeking a middle ground between these two ends: different parts of neuroscience work synergistically, attaining progress by constraining and informing each other, jointly assisting the common investigative cause. Thus they are mutually supportive (hence not strongly dis-unified) but retain their respective concepts and methods (hence not strongly unified). What the mosaic picture does not require is the existence of *shared theoretical content*: general concepts, principles and explanatory schemas applying across a range of neuroscientific phenomena. Such a situation would go beyond synergy and represent a stronger form of unity. Mixing metaphors, Craver's mosaic account presents the different parts of neuroscience as members of an alliance—independent nations joining efforts. Unity in a stronger sense, the sense to which the classical view aspired though didn't quite live up to, views neuroscience as a corporate body. Even short of the homogenous nation-state of the classical picture, we may speak of "federal" ideal of unity, akin to a set of partially distinct states, united by common principles.

Now it is of course possible that an alliance/mosaic is the best we can hope for. Perhaps the world (that is, the brain) will not allow for a greater degree of unity. Needless to say, we cannot at present rule out that possibility. But nor should we rule out the possibility of attaining a stronger kind of unity. I have surveyed reasons why we must abandon the classical view of unity, and why the mechanistic outlook does not offer a sufficiently ambitious alternative. But just because *those* pictures do not succeed, it does not follow that one cannot define and pursue a fairly strong kind of unity, a unity of shared concepts and principles. In the remainder of the article I elaborate on this idea.

4 Motivating examples: motifs, CNCs, sparsify

In developing an alternative view of unity I will start, in this section, with three examples from recent theoretical work in neuroscience—neural network motifs, CNCs and so-called design-principles, such as Sterling and Laughlin's "Sparsify". In the next section I extract the philosophical view they are suggestive of, and situate it relative to other notions of unity.

Let me highlight two points at the outset. First, the view I'll be discussing is to a significant extent speculative, in the sense that it relies on cutting edge ideas in neuroscience that may turn out to have been off the mark or of exaggerated importance. Thus, I am not presupposing that these theoretical ideas are correct, but proposing a view of unity that they *may* embody *if* correct. Secondly, the suggested view isn't incompatible, logically speaking, with the mechanist ideas presented earlier. It is meant to represent an ideal that is more ambitious than the one presented by mechanists, but compatible with the perusal of evidential and methodological synergy as well as related middle ground options. I will come back to both points in Sect. 5.

4.1 Neural network motifs

Network motifs are small sub-graphs within a broader network that occur frequently and exhibit interesting functional dynamics. Motifs were first characterized in the context of gene regulatory networks (Alon 2007). Since then, they have been described in a range of other contexts. Specifically, it has been shown that neural networks also exhibit motifs, and various ongoing efforts to detect and characterize neural motifs are under way. Figure 1 depicts two motifs that have been found to be very common across a range of neural systems, according to a comprehensive analysis performed by Sporns and Kötter (2004; See also Sporns 2010). These and other sub-graphs are unusually common in the brains of C. elegans, cats, monkeys and other animals. Notably, motifs like these occur *within cells*, especially in gene regulatory networks, as well as *among cells* and also in the pattern of connections among *whole regions* of the brain. Thus, motifs span size scales and functional distinctions.

The simplest approach in the study of motifs treats them as basic units of connectivity, and seeks to understand their dynamics, and consequently their function, simply on the basis of topology. For some purposes such a skeletal description suffices. For instance, a feed forward loop (FFL—Fig. 1b) will generate pulses of output; or, in a subtly different version, may behave as a noise buffering element. These behaviors are due principally to how the nodes are connected and depend relatively little on structural features of the nodes themselves or on the details of their interactions (Alon 2007; Sporns 2010; Levy and Bechtel 2013). This is a powerful approach, especially given that FFLs exist on a variety of scales and in different contexts. However, as noted, looking at connectivity alone is the simplest approach and for many purposes a richer

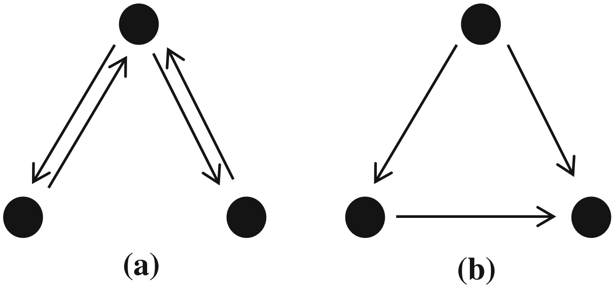

Fig. 1 Network motifs. **a** A motif identified by Sporns and Kötter (2004), labeled motif # 9; **b** a feed forward loop (Alon 2007)

model is required. There are various ways of enriching the analysis. One option is to treat the motif as situated within a structured environment, such as a field or a broader network. A recent case in point: Liu and Li (2013) modeled an FFL in which the nodes behave in accordance with the Hodgkin–Huxley equations, and, in addition, treated it as embedded within a field of astrocytes. They found that this enriched model allows one to account for the important phenomenon of stochastic resonance—in which the addition of "white noise" to a system enhances a sub-threshold signal such that it becomes detectable. Much more could be said about motifs, including other ways of enriching a motif-based model and the range of contexts in which such analyses may be fruitful. The preceding gives a rough idea of this framework, I hope.

To be sure, the study of neural network motifs is at an early stage and has been subject to criticism (Marder 2012; Sompolinski 2014). These criticisms, and the initial but as yet limited predictive success of motifs, should surely be taken seriously. Therefore, I offer a conditional claim: if, and to the extent that motif-based analysis proves important and informative, it will be a distinctively non-level-like way of thinking about the brain. For, as noted, motifs operate in a variety of different contexts—in molecular settings as well as among cells and even among large-scale brain regions. Motifs are seen to exist at different size and temporal scales and at different structural and information processing contexts. Thus, this kind of analysis is, as it were, flat: it generalizes over abstract organizational features of systems that occupy different scales, compositions and contexts. If motifs are for real, then they embody the potential for a different sort of unification of neuroscience. More on this potential below, after we look at two further examples of a broadly similar sort.

4.2 Canonical neural computation

A recent coinage by Mateo Carandini, canonical neural computations (CNCs) are "standard computational modules that apply the same fundamental operations in a variety of contexts." (Carandini and Heeger 2012, p. 51). The key idea is that some operations, characterized in formal-computational terms, are used over and over across different brain regions, modalities and scales. Like motifs, CNCs span different size and information processing tasks. Also like motifs, it appears that some complex tasks can be accomplished by a composition of different CNCs (Carandini and Heeger 2012).

One important CNC is *divisive normalization* (DN). Essentially, in DN the response of a neural element (say, a neuron) is scaled against the response of a population of related elements (neurons), either ones with a similar selectivity, (i.e. sensitivity to similar stimuli), or ones with a different selectivity. The simplest DN model takes the following form:

$$R_j = \gamma \frac{D_j^n}{\sigma^n \sum_k D_k^n}$$

R_j represents the response of element j, given a driving input D over the pool of elements k.

Here γ, σ and n are parameters determining overall responsiveness, saturation, and amplification of individual elements, respectively. The equation entails that total response rises with driving input D, but is attenuated by the population's behavior, represented as a sum in the denominator. Thus we have a kind of averaging that smooths out some of the peculiarities of the response and scales relative to its surrounding field. Such a model can be extended to add a baseline response, to distinguish different sub-pools and so on. Caradini and Heeger argue that, depending on its specific form and context of operation, normalization can allow a neural system to preferentially filter out some stimuli, enhance sensory adaptation, permit the fine-tuning of an input system's gain, as well as generate a bi-modal, winner-takes-all dynamics. Most of the work on normalization to date is theoretical, but there is an increasing number of empirical results that suggest that normalization is present in different parts of the mammalian visual system—both at the scale of photoreceptors, and in several cortical regions—as well as in other organisms, such as fruit flies (Olsen et al. 2008; Wilson 2013). Oshiro et al. (2011) have argued recently that normalization can uniquely account for certain empirically established features of multisensory integration. Multisensory integration is a context in which another aspect of work on CNCs can be seen: Van Atteveldt et al. (2014) suggest that a combination of divisive normalization and another CNC—phase resetting (PR)[8] can account for a variety of properties of sensory integration.

4.3 Sparsify

A third example comes from a very recent book—“Principles of Neural Design” by Sterling and Laughlin (2015). The book suggests ten principles that “apply across a range of spatial and temporal scales… from nanometers (protein folding, structure of ion channels) to micrometers (structure of synapses and local circuits) to a meter (structure of long tracts). [And across] a 10^{13} range of temporal scales—from microseconds (transmitter diffusion) to decades (human memory). If the principles are correct” add Sterling and Laughlin, “they should apply across brain regions and across brains of diverse species—and they do.” (Ibid, p. 440).

To illustrate, let me briefly discuss one of the principles proposed by Sterling & Laughlin, which they dub “sparsify” (the name is intended as a metaphorical instruction to a would-be designer of a brain[9]). The basic idea is that the brain is often organized in a sparse manner: a small proportion of elements is highly active at any given moment. Sterling and Laughlin argue that this kind of organization is efficient and reliable. It is essential if the brain is to perform well given its size and various physiological

---

[8] PR involves a change in the period of an oscillatory pattern of firing within a neuronal population, resulting in the amplification or dampening of further incoming stimuli, depending on where in the (reset) phase they “hit”.

[9] The references to design, here and in the book's title, are not incidental. Sterling and Laughlin believe that many of the principles stem from the design-like character of natural selection, and appeal to many ideas from electrical engineering, computer science and related areas in the course of the book. In this respect too there are similarities between this approach and the motifs and CNC cases discussed earlier, in which the notion of design plays a central motivating role. But a discussion of this interesting connection won't be possible here.

constraints such as energy use, the time scale of key chemical processes and the materials out of which a living tissue is made.

This principle is seen to apply in various contexts and at diverse scales: “information is concentrated at every scale: a few transmitter puffs open a few ligand-gated ion channels that allow a few analogue currents that sum to cause a few spikes in a few active neurons; and with this design information is sent and received efficiently. Certain cases are even shown to be optimal, but all follow the principle *sparsify*” (Ibid, p. 438; italics in the original).

Sterling and Laughlin are not alone in suggesting that there are powerful principles of efficient design that may apply to the brain as a whole, as well to its sub-systems at different scales. Another prominent example is Bialek (2012), who employs a statistical physics approach, and argues for a series of principles of efficient design in diverse neural systems, across different scales.

Like motifs and CNCs, *sparsify* and other so-called design principles are patterns that recur across spatial and temporal scales, and across different biological systems. Let me re-emphasize that these examples, and the research programs they’re associated with, are in early stages. Neuroscientists may well discover that their apparent promise was illusory. If successful, however, such theoretical developments may pave the way for an alternative kind of unity in neuroscience. Next, I want to discuss this possibility in more detail.

5 Unity without levels

The examples we have looked at—network motifs, CNCs and the Sparisify principle—are rather different beasts. A motif is a mini-network, a layout of causal connections; a CNC is a formally described algorithm; and a principle like Sparsify concerns the distribution of activity within a set of elements. But there is a commonality which is relevant to our discussion of unity: all three proposals embody the idea of a recurrent pattern, one that transcends spatial and temporal scales and applies in a range of neural systems. They depict phenomena at different sizes and contexts as, in an important sense, unitary—as exhibiting one and the same pattern and obeying a common principle. This, I suggest, points to an alternative picture of unity in neuroscience, one that offers a stronger ideal than mechanistic synergy, yet is not committed to levels in the classical sense.

In a nutshell, this is a picture in which theory is organized around a relatively small set of explanatory patterns which are found in neural systems across scales, modalities and anatomical structures. Such an alternative picture is “flat”: it does not presuppose a hierarchy in which elements of the brain are ordered. Of course, there are biological, chemical and physical structures underlying any given CNC, network motif or sparsely activated system. And the presence of the common pattern does not entail that the underlying structures are the same. That, obviously, is not denied in the alternative picture. To the contrary—how, say, divisive normalization is realized in the brain is an interesting research question that is actively pursued in the field (Carandini and Heeger 2012). So on this picture there are common, context- and scale-independent patterns, but also a diversity of underlying structure. But the key, for present purposes, is that such a view does not presuppose that the brain can be organized into a hierarchy of disjoint comprehensive sets. Size and other physical scales, compositional relations and underlying structure more generally—these obviously matter. But they do not, on the present view, generate a hierarchy of levels upon which a notion of unity rests. Instead it is the commonality of patterns and structures, which is orthogonal to scale and compositional relations, that serves as the foundation for unity. Let me further elaborate this picture and compare and contrast it with other approaches.

5.1 Unified (causal) explanations, but not explanation-as-unification

The examples above are connected with explanation. An FFL serves in an explanation of a dynamic such as a pulse of transcription activity or an effect such as stochastic resonance. A CNC like divisive normalization helps elucidate how, say, a receptive field in the retina fine-tunes a signal's gain. The fact that a network in the cortex is sparsely connected contributes to our understanding of how the relevant network can accurately represent a range of stimuli with limited resources. And so on for related examples. Such explanations are, in one way or another, causal explanations. That is, they depict the factors impinging upon the system and/or the interactions within it (if in an abstract, non-detailed, way) and outline the manner in which these interactions give rise to the system's overall behavior. [10] A feed forward motif explains pulse generation because this arrangement of elements, given the right circumstances, *produces* pulses.

Without entering into a detailed discussion of explanation and causation, let me give an indication of how I view these issues and how they connect to the claims I am making about unity. Here I assume a broadly causal view of explanation, though not any specific version of it.

A basic task of a view of explanation is to provide an account of the explanatory relation. On a causal view the explanatory relation is taken to be, as the name suggests, a relation of causal relevance. One can then specify additional requirements, namely explanatory virtues: properties that make a particular explanation better. Some such virtues pertain *to a particular explanation, of a given phenomenon*. But we may also ask about features of the *overall stock* of explanations in some domain. Suppose we have a model in some area of neuroscience that provides causal information and scores highly in terms of its "individual" explanatory virtues. We may ask, in addition: does this explanation, or something very much like it, apply to *other systems in the brain?* To the extent that it does, we have here a kind of unification of different systems under a common causal explanation. In like manner, we could inquire about the entire stock of explanations in the domain. We could ask: can we explain the various phenomena in this domain using a small number of (good) causal-explanatory models? We have a unified explanatory system to the extent that the "ratio" of explanatory models to phenomena explained is high. I should immediately clarify that the scare quotes around 'ratio' indicate that the quantitative formulation in the last sentence is, to a considerable extent, metaphorical. I do not intend to commit to the claim that there

---

[10] This is not to say they are mechanistic explanations—a matter which is controversial, especially with regards to CNCs (Chirimuuta 2014). But this will not make a difference here.

is a straightforward way of "putting a number" on the degree of unity of a body of explanations (certainly not a simple fraction). Perhaps this is possible. But I am only relying on a more basic idea, namely that we have unification to the extent that we can explain more phenomena with less explanatory resources. It is this sense of unification that, I claim, we see in the examples discussed above. We see a situation in which, say, a motif that occurs in different systems explains the generation of pulses in a similar manner in different systems, and across different scales. So we have the same causal model, capturing a similar causal pattern, operative in different contexts.

In contemporary philosophy, when one mentions the terms 'unity' or 'unification' and 'explanation' in the same paragraph, people are (rightly) reminded of the unificationist conception of scientific explanation, and especially of the of work Philip Kitcher (1981, 1989). There are some similarities between Kitcherian unification and the conception just outlined, but there are also important differences. Let me highlight both, starting with the similarities. For Kitcher, too, unity consisted in the possibility of accounting for a variety of phenomena under a small set of patterns. Kitcher speaks in terms of an 'explanatory store' and understands unification in terms of the degree to which the explanatory store "makes the best tradeoff between minimizing the number of [explanas] employed and maximizing the number of [explananda accounted for]" (Kitcher 1989, p. 82). This is a fundamental idea of Kitcher's that I embrace.

Now to the (important) differences. First, Kitcher understood explanations as derivations—formally specifiable argument schemas that have a statement of the phenomenon-to-be-explained as their conclusion. (This is why the above quote contains square brackets—I have replaced Kitcher's allusion to derivations with the more neutral terms 'explanans' and 'explanandum'.) Like most contemporary philosophers of science, I do not accept such a formalist-syntactic account of explanation. Explanations can but need not be arguments—they may also take the form of a description of a mechanism, an image, even a physical model.

A more important difference between the present view and Kitcher's unificationism is that the latter is *an account of explanation*. Roughly, Kitcher's idea is that the more unified our overall stock of explanations, the more explanatory it is. Explanatoriness can only be assessed at the global, explanatory store level, and corresponds to the degree to which that store is unified. In contrast, I am not supposing that explanation *is* unification. Indeed, I have stated that we can and typically do assess explanations one by one—in terms of their causal credentials, as outlined above. But I am also suggesting that we can ask how unified our stock of (causal) explanations is, and that we can do so by a measure of unification that is broadly similar to Kitcher's. In other words, while Kitcher *identifies* explanation with unification, my proposal presupposes a causal conception of explanation, and defines a notion of unity in terms of the degree of unification of the stock of causal explanations.

I should note a feature that is lacking in the account I just gave: while it relies on the idea of minimizing the "ratio" of explanations to phenomena explained, it does not provide a criterion for individuating explanantia and explananda. This would be required, strictly speaking, before we can evaluate the degree of unity of a stock of explanations. Moreover, such an individuation criterion must not be too stringent; two explanations that differ in minute or trivial details should be counted as the same. Otherwise we would hardly ever be able to say that the same explanation applies to different phenomena (contra the many cases in which this occurs in science). On the other hand, such a criterion cannot be too liberal, identifying loosely similar explanations and treating them (implausibly) as the same. To some extent, this problem can be handled by formulating explanations in a standard format, within which a metric of similarity can be defined. Woodward (2003), for instance, formulates many of his claims about explanation with the aid of causal graph theory and associated structural equations. Such a formulation will allow fairly precise individuation criteria. But it is not clear that all explanations can be handled this way, and I do not have a better framework to offer. That said, I think the examples discussed in Sect. 4—motifs, CNC's, etc.—show that scientists engage in assessments of sameness of explanations, and seem to agree in their judgements. So I am optimistic that a criterion can be found, at least one that is suitable for our purposes. I leave further investigation of this issue for another time.

5.2 Relationship to the mechanistic outlook

To further clarify the flat view of unity I am proposing, let me compare it to the mechanistic outlook discussed above. As we'll see, the flat view is compatible with the mechanistic outlook. But as promised, it is more ambitious with regards to unification.

Consider a network motif such as an FFL (Fig. 1b), used to explain pulse generation in a gene regularity network within a neuron. This is a kind of mechanistic explanation. The motif itself can, in turn, be explained mechanistically: nodes can be identified and decomposed into their elements (say, into structural sub-units of the relevant transcription factor) and this may explain why they interact with the other nodes (say, via non-covalent bonding to a nucleic acid) in the manner and timing at which they do. Recall that the mechanistic view construes levels in terms of mechanistic part-whole relations. Two elements of a mechanism are on the same level if and only if they interact within the mechanism. They are at different levels if one is a component in the mechanism for the other; specifically, the component belongs to a lower level than the whole. So the FFL-based explanation licenses talk of mechanistic levels: the nodes in the motif are at the same (mechanistic) level, whereas the node and its components are on different (mechanistic) levels. None of this is in conflict with saying that the same motif may appear in a different context—say, as a pattern of connections among whole cells—and that this fact constitutes a unification of our explanatory repertoire. The latter statement doesn't presuppose any claim about levels, and it is consistent with thinking that *within* each mechanistic-explanatory context, there is a way of demarcating levels.

Perhaps somewhat surprisingly, the view I have offered is compatible with the mechanistic view of unity, too. As described above, that view treats unification as synergy—wherein different investigative approaches combine concepts and methods to "zero in" on a single mechanism. There is nothing in what I have said that is in tension with such synergistic investigations. Indeed, I think they are laudable. But the view I am proposing puts forward an additional, and in my view more ambitious, goal of unity. For it suggests that neuroscientists pursue the possibility that different features of the brain can be explained in a similar way—via a similar mechanism, if you will. I regard this view as more ambitious because it suggests that there is a real sense in which to pursue unity is to aspire to more than cooperation between different parts of neuroscience. It suggests the pursuit of common underlying principles. So the views are compatible because one can pursue evidential synergy as well as common explanatory patterns. But the flat view is more ambitious because it aspires for a tighter integration—at the level of shared theoretical content.

Finally, some might wonder about the metaphysical implications of the two views of unity. Both give up on the idea that the brain can be neatly partitioned into levels, at least on an Oppenheim and Putnam style of levels. In addition, both views accept that in a given system or mechanism within the brain, relations of constitution can be discerned. In these respects the views are metaphysically akin. But doesn't the flat view, with its invocation of recurrent explanatory patterns such as CNCs and motifs, posit additional metaphysical structure? This is a complex question, which I will not be able to discuss here. Suffice it to say that resolving it requires taking a stand on how to move from explanation to ontology. My own view is that one cannot read off metaphysical commitments from scientific explanations in any straightforward way. So I would caution against taking the flat view to have any immediate metaphysical implications. But I will have to leave a fuller discussion of this issue for another day.

5.3 Generality

Before concluding, I want to discuss two questions connected with my account's generality. First, given that I motivated the flat view by way of examples, why think it generalizes beyond those examples? Second, given that the examples involve patterns that, in some cases, occur outside of neurobiology, why think the view isn't overly general? I look at these issues in turn.

How general is the account I offer? The honest answer is that I do not know. Nor do I think anyone else does. Where in the brain can we locate motifs or CNCs? How powerful are principles like Sparsify? These are subjects of ongoing investigation. The significance of the flat unity I have outlined will depend, to a considerable extent, on the outcomes of these investigations. There is little point pursuing a kind of unity that isn't likely to be had. But the neuroscientists I rely on are optimistic and have more than a few empirical and theoretical results to buttress their optimism. The "flat" conception of unity can be seen as a philosophical complement to that optimism.

The second question concerns the specter of over-generalization. Take network motifs: as noted earlier, motifs can be found in various contexts, including non-neural biological systems. Indeed, such motifs have also been detected in social networks and in the structure of the Internet. Does the view I am suggesting imply that these phenomena can be unified with neural phenomena? More generally, does the view entail that it suffices for unification that we discover a similar pattern across different contexts? And if so, does that not serve as a kind of *reductio* of the view, showing it to result in an implausibly strong kind of unity?

Let me offer three remarks in response to this concern. First, with respect to the case of network motifs. As the reader will recall from the initial discussion in Sect. 4, an analysis of motifs just in terms of their layout, the pattern of connections they embody, is the simplest possible option. It yields results, but more sophisticated analyses yield stronger, and concomitantly less general, results. So unification via the mere sharing of a pattern of connectivity can only go so far. Second, let me emphasize that the view on offer is not a Kitcher-style view, in that it places a further and very important constraint: the unity in question pertains to causal structures. So it does not suffice for flat unity that two systems can be represented by the same equation or by a similar graph. They need to share a common causal structure. If such a common causal structure is indeed identified then, I suggest, there is a real sense in which we have unified the two systems. Third and last: it is possible that some causal structures present in the brain are indeed sufficiently similar to causal structures in other systems. Sufficiently similar, that is, that their key features can be explained in a common fashion.[11] If that is the case then I think there is good reason to treat the common explanation as a unification and as a striking scientific achievement.

6 Conclusion

I have offered a novel view of unity in neuroscience. The discussion leading up to it was critical: I looked at problems with the classical, Oppenheim-and-Putnam view of unity, which presupposes a strong notion of levels. There does not seem to be a criterion for demarcating levels that meets the requirements imposed by this strong notion. Discussing the main alternative, the mechanistic "mosaic" account, I suggested that it is minimal in a way that makes it less attractive as an account of unity. Against this background I presented my "flat" account, in which unity is attained via the identification of recurrent explanatory patterns. Such an account differs from the classical picture in that it does not rest on a notion of levels. Its main motivations are, first, that it presents a suitably ambitious ideal of unity (unlike the mechanistic view); and, second, that it dovetails with a spate of recent work in theoretical neuroscience, including work on motifs, CNCs and design-principles. I acknowledge that, in this latter respect, it should be read not as a retrospective of current science but as a prospective picture for its future. To echo Oppenheim and Putnam, we may treat the possibility of "flat" unity as a working hypothesis.

References

Alon, U. (2007). Network motifs: Theory and experimental approaches. *Nature Reviews Genetics*, *8*, 450–461.

Baumgartner, M. & Casini, L. (forthcoming). An abductive theory of constitution. *Philosophy of Science*.

Bechtel, W., & Craver, C. F. (2006). Top-down causation without top-down causes. *Biology & Philosophy*, *22*, 547–563.

Bechtel, W. (1984). Reconceptualization and interfield connections: The discovery of the link between vitamins and coenzymes. *Philosophy of Science*, *51*, 265–292.

---

[11] As noted in footnote 10, part of the motivation for work on motifs, CNCs, and design principles such Sparsify is the thought that they are subject to constraints akin to those under which engineered (manmade) devices are made. From this point of view, the prospect of a unified account of neural systems and, say, the internet is not problematic. It is a prediction of such approaches and its confirmation supports it.

Bechtel, W., & Hamilton, A. (2007). Reduction, integration, and the unity of science: Natural, behavioral, and social sciences and the humanities. In T. Kuipers (Ed.), *Philosophy of Science: Focal Issues*. New York: Elsevier.

Bialek, W. (2012). *Biophysics: Searching for principles*. Princeton, NJ: Princeton University Press.

Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, *13*, 51–62.

Cartwright, N. (1983). *How the laws of physics lie*. Oxford: Oxford University Press.

Cartwright, N. (1999). *The dappled world: A study of the boundaries of science*. Cambridge: Cambridge University Press.

Cat, J. (2013). In In E. N. Zalta (ed.), *The unity of science. The Stanford Encyclopedia of Philosophy* (Winter 2014 Edition). http://plato.stanford.edu/archives/win2014/entries/scientific-unity/.

Chirimuuta, M. (2014). Minimal models and canonical neural computations: The distinctness of computational explanation in neuroscience. *Synthese*, *191*(2), 127–153.

Churchland, P. S., & Sejnowski, T. J. (1988). *The computational brain*. Cambridge, MA: MIT Press.

Craver, C. F. (2007). *Explaining the brain*. Oxford: Oxford University Press.

Craver, C., & Darden, L. (2013). *In search of mechanisms: Discoveries across the life sciences*. Chicago: The University of Chicago Press.

Darden, L., & Maull, N. (1977). Interfield theories. *Philosophy of Science*, *43*, 44–64.

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

Eronen, M. I. (2013). No levels, no problems: Downward causation in neuroscience. *Philosophy of Science*, *80*(5), 1042–1052.

Eronen, M. I. (2014). Levels of organization: A deflationary account. *Biology & Philosophy*, *30*(1), 39–58.

Glennan, S. (2010). Mechanisms causes and the layered model of the world. *Philosophy and Phenomenological Research, LXXX*, *1*(2), 362–381.

Harbecke, J. (2015). The regularity theory of mechanistic constitution and a methodology for constitutive inference. *Studies in History and Philosophy of Science Part C*, *54*, 10–19.

Kim, J. (2002/2010). *The layered world: Metaphysical considerations. Essays in the metaphysics of mind*. New York: Oxford University Press.

Kitcher, P. (1981). Explanatory unification. *Philosophy of Science*, *48*, 507–531.

Kitcher, P. (1989). Explanatory unification and the causal structure of the world. In P. Kitcher & W. C. Salmon (Eds.), *Scientific explanation*. Minneapolis: University of Minnesota Press.

Levy, A. (2014). What was Hodgkin and Huxley's achievement? *British Journal for Philosophy of Science*, *65*(3), 469–492.

Levy, A., & Bechtel, W. (2013). Abstraction and the organization of mechanisms. *Philosophy of Science*, *80*(2), 241–261.

Liu, & Li,. (2013). Stochastic resonance in feedforward-loop neuronal network motifs in astrocyte field. *Journal of Theoretical Biology*, *335*, 265–275.

Marder, E. (2012). Neuromodulation of neuronal circuits: Back to the future. *Neuron*, *76*(1), 1–11.

McShea, D. (1991). Complexity: What everyone knows. *Biology & Philosophy*, *6*, 303–324.

Olsen, S. R., Bhandawat, V., & Wilson, R. I. (2008). Divisive normalization in olfactory population codes. *Neuron*, *66*, 287–299.

Oppenheim, P., & Putnam, H. (1958). The unity of science as a working hypothesis. *Minnesota Studies in the Philosophy of Science*, *2*, 3–36.

Oshiro, T., Angelaki, D. E., & DeAngelis, G. C. (2011). A normalization model of multisensory integration. *Nature Neuroscience*, *14*, 775–782.

Potochnick, A., & McGill, B. (2012). The limitations of hierarchical organization. *Philosophy of Science*, *79*(1), 120–140.

Sompolinski, H. (2014). Computational neuroscience: Beyond the local circuit. *Current Opinion in Neuroscincę*, *25*, 1–6.

Sporns, O. (2010). *Networks of the brain*. Cambridge, MA: MIT Press.

Sporns, O., & Kötter, R. (2004). Motifs in brain networks. *P*Los Biology, *2*(11), e369.

Sterling, P., & Laughlin, S. (2015). *Principles of neural design*. Cambridge, MA: MIT Press.

Van Atteveldt, N., Murray, M. M., Thut, G., & Schroeder, C. E. (2014). Multisensory integration: Flexible use of general operations. *Neuron*, *81*(6), 1240–1253.

Wilson, R. I. (2013). Early olfactory processing in Drosophila: Mechanisms and principles. *Annual Reviews in Neuroscience*, *36*, 217–241.

Wimsatt, W. C. (1976a). Reductionism, levels of organization, and the mind-body problem. In I. Savodnik (Ed.), *Consciousness and the brain: A acientific and philosophical inquiry* (pp. 202–267). New York: Plenum Press.

Woodward, J. (2003). *Making things happen: A theory of scientific explanation*. New York: Oxford University Press.