Axiomathes (2011) 21:33–56
DOI 10.1007/s10516-010-9128-6

ORIGINAL PAPER

Physicalism, Emergence and Downward Causation

Richard J. Campbell · Mark H. Bickhard

Received: 17 September 2010 / Accepted: 18 October 2010 / Published online: 5 November 2010

Abstract The development of a defensible and fecund notion of emergence has been dogged by a number of threshold issues neatly highlighted in a recent paper by Jaegwon Kim. We argue that physicalist assumptions confuse and vitiate the whole project. In particular, his contention that emergence entails supervenience is contradicted by his own argument that the 'microstructure' of an object belongs to the whole object, *not to its constituents*. And his argument against the possibility of downward causation is question-begging and makes false assumptions about causal sufficiency. We argue, on the contrary, for a rejection of the deeply entrenched assumption, shared by physicalists and Cartesians alike, that what basically exists are things (entities, substances). Our best physics tells us that there are no basic particulars, only fields in process. We need an ontology which gives priority to organization, which is inherently relational. Reflection upon the fact that all biological creatures are far-from-equilibrium systems, whose very persistence depend upon their interactions with their environment, reveals incoherence in the notion of an 'emergence base'.

Keywords Physicalism · Supervenience · Emergence · Reduction · Processes · Downward causation

Physicalism, in one form or another, has been the dominant metaphysical position in recent decades. That dominance has been much aided by the perception that its only serious rival is some sort of Cartesian dualism. Less well recognized is how closely both these positions are linked. Whereas dualism holds that fundamentally the world is made up of just two kinds of things—mental substances and physical

R. J. Campbell (✉)
The Australian National University, Canberra, NSW, Australia
e-mail: rcampbell@netspeed.com.au

M. H. Bickhard
Lehigh University, Bethlehem, PA, USA

substances—physicalism simply denies the existence of mental substances. Both presuppose that what basically exist are things (entities, substances); their disagreement is about whether there are two kinds, or just one.

Lately there has been renewed interest in the project of developing a genuine third alternative by drawing on the idea of emergence. Unfortunately, much of the literature on this topic has proceeded within an intellectual framework which perpetuates the metaphysical assumptions common to physicalism and Cartesianism. We maintain that such a project can only succeed if those shared presuppositions are rejected—in particular, the assumption that what is ontologically primary are things, entities. We contend, on the contrary, that what humans recognize as stable, enduring things themselves emerge from certain kinds of complex processes, of which many depend for their very existence upon their interactions with their environment.

Manifestly, our contention depends on there being a viable and powerful notion of emergence. A recent article by Kim highlights many of the issues which need to be sorted out in developing a coherent account of emergence (2006). His analysis of what would be required reveals how physicalist assumptions confuse and vitiate the whole project, and must be rejected if progress is to be made.

We argue that three issues in particular confuse the discussion, and, when they are resolved, a much clearer perspective from which emergence can be understood is found. The three issues are:

1. Does emergence occur in certain special organizations of phenomena, or does emergence require more than that, as some British emergentists held (e.g., Broad)?
2. Does emergence require supervenience on some base of particulars and their properties, or does it require including the configural relations in the supervenience base?
3. Can even a fully relational construal of a supervenience or emergence base suffice for paradigmatic cases of emergence?

We show that these issues limn a background of assumptions, usually only partly explicit, that puzzle and perplex much of the discussion. We argue that emergence is best understood as inhering in special organization as such, that the most powerful sense in which emergence requires supervenience requires that emergents be understood as being *identical to* the higher level relations among the constituents in the supervenience base (contrary to a British emergentist notion), and that even this strengthened notion of supervenient emergence fails for familiar cases such as life and mind. Instead, a process-based understanding of emergence is required, and a shift to a metaphysics of process alters the entire framework of assumptions within which the discussion has proceeded.

1 Emergence and Supervenience

Kim begins his analysis by observing that "the intuitive idea of an emergent property stems from the thought that a purely physical system, composed

exclusively of bits of matter, when it reaches a certain degree of complexity in its structural organization, can begin to exhibit genuinely novel properties not possessed by its simpler constituents" (2006, p. 458). We dispute the gloss that a purely physical system is "composed exclusively of bits of matter", but otherwise that *is* the idea of an emergent property.

He then argues that supervenience is a component of emergence. Consider two wholes which have identical microstructure (i.e., they are composed of identical basic physical constituents configured in an identical structure) but differ in respect of their supposedly emergent properties. If such cases were possible, he says, the connection between emergent phenomena and their supposed substrate could be irregular, haphazard, or coincidental. There would be no point in saying that the former 'emerges from' the latter, rather than being randomly associated with various configurations of substrates. So he rejects the view that an emergent property of a whole is not determined by the properties and relations characterizing its parts, "or, to put it another way, [that] an emergent property of this sort does not *supervene* on the microstructure of the object which has it" (p. 549). He concludes that we need to accept the following proposition:

> *Supervenience*: If property *M* emerges from properties *N*1, …, *Nn*, then *M* supervenes on *N*1, …, *Nn*. That is to say, systems that are alike in respect of basal conditions, *N*1, …, *Nn* must be alike in respect of their emergent properties. (p. 550)

At first blush, this proposition might seem to capture the intuition that the emergence of higher-level properties and powers is not haphazard or coincidental. On closer examination, however, it becomes evident that it is fraught with problems, masked by its vagueness. Presumably, by "alike" Kim means that systems that have basal conditions of the *same* type must have the same types of emergent properties. More seriously, it is far from clear what is included in "basal conditions". Shortly after arguing for the above-quoted proposition, he defines supervenience as:

> *Supervenience/determination*: Property *M* supervenes on, or is determined by, properties *N*1, …, *Nn* in the sense that whenever anything has *N*1, …, *Nn*, it necessarily has *M*. (p. 550)

Note that there is no mention of relations in this definition. We will address this briefly below, and consider this omission in more detail in Sect. 2.

The argument provided is couched in terms of two wholes which have identical micro*structures*, not just identical constituents. Accordingly, it seems that "basal conditions" *is* meant to include relations, contrary to the above definition. But if so, in what sense could some relations be 'basal' relative to others? (This problem becomes severe if, as we have foreshadowed, everything is process, for that means that *everything* is relational.) Some careful attention to this issue uncovers serious problems.

If, as we suspect, the phrase "basal conditions" refers only to the micro-physical-scale relations amongst smaller-scale micro-constituents, then the definition becomes better defined at the expense of making a false assumption—the assumption that there is some privileged 'micro-scale' below which relations are

part of a relevant supervenience base, and above which relations are merely boundary conditions. But there is no such privileged scale; quantum field theory can be manifest at scales of indefinite size—superconductivity, for example. That is, there is no privileged quantum scale below which relations can be included in the 'basal conditions' and above which those relations are merely boundary conditions.

Perhaps the proposition is assuming some sort of mereological inclusion of lower-level 'basal' patterns in larger patterns. Then at least two problems arise:

1. The 'basal' patterns would have to be *instances* of patterns, not just pattern types; *types* of patterns of processes are not processes, and, therefore, are not parts of processes. But then open systems become counter-examples. In open systems, the 'basal' instances are not constant. Fires, for example, consume their component fuel and the oxygen they suck in. Living organisms turn over the cells within them many times during their lifetime. Note that this point holds also in the case in which those "instances" are considered to be particles or entities of some sort. We will have more to say about these kinds of cases later.
2. An emergence base must be bounded, but it is not clear what privileges certain pattern instance *boundaries* over others—what determines the relevant micro-constituent pattern types? What, for example, is the relevant supervenience boundary of a candle flame? The hottest colour in the center of the flame? The cooler but still visibly radiating boundary outside of that hottest domain? The collective boundary of the candle and one of those colour phase change boundaries? The cubic meter centered on the candle? And so on. Open systems, such as biological organisms, depend for their continued existence upon their interactions with their environments. Even living organisms of the same general type, such as human bodies, do not interact with their environments in exactly the same way (we all have to eat, but not all of us eat the same foods). So what constitutes the boundaries of such 'systems'? If boundary specifications are just arbitrary, then what justifies even the restriction to honouring inclusion relations among pattern instances? If the privileging of certain pattern types is *not* arbitrary, what justifies the privileging of those types? Probably Kim is implicitly assuming that the 'basal conditions' of such an organism are bounded by its skin. But if so, it is not true to say that the properties and powers of the higher-level system, the organism's body, are determined solely by what is within its skin, since it is an open system. In either case, "basal conditions" turn out to be not well defined, and, therefore, "supervenience" is not well defined.

There is a further puzzle. The notion of a micro-based property of an object is one Kim had introduced in his recent books (1998, 2005). A property characterizing an object's microstructure, he says, "tells us what sorts of microconstituents the object is made up of and the structural relations that configure these constituents into a stable object with substantival unity" (2005, 57). Such micro-based (or micro-structural) properties of an object, he rightly insists, are its macro-properties—that is, they belong to the whole object, not to its constituents. As he says, it follows that *they do not supervene* on (just) the micro-constituents and the *properties* (not including the relational configurations) of the object's micro-constituents. This

conclusion is of crucial importance for him, for it allows causal powers to be located at the macro level, and not be supervenient. Otherwise, on his reductive model, they would seep downward from level to level, from macro to micro, and, should it turn out that there is no bottom level, causation would drain out of the world altogether.

This conclusion is important for another reason, which he fails to recognize. We agree that the 'microstructure' of an object belongs to the whole object, *not to its constituents*. But thereby Kim undermines his thesis that supervenience is a component of emergence. If the causal powers of a stable system emerge at the macro level from the holistic organization of its micro-constituents, it is precisely that macro *organization* which gives the system its distinctive properties, properties that its micro-constituents (and their properties) do not have. In that case, those macro properties would be genuinely novel—properties not possessed by its simpler constituents. Therefore, according to his starting point, they are emergent. Yet Kim has said that they do not supervene on the properties of those micro-constituents. So he has himself provided a powerful argument against supervenience (given his definition) being a component of emergence.[1]

2 Supervenience and Physicalism

These considerations would seem to make a precise rendering of the intuitive idea of an emergent property rather difficult. At least, it seems that the concept of supervenience, as generally understood, will not do the trick. We hardly need to point out the many differing attempts to develop a clear and cogent definition of supervenience, nor that those attempts have been fraught with conceptual difficulties. We are not aware of any definition in the literature on supervenience, in its intended interpretation,[2] which would avoid these all problems, and we will not attempt to explore that mare's nest here.

Nevertheless, supervenience is often taken to justify the reduction of all higher-level properties to physical ones. If that works, Kim's contention that supervenience is a necessary component of emergence would imply that emergence offers no alternative to physicalism. Kim explains that ontological physicalism is "the view that bits of matter and their aggregates in space–time exhaust the contents of the world" (2005, p. 71). This characterization is indeed the common understanding of the physicalist position. And Kim does argue that supervenience delivers physicalism, as we will discuss in Sect. 3.

So let us consider the 'strong' definition of supervenience which we passed by earlier (because it omits any mention of relations). This will throw some light on the contradiction to which we drew attention in the previous section.

Since this definition omits any mention of relations, a fortiori it does not include structural relations amongst the 'basal conditions'. We surmise that no relations are

[1] Note that here Kim has avoided his own earlier "causal regularities" argument (Kim 1991) only by ad hoc stipulation in his definition that relations, configurations, are not included in the base.

[2] The caveat about 'intended interpretations' is because, given its vagueness, it might be possible to develop a non-intended interpretation of Kim's first supervenience proposition which could avoid these problems.

mentioned because Kim rightly insists that the structural relations which configure constituents into a stable object belong to the whole object, not to its constituents. And as noted, such micro-structural properties are not supposed to *supervene* on their constituent properties, taken individually or as a group. Rather, as Kim says in the earlier book,

> they supervene on specific mereological *configurations* involving these micro-properties—for a rather obvious and uninteresting reason: they *are* identical with these micro-configurations. (1998, pp. 117–8)

Since these configurations are macro-properties, they are *not on the same level* as the micro-properties thus configured. Thereby he commits himself to a micro-macro hierarchy which is not tracked by the supervenience relations; the latter remain at the same level in the micro-macro hierarchy as their supervenience base. When it is said that a certain property *P* of something *s* supervenes on microphysical facts about *s*, what that means, Kim says, is that the fact that *s* has *P*, or whether or not *s* has *P*, is fixed once the micro-constituents of *s* and the properties and relations characterizing these constituents are fixed—and those relations are *not* included in the base (p. 85).

So, when we talk of supervenient properties and the base in which they are realized, Kim concludes, there is no movement downward, or upward, in the hierarchy of entities and their properties ordered by the micro-macro relation. Therefore, the realization relation does not 'drain' causation down to entities of the lowest level, or perhaps away altogether. Rather, causation remains at the same level—the level at which the supervenient property is realized.[3]

This seems like a clever move, saving physicalism while allowing for a certain sort of emergence (although not the sort advocated by the so-called British emergentists of the 1920s). We believe that Kim has reached an important insight in recognizing that *configurations of constituents are what generate the emergence of higher-level causal powers*. Indeed, Kim himself now can speak of ways "in which a macro-property may be emergent" (p. 117). He has not, however, fully thought through its implications. His concern is solely to prevent all higher-level causation from being drained of causal potency.

But now, if supervenience does not track the micro-macro hierarchy, and it is in that hierarchy that higher-level causal powers emerge, what has happened to the thesis that supervenience is a component of emergence? Once again he has himself provided an argument against that thesis.

We surmise that what has prevented Kim from recognizing the full significance of his endorsement of emergence is that he has focussed his discussion exclusively on the British emergentists, while remaining committed to physicalism. That brings us back to the above characterization of physicalism—the one that omits relations from the basal conditions.

What is amazing is that this characterization of physicalism is inconsistent with Kim's argument about the status of structural relations. The problem is that the

[3] Note the *essential* role in this reasoning of the assumption that the supervenience base does not include the configurational relations among the constituents.

holistic relations which structure an object's constituents are not generated by simply aggregating the properties of those constituents. An aggregate is not a structure. *Kim's dilemma is that he needs the ontology of structure as a potentially causally powerful ontology in order to avoid causal powers draining downward. Yet he has defined physicalism in such a way that structure plays no significant role*. It takes structure to block causal drain, but he still thinks that is consistent with mereological physicalism. Except for a few hints in the 1998 book, he does not seem to recognize that he has thereby carved out a 'new' ontological realm, that of structure or organization.[4] We will return to explore this new realm in Sect. 6.

3 Emergence and Reduction

The invocation of supervenience in order to deliver physicalism raises the issue of reduction. Clearly, if higher-level properties can be reduced to lower-level properties they are not emergent. What is not so clear is what exactly "reduced" means in this context.

One possible interpretation is that irreducibility means that emergent properties cannot be *predicted* or *explained* purely in terms of their basal conditions. That is the understanding of emergence which Kim extracts from the British emergentists. Alternatively, irreducibility can be interpreted more weakly, as requiring only that emergent properties cannot be shown to be *identical* with the properties of the constituents in their emergence base—not that they cannot be *explained* in terms of them.

The first interpretation is itself in need of clarification. Some physicalists take Nagel-style 'bridge laws', correlating higher and lower-level phenomena, as sufficing to reduce the former to the latter. Kim, however, denies that. Consider the following inference:

Jones's C-fibres were stimulated at *t*.

Anyone whose C-fibres are stimulated will experience pain.

That is why Jones experienced pain at *t*.

Kim argues that inferences like this are compatible with emergence. The second premise invokes empirical facts about the emergent property in question, linking phenomena at the two levels. To claim that 'bridge laws' like this 'reduce' higher-level phenomena to lower-level ones is to beg the question of emergence. An effective *reductive* explanation, he claims, should invoke only base-level conditions.

But there is another model of explanation which, Kim argues, *would* be incompatible with emergence. That is reduction through functionalization. Functional reductions, if they work, rely on conceptual *definitions* of upper-level

[4] This may be because he is so focused on British emergentism as the only non-physicalist alternative to dualism. That is, the three options he recognizes are either dualism, or British emergentism, or mereological physicalism. But an ontology constituted by levels of organization is different again.

properties in terms of their roles as causal intermediaries between sensory inputs and behavioural outputs. Again, his example is pain. Suppose that being in pain is defined as being in a state which is typically caused by tissue damage and trauma and which typically causes aversive behaviour. Then, provided this functional definition is sound, empirical investigation might discover that in humans and mammals certain neural conditions are typically caused by tissue damage and trigger aversive behaviour. In this way, pain is reduced to neural conditions. If we understand mental properties and powers to be functional properties, a physicalist can maintain that the only potential occupants, or 'realizers', of these causal roles are physical properties. In this view, functionalization is 'merely' the picking out of some causal property (or properties) as being the functional one(s); in that sense, it is a second order property and adds nothing beyond the causality of the first order property selected.

Kim himself does not believe that an emergentist can accept this kind of reductive explanation. If pain can be given a functional definition or interpretation (something Kim in fact denies), and if it can be identified with neural states by empirical investigation, then functional reduction also accomplishes ontological reduction. So he proposes irreducibility as a second condition of emergence.

Let us now turn to the alternative, weaker interpretation of irreducibility. In his 'Reply to Kim', Marras disputes Kim's claim that functional reductions are substantially different from bridge-law explanations (2006). Marras loosens up the understanding of reductive explanation to requiring only an isomorphism between properties and laws at the lower level and those at a higher level. Then he argues that on either model, properly understood, it is possible to draw a distinction between a property's being *reductively identified* with its base property and a property's being *reductively explainable* in terms of it.

Now, the starting point of this whole debate is that properties are emergent if they are truly novel properties, i.e. ontologically distinct from the 'base' properties from which they emerge. This, Marras says, only requires that emergent properties are *not identical* to their base properties, not that they cannot be reductively *explained* in terms of them.

Now, in arguing that functionalizable properties would effect a reduction, Kim is assuming that properties can always be functionalized one property at a time—he defines functionalizing that way. Marras' discussion shows that this is at best an artificial and implausible restriction, except perhaps for a few cases like 'light switch'. But, if the functionalization of mental phenomena (or any other) requires the sort of *network* process and consequence isomorphisms for which Marras argues, then there is no longer any difference between functionalization and micro-macro configuration relations—neither would be supervenient so long as relations are not part of the supervenience base. Even if 'switches' might be functionalizable one at a time, computers, for example—even if defined in terms of switches—cannot; they are inherently relational. 'Network' functionalization would not be 'at the same level' as the micro-conditions in which it was realized. We will not pursue any further here Marras' dispute with Kim about whether functional reductions are substantially different from bridge-law explanations, and will confine our discussion to Marras' distinction.

The criterion for demarcating emergent from non-emergent properties, Marras proposes, is this: emergent properties are those supervenient properties which, though reductively *explainable* in terms of base properties, are not reducible to—i.e., reductively *identifiable* with—them. He accepts Kim's claim that pain is not a functional concept—that there is no functional analysis of pain—but insists that that is irrelevant to the question of its reducibility, via bridge laws, to neural conditions. That is, he accepts that there may in principle be reductive *explanations* of pain, but claims that if the relevant bridge-law is one-way, not a biconditional, there is no basis for claiming property identity. (*Perhaps C*-fibre firing is not the *same* as pain, but it might still be a way in which pain can be *realized*.)

We believe that Marras is correct in insisting that emergence requires only that higher-level properties and powers be novel, and that this does not exclude the mirroring kind of explanation of them which he sketches. But in other respects his position is vulnerable. Marras characterizes it as a form of non-reductive physicalism while agreeing with Kim that supervenience is a necessary component of emergence.

Kim has an oft-repeated argument which shows that non-reductive physicalism, which distinguishes between higher-level properties and their physical realizations, collapses into reductive physicalism. Suppose M is a mental property, with causal powers, and that some instance of it is causally efficacious in bringing about an instance of another mental property, M^*. But, *ex hypothesi*, M^* is physically realized in its physical base P^*, without which M^* would not be present. So, M must have brought about P^*. The only coherent story, Kim maintains, is that the instance of M caused M^* to be instantiated *by causing its realization base, P^*, to be instantiated*. But M has its own physical realization base, P. And if M supervenes on P, the presence of P is sufficient for the presence of M. It follows by causal transitivity that if M is causally sufficient for the presence of (an instance of) P^*, and thereby M^*, P is causally sufficient for both P^* and M^*. Accordingly, the hypothesized causal efficacy of M is superfluous; its physical realization base is what does all the causal work (1998, p. 37ff).[5]

What reason, Kim then asks, is there for not bypassing mental causes altogether and treating them as mere epiphenomena? The negative answer to this rhetorical question is firmed up if one accepts, as a physicalist is bound to, the causal closure of the physical world; mental events or properties can only bring about physical changes if they are realized in some physical base-properties. The purported mental causes are superfluous. So, either physicalists must give up their anti-reductionism, or else they must reject the possibility of psychophysical causal relations.

Kim then points out that there are two ways in which psycho-physical causation can be denied: one is to deny that there are mental events; the other is to keep faith with mental events but concede that they never enter into causal transactions with physical processes. So either a physicalist has to espouse so-called 'eliminative materialism', or else to move further in the direction of dualism, a dualism that

[5] Note how Kim's use of the letters "M" (for "Mental") and "P" (for "Physical"), which we have followed here for ease of exposition, illustrates how these debates still proceed within a Cartesian dichotomy.

posits a realm of the mental in total causal isolation from the physical domain. Kim's argument therefore poses a physicalist with an invidious choice: either there is no emergence of causally efficacious properties above the base level, or the physical domain is not closed. For any serious physicalist, the latter is not an option. Hence, once supervenience is admitted, if one wants to be a physicalist, one cannot be a non-reductive physicalist; one has to embrace some kind of 'eliminative' or 'reductionist' program.

As Kim argued in earlier writings, to explain mentality in these terms implies that supervenience lapses into epiphenomenalism. There can be no serious causal work for any mental event or property to do—that work is already effected by its physical 'realizer'. So, any 'supervenient' mental property or power, on these assumptions, is superfluous, causally impotent. And the basic objection to epiphenomenalism was well expressed by Samuel Alexander (1927, p. 8):

> It supposes something to exist in nature which has nothing to do, no purpose to serve, a species of noblesse which depends on the work of its inferiors, but is kept for show and might as well, and undoubtedly would in time, be abolished.

Of course, this whole argument against non-reductive physicalism only works if emergent properties and powers are supervenient upon their physical base. So any physicalist who believes that emergent phenomena exercise their own distinctive causal powers faces an uncomfortable choice: either those powers do not supervene on the properties and powers of their lower-level constituents—which places physicalism in jeopardy—or they can be eliminatively reduced to the properties and powers of those constituents. We agree with Kim in denying that emergent powers supervene on the properties and powers of their lower-level constituents,[6] and therefore reject Marras' conclusion that the set of emergent properties may well include mental properties as conceived by non-reductive physicalists.

4 The Possibility of Downward Causation

As Kim notes, it is critically important that if there are 'emergent' properties and powers, they should have their own distinctive causal powers, irreducible to the causal powers of their base properties. Otherwise they would be mere epiphenomena. But, as his argument against non-reductive physicalism shows, this implies that emergent properties not only bring about changes in other emergent phenomena on the same level, but such changes are also accompanied by changes in the relevant base-level properties. That is, anyone who maintains that there are emergent properties and powers is committed to the possibility of 'downward causation'.[7]

Certainly, there seems to be no shortage of examples of downward causation. Certain psychological states (e.g., prolonged anxiety, embarrassment) can cause physiological effects (heightened blood pressure, eczema, blushing) in a human body. McClelland's experimental studies of human motivation showed that

[6] Note once again the absence of relations.

[7] The term "downward causation" comes from the American psychologist D.T. Campbell (1974).

affiliative motives (the capacity to love and be loved) promote better health.[8] Another example: the functional molecules (DNA, proteins, fatty acids, etc.) within a cell are fabricated within internal processes of the cell itself; they are generated through the web of interactions of the whole system (Moreno and Umerez 2000). That downward causation occurs is a fact; how to understand the phenomena is the contentious issue.

Kim, however, adapts his argument against non-reductive physicalism in order to show that emergence entails downward causation, but downward causation is incompatible with the thesis that the emergent properties of a whole are determined by the properties and relations characterizing its parts. Any causal work effected by an emergent property is also effected by its physical base, but then emergence is no better than non-reductive physicalism. So, he argues, emergent phenomena cannot fulfil their causal promise; anything they causally contribute can be, and is, contributed by a physical cause, and if it be insisted that nevertheless emergent phenomena also are genuine causes, then their effects are over-determined.

Suppose once again that M is an instance of an emergent non-physical property which brings about the instantiation of another non-physical property M^*. This would be an instance of 'same-level' causation. Now, since M^*, is also an emergent property, instances of it which come into being will be accompanied by changes in the physical base of this system. Let us call the new physical condition thus brought about an instance of P^*. So M not only brings about M^*; it also brings about P^*. Kim's crucial step then follows:

> Now we face a critical question: if an emergent, M, emerges from basal condition P, why cannot P displace M as a cause of any putative effect of M? Why cannot P do all the work in explaining why any alleged effect of M occurred? If causation is understood as nomological (law-based) sufficiency, P, as M's emergence base, is nomologically sufficient for it, and M, as P^*'s cause, is nomologically sufficient for P^*. It follows that P is nomologically sufficient for P^* and hence qualifies as its cause. The same conclusion follows if causation is understood in terms of counterfactuals—roughly, as a condition without which the effect would not have occurred.

He concludes that this result, unless it is successfully rebutted, threatens to bankrupt one of the central claims of emergentism. "If downward causation goes, so goes emergentism" (2006, p. 558).

Five significant objections, however, can be brought to rebut this argument:

1. It begs the question. To assert that P is nomologically *sufficient* for M is precisely the distinctive claim of physicalism, and so cannot be invoked as a premise in an argument against the possibility of causally significant emergence. And on Kim's own account, this 'sufficiency' could hold only if P *included* the configuration of its constituents, which violates his definition of

[8] D. McClelland (1987, pp. 366–368). His research found that the salivary immunoglobulin A levels of subjects were significantly increased when they viewed a film of Mother Teresa designed to arouse affiliative motives.

supervenience. No-one who believes in emergence need accept this premise. Indeed, they had better not!

2. As Kim himself acknowledges, citing the British emergentist Morgan (1923), those who defend emergence expressly *deny* that the relation between an emergence base and the properties which emerges from it is a causal one. Furthermore, Kim's claim that *P* is nomologically *sufficient* for *M* is also contradicted by his own insistence, discussed in Sect. 1, that micro-structural properties do not *supervene* on their constituent properties taken individually or as a group. As we saw earlier, he takes micro-structural properties to supervene on *configurations* of these constituent properties.
3. Even if micro-structural properties are identical with the configuration of micro-constituents, it does not follow that the causal powers of such configurations can be attributed to the causal powers of their constituents. A configuration is a property of a *set* of constituents, and sets are not identical to their members, nor to any aggregation of them. Hence, it does not follow that *P* is causally sufficient for *M*.
4. As we will elaborate later, the constituents of self-maintenant phenomena are not present within them for the same period of time as the phenomena themselves persist. Fires (including candle-flames) and biological organisms persist for significant periods, while consuming and/or expelling their constituents. The properties of those constituents cannot be nomologically sufficient for the properties and causal powers of those phenomena, which must therefore be attributed to their organizational form.
5. While Kim argues that emergence involves accepting some sort of supervenience, his invocation of nomological sufficiency is not justified by that (deficient) sense of supervenience. His correct point—that the connection between emergent phenomena and their supposed substrates cannot be irregular, haphazard, or coincidental—does not justify his characterizing that connection as nomological sufficiency, especially not in both directions; multiple realizability refutes the *M* to *P* direction. So, if relations are not included in the base, then, on Kim's own account, *P* (which does not include relations) is not sufficient for *M*, and if the possibility of multiple realizability is recognized, then *M* is not sufficient for *P*. All that an advocate of emergence need accept is that *some* physical base is necessary, but not sufficient, for the emergent property.

We conclude that, contrary to Kim, the possibility of downward causation is not defeated by this argument.[9]

[9] Systems that are necessarily in open interaction with their environments—that are (e.g., a candle flame) constituted in such interactive flows—pose even deeper problems for supervenience. We will consider these later.

5 The Failure of Physicalism

To deny some sort of emergence seems foolhardy—even for a philosopher! Emergence appears to be ubiquitous throughout the world. Just about everything that exists now which is at least a potential topic of scientific interest has emerged since the Big Bang. So, any purported scientific model of any phenomenon must be able, at least in principle, to account for the ontological and historical emergence of that phenomenon since the Big Bang. Any model that does not provide such an account has to be reckoned as thereby incomplete. In the light of this, we can turn around physicalism's exclusion of the very possibility of novel emergence and use such exclusions as a powerful negative criterion for assessing scientific theories. That is, *any purported model of X that makes the emergence of X impossible is thereby self-refuting*.

The failure of Kim's argument against downward causation does not suffice, of course, to establish its possibility, and consequently the possibility of emergence. Before taking up that challenge, however, we want to prepare the ground by exploring some of the basic assumptions adopted, commonly but uncritically, by physicalists.

The physicalist claim is that all facts obtain *in virtue of* the distribution of the fundamental entities and properties—whatever they turn out to be—of completed *fundamental physics*. These fundamental entities are basic particulars of some sort: bits of matter, concrete events, or whatever, which are the fundamental constituents out of which everything in the world is composed, and whose properties and relations are sufficient to determine everything that is true about the world.

The most obvious candidates for these basic particulars are the 'elementary particles' of which physicists still continue to speak, loosely, when describing quantum mechanical phenomena. For nearly a century physicists have struggled to reconcile the facts that these phenomena exhibit, in different experimental circumstances, both particle-like and wave-like behaviour. With so much being discovered that physicists themselves confess to finding weird, the habitual ways of talking about particles persist even though its use is confused and confusing. Even reputable physicists often give explanatory descriptions in terms of 'particles' in a way which, if taken seriously, would be incompatible with the physics they are trying to explain.

From the beginning of the twentieth century, physics has been dogged by a series of deep theoretical inconsistencies that are not yet fully resolved. But enough theoretical progress has been made for some conclusions of metaphysical significance to be drawn. What our best contemporary physics reveals is that *there are no elementary 'particles'*, elemental events, or some such particulars; *everything* is composed of quantum fields, of various scales and complexity (Weinberg 1977, 1995, 1996; Davies 1984; Saunders and Brown 1991; Brown and Harré 1988; Cao 1999). Quantum field theory shifts the basic ontology of the universe from micro-particles to quantum-fields-in-process. What have seemed to be 'particles' are now conceptualized as particle-like processes and interactions resulting from the quantization of field processes and interactions. Those are no more particles than are the integer number of oscillatory waves in a guitar string.

Each of the apparent particulars assumed by a physicalist ontology is a quantized field process.

It is just as well that the physicalists' metaphysical model, which would reduce everything to interactions amongst basic particulars, is not supported by recent physics, since it is arguably incoherent. Either the basic 'particles' are tiny regions with no internal differentiation or they are points. That is, either they have extension, or they do not. If they have no extension, then a field view is forced in order to account for interactions amongst them, since the probability of such 'particles' ever actually hitting each other is zero. If, on the other hand, these elementary 'particles' do have finite extension, they pose intractable problems for physical theory. They could not be compressible; the velocity of transmission of force through their diameter would therefore be instantaneous; there would be extreme difficulty in explaining differing kinds of interactions (gravity, electricity, etc.); and so on.

If a move is made to a combination of 'particles' and fields (the typical contemporary semi-sophisticated view), then all the significant issues are already granted anyway in the granting of fields. There are no 'particles', but, even if there were, so long as fields are granted at all, the micro-reduction motivation fails—and a strict particle view is not only factually false, but conceptually incoherent as well (Bickhard and Campbell 2000, p. 332). This stage of our argument turns on the point that an ontology of basic particulars is simply not workable: without fields, nothing would ever happen; and there are no such 'particles'.[10]

Once we have made the conceptual shift required to free ourselves of the age-old prejudice of a substance metaphysics—of which a metaphysics of basic particulars is the most recent manifestation—it is far from clear that there *is* any basic level. That is, there might well be no fundamental plane of organization, 'lower' than which it is not possible to go. Even if there were a level of organization that is the lowest, that would be a contingent fact. Without any assurance of such a basic level upon which to ground his position, the physicalist's claims become rather hollow.

Physicalists might reply to all this that they have always been open-minded about which particular theory, or set of theories, physical science will eventually accept as the best confirmed. Their position is simply that, whatever that account should turn out to be, every phenomenon will, in principle, be determined by the properties and powers of the basic particulars posited by that account. So, if some refinement of quantum field theory should prove to be the best that physics can provide, why not accept quantum fields as the ultimate reality, and regard everything else as supervenient upon and reducible to their behaviour? By adjusting one's physics, one can retain one's physicalist metaphysics.

[10] It might be objected to this argument that Newtonian mechanics was not incoherent. But that would be to overlook the fact that Newtonian mechanics is *not* a purely particle ontology; it admitted forces in addition to the particles, e.g., gravity. With the restriction to a finite speed (the speed of light) and conservation of energy, those forces have to be fields. Take away the forces (forget fields for the moment), it would be logically/mathematically consistent, but nothing would ever happen because points never hit each other. Whether or not a Newtonian system would support the kind of anti-emergence arguments that a pure particle ontology requires would depend on how the metaphysics of those forces is understood. Significantly, it was various non-linear resultants of force relations that generated the first notions of emergence.

The adjustment in one's physics, however, is not so simple, nor so innocent; the 'particle' view is intimately related to the micro-reductionist position. If 'particles' are elementary, they do not *have* a configuration. They do, however, *participate* in configurations relative to each other. But if 'particles' are the purported locus of causal power, their organization can only be a boundary condition, with no causal power of its own. *This is the crucial point*. On this view, organization is factored out as a legitimate locus for causal power—it is just stage setting[11]—and such delegitimation seems to succeed because there is a non-configurational candidate available to be such a causal locus: particles. The 'particle' view supports the reductionist view by motivating the elimination of configurations as legitimate loci of causal power. So, abandoning a particle physics in favour of a quantum field physics, a process metaphysics, is not an innocent choice with respect to the issues at hand.

It might be thought easy enough to fix up the account of physicalism to avoid this, e.g., by stipulating that physics contains not only the basic particulars and their properties, but also their *configurations*. That is how Kim tries to render his position consistent in his 1998 book and subsequently. Accordingly, he defends a wide sense of the word "physical", to include the biological and neurophysiological, indeed every emergent property short of conscious experience. Accordingly, he takes all of the following as 'physical':

> first, any entity aggregated out of physical entities; second, any property that is formed as micro-based properties in terms of entities and properties in the physical domain; third, any property defined as a second-order property over physical properties; and perhaps conjunctive properties, if we want them. (1998, pp. 114–5)

This is implausible. For one thing, controversial metaphysical conclusions cannot be achieved simply by stipulation. And if it is recognized in this formulation that causal power can emerge in new organization (relations) the physicalist has given everything away.[12] It is not surprising that someone who believes that every phenomenon is determined by the properties and powers of bits of matter and their aggregates would take it as obvious that higher-level properties and powers generated by configurations of these bits should also be classified as physical. But it is remarkable that Kim, who from his 1998 book onwards has placed such store on the thesis that configurations generate new causal powers, should not have realized that this thesis undermines physicalism as he has defined it.

Since they are holistic, configurations are properties of a *set* of constituents, as we said above—not of those constituents themselves, nor of any aggregate of them. Ever since Bertrand Russell wrote his famous letter to Gottlob Frege in 1902, the logical perils of regarding a set and its members as being on the same level have been clear. It follows that configurations and their constituents cannot coherently be described in the same vocabulary. A hierarchy of levels, with novel properties emerging at higher and yet higher levels of organization, does not raise this sort of difficulty. That much Kim has seen. But it also follows that the properties of sets

[11] That explains why characterizations of physicalism typically omit relations from the physical base.

[12] Except to dualists or British emergentists.

which are generated by their configuration cannot be treated as on a par with the properties of their constituents. Not only is the set of apples not an apple, but the properties of apples can neither be affirmed or denied of the set of apples. Likewise, a set of physical entities is not a physical entity, and physical properties can neither be affirmed nor denied of a set of physical entities.

The only way in which this formulation could still be a physicalism would be if somehow emergence were blocked or simply denied, even though this revision is supposed to include in 'physics' all properties, relations, etc. needed to account for everything else. What would not be included in 'physics' on this account (unless some other way of delegitimating some levels of relations from the possibility of having emergent causal power is invoked)?

Kim's motivation for defining the physical so widely is to retain the causal closure of the physical domain, that is, the belief that tracing the causal ancestors and consequents of any physical event will never take us out of the physical domain. Assuming causal closure of a world consisting of aggregations of physical entities, higher level organization is causally superfluous relative to the working out of the causal powers of the most basic constituent 'particles'—unless organization, configurations, or relations can themselves be legitimate loci of genuine causal power. If they can, then new organizations can yield emergent causal power, but, in that case, not all causal power is resident in particles, or in whatever micro-particulars are preferred. But that is to give up the belief that everything in the world is determined by the properties and powers of bits of matter and their aggregates. It is to (stipulatively) retain the causal closure of the *physical* world, at the cost of the causal closure of the *micro*-physical world, as Hansen has point out (Hansen 2000).

Another manoeuvre which has been suggested is to allow that the fundamental ontology consists of 'quantum fields' but to insist that particle talk supervenes on it, so there could be a two-step account of the supervenience of macro objects and properties. This objection seems to be suggesting that (a) particles supervene on quantum fields, and then (b) higher-level properties and powers supervene on particles in the standard particle-physicalist way—as if the first step has no consequences for the rest. But that is surely wrong.

Firstly, the emergence of properties and powers at level n from the organization of processes at level $n - 1$ requires contextual constraints that are not themselves adequately describable in terms of the processes at level $n - 1$. Secondly, in the case of quantum fields it is their organizational properties that 'do all the work' in step (a), but thereafter organizations are relegated to the role of being simply relational properties of the entities at level $n - 1$. That is a serious conceptual disjunction in this proposed model. The power of organization cannot be admitted for levels below particles, and then get rid of it above them, for at least two reasons: (1) it is logically inconsistent to deal with relations/organization in this manner, and (2) there is no level like that envisaged which emerges out of quantum fields—even particle talk, insofar as it is consistent at all, must still also address fields in order to address any sorts of interactions among 'particles'.[13] So this manoeuvre has simply

[13] Strictly, quantum field *interactions* are quantized and usually localized, and those two properties are all that remains of particles.

raised a prima facie conceptual possibility that is wrong both in terms of the physics and in terms of the logical and conceptual coherence of the 'position' outlined.

What is absolutely critical here is that quantum field processes have no existence that is independent of their configurations: quantum fields *are* processes, and can *only* exist in various patterns. Those patterns come in many sizes, of many different physical and temporal scales, some as large as a human person, or a social institution—but they are all equally patterns of processes. There is no 'bottoming out' level in quantum field theory—it is *patterns* of process all the way down, *and* all the way up.

That is the rub. To be a physicalist is to believe that 'higher-level' entities are nothing other than complex configurations of lower-level entities, in such a way that the higher-level properties and powers are in principle identical to the properties and powers of the lower-level entities (or at least, wholly determined by them). But now the supposed base-level entities are nothing but configurations of process as well! If there is no 'bottoming out', there are no bases to which all other phenomena can, even in principle, be reduced. Our physicalist has lost the ground on which he wants to stand. *If being configurational makes a property or power epiphenomenal, then everything is an epiphenomenon*. That is the reductio ad absurdum of this position.

6 A Process Model of Emergence

In the light of these considerations, we can now outline a different model of emergence. Taking emergence seriously means taking seriously the standard multi-layered hierarchy of the different sciences. Each has its distinctive theoretical vocabulary with which it describes its field, with its distinctive properties and powers. But here we must be careful. In his 1998 book Kim characterized this hierarchy as ordered simply by mereological relations, which is highly prejudicial. There are two assumptions built into this seemingly innocuous account, both of which deserve to be challenged.

The first is that by generating the hierarchical structure by the part-whole relation, the levels are characterized primarily in terms of entities, which can be exhaustively decomposed into other entities that are their proper parts. The picture is of *things*, composed of smaller *things*. At the bottom (if there is one), there are the tiniest things: elementary 'particles'. The account is loaded in favour of a particle-like metaphysics.

Secondly, and not surprisingly, this standard picture assumes that the *relations* into which things are organized play no constitutive role (although, as discussed, from 1998 onwards Kim has been departing somewhat from the standard picture on this point). It supposes that molecules, for example, can be decomposed, without remainder, into atoms. Of course, no-one wants to deny that some force holds the atoms in a molecule together. Nevertheless, the implicit claim is that the proper parts of, say, a molecule of water, are two atoms of hydrogen and one atom of oxygen, *and nothing else*.[14] So, this apparently harmless characterization of the

[14] A more sophisticated and informed version of this claim would acknowledge the forces within the molecule, but give them a particle interpretation. Thus, the proper parts of a molecule of water would be

multi-layered hierarchy of the sciences already implies a clear denial that the spatiotemporal *organization* of those atoms has any role to play—other than being boundary conditions—in the emergence of the characteristic properties of water. Accordingly, higher-level emergent properties are rendered causally superfluous in this ontology; they have been designed out.

Processes exist *only* in some organization or other. Some organizations of process are fleeting, such as Newton's legendary falling apple. Others are stable—or at least relatively stable—and may persist for eons. Stability over time and against perturbation manifests the *cohesion* of an overall organization of process. The articulation of the required model, which we can only sketch here,[15] takes seriously the observation that there are fundamentally two forms of process stability: (1) energy well stability; and (2) far-from-equilibrium stability.

The former is exhibited when some process remains in or near thermodynamic equilibrium for a significant period. Such 'energy wells' require a significant level of energy to change their organization, and they can only be disrupted by a higher level of energy than they typically encounter in their ambient environment. Hence, they are typically very stable and robust. Atoms are straightforward examples; they are a furious process of electron waves around an even more furious dance of quarks and gluons, a process that takes a great deal of energy to destabilize.

Combinations of such stable 'energy well' processes exist at the macroscopic level, and some of the properties which such combinations manifest, such as mass, are the resultant of aggregating the properties of more microscopic processes that are their constituents. (The organization of certain stable processes to form energy wells can themselves also manifest emergent properties: e.g., van der Waals forces among molecules, or differing molecular properties that manifest quantum interactions within the molecules.) Note that aggregation is itself a form of spatiotemporal organization, and that more is involved in being a causally effective aggregate than simply the logical sum (a bare conjunction) of its constituents. The components have to *stick together*.

More remarkable is the second kind of stable organization. *Far-from-equilibrium stability* occurs in an organized process when, despite not being in thermodynamic equilibrium, it persists in that state for some significant time. What enables any far-from-equilibrium system to survive is the interaction of its intrinsic processes with its ambient environment which counter the physical tendency to move toward equilibrium.

Consider, for instance, a candle flame. It can manifest persistence and stability for hours on end, but it can do so only as long as fuel and oxygen continue to be drawn into the process of burning. Cut off either, and the flame is put out; it ceases to exist. In short, the persistence of far-from-equilibrium processes is possible if and only if such crucial interchanges with the environment (are able to) persist: they are *necessarily open* processes.

Footnote 14 continued
two atoms of hydrogen and one atom of oxygen, plus the elementary particles whose exchanging holds the molecules together.

[15] For a more detailed elaboration of this model, see Campbell (2009).

A candle flame illustrates another feature of this model also. For a candle flame is a familiar example of a process that makes several active contributions to its own persistence. It maintains its temperature above the combustion threshold; it vaporizes wax into a continuing supply of fuel; and in standard atmospheric and gravitational conditions, it induces convection currents, thus pulling in the oxygen it needs and removing the carbon dioxide produced by its own combustion. Processes like this tend to maintain themselves; they exhibit *self-maintenance*. The ability to be self-maintaining is an *emergent causal power* of the organization of the candle flame; it cannot be explained simply as the physical resultant of the causal properties of its distinct constituents.[16] Of course, in one sense its persistence is dependent upon its constituents: when the candle flame has burnt all its fuel, or it is deprived of oxygen, it ceases to be. But so long as the boundary conditions are fulfilled—its external requirements for fuel and oxygen continue to be satisfied—it continues to contribute to its own persistence. It succeeds in maintaining its own process of burning.

A further level of complexity is exhibited by systems that can maintain stability not only in certain ranges of conditions, but also within certain ranges of *changes* of conditions. That is, they can switch to deploying *different* processes depending on conditions they detect in the environment. A relatively simple example is a bacterium that can swim up a sugar gradient, but tumble if it finds itself to be swimming down a sugar gradient. These two kinds of activity—swimming and tumbling—are different ways for the bacterium to act appropriately to its environmental conditions—appropriate in the sense that each contributes to its self-maintenance in the differing circumstances. The bacterium's ability to detect sugar-gradients, and to respond by switching between its two modes of action, means that it exhibits a kind of maintenance of its own ability to be self-maintenant; it is able to switch between its self-maintenant processes as the environment changes. That is, it exhibits *recursive self-maintenance* (Bickhard 1993, 2009). Note that even in these relatively primitive examples, the description of these systems of organized process has to use *self-reflexive* locutions and to speak of its *abilities* and *actions*.

These considerations entail a radically different ontology. Biological systems—including humans—are *not* aggregations of cells (smaller things), which in turn (after a few more reductions) are aggregations of elementary particles. Again, an aggregate is not a structure, and certainly not a process. And such creatures are open, organized action systems, in *essential* interactions with their environments, such that we cannot say what they *are* without taking those interactive processes into account.

This model of emergence enables us to see clearly what has gone wrong with the concept of an 'emergence base'. Consider first the case of stable 'energy wells'. These are processes which are relatively self-contained, and do contain smaller but likewise stable 'energy well' processes within them. It makes sense to speak here of 'constituents'. If the concept of an 'emergence base' of constituents were to have

[16] Of course, combinations of far-from-equilibrium systems can also manifest aggregative properties, e.g., mass, but they are not what is remarkable about such systems.

valid application, it would be to these cases. However, to take one of the simplest examples, there is more to an atom of hydrogen than an electron and a proton; there is the *pattern* of the relationship between them, and that pattern of the process, its organization, is what is crucial to the emergent properties of hydrogen.[17] Consequently, explaining the emergence even of stable energy wells requires attributing causal power to the organization of their constituents; the constituents alone, with their properties and powers, are not sufficient. So, while we can make sense of the proposition that a higher-level 'energy well' process, like a hydrogen atom, has lower level constituents, the concept of the atom's 'emergence base' is far from clear once we take the organization itself, a holistic relational fact, into account—the atom *is* that organization of processes between the nucleus and the electron.

When we turn to the other kind of stable system, far-from-equilibrium systems, the unsoundness of the inference is even more obvious. For the latter kind of systems can maintain their stability *only* by dint of their interactions with their surroundings. Their very existence—and their persistence—are dependent upon their relations with *external* factors in that environment, from which they keep drawing sustenance.

Consider once more a candle flame. What is its 'lower-level emergence base'? One answer (somewhat simplified) we might try is: the molecules of vaporized candle wax and of oxygen. But those molecules which are present at any one time within the visible area of the flame are soon consumed, and the products of that combustion—mainly heat, light, carbon dioxide, and water—are largely either radiated or carried away by convection. All that *persists* within the flame is the wick, but it is not, in any relevant sense, a basal constituent *of the flame*, and it too is progressively consumed.

It is true that some of the energy released by the burning at any one time is used to cause the combustion of succeeding molecules of molten wax and oxygen. But if this argument were sound, the *molecules* of wax and oxygen within the area of the flame at any one time would be nomologically sufficient for (i.e., be causes of) the combustion of other molecules within the flame, at a later time. That is nonsense. It is the *flame*, the *process* of burning, which causes the molecules even to be present within its own area for a brief period, before they are in turn burnt. The flame has *no* stable set of constituents that could sensibly be called *its* emergence base (not even the wick). There is just an inflow of waxen and oxygen molecules, which are consumed, and an outflow of carbon dioxide and water molecules (ignoring other trace by-products, which do not affect the argument), together with a release of energy.

At bottom, the problem with the argument we are considering is its being framed in terms of entities and their properties, at both higher and lower levels. Its terminology of emergence bases and higher-level properties necessarily envisages only internal structures, not open processes. This is necessarily so, because an emergence base could not include any relations external to the system.

[17] Consider a proton at one location and an electron a light year away. The mereological sum of these does *not* constitute a hydrogen atom.

To see this, consider the property of being the longest pencil in a box. The pencil's having this property has nothing to do with the molecules and physically internal relations that make up that pencil (Bickhard and Campbell 2000, p. 333).[18] It is a relational property of that pencil, which it would lose if a pencil longer than it were added to the box. Now, adding another pencil to the box would not alter in any way the molecules and internal relations of the former pencil. The property of being the longest pencil seems trivial, but the logic of the case would apply equally to any extrinsic, relational property. So, while any distinction between relational and non-relational properties seems context-dependent—and not purely formal—on any account extrinsic, relational properties such as these could not be part of any 'emergence base' of the type presumed by the argument.

On the other hand, we cannot say what a candle flame *is* without mentioning its *relations* with *external* elements in its ambient situation. The very being of the flame, then, is a function, in part, of these external relations. These *physically external* relations are *logically internal* to any flame; they are constitutive of its being. If, say, the temperature of the atmosphere around the candle were gradually raised (independently) towards that of the flame itself, the convection currents required to suck in new oxygen and remove carbon dioxide would progressively become less effective. Either the flame would go out, smothered by the carbon dioxide it had been producing, or the entire candle would first melt and then vaporise.

"—Oh", might come the reply, "that is just a matter of the boundary conditions; no-one would seriously suggest that the temperature of the ambient atmosphere is a constituent!" But a significantly lower temperature of the ambient atmosphere *is* a necessary part of the nomologically sufficient conditions for the existence of the candle flame; fail to take those conditions into account, and the argument we are assessing collapses. A candle flame is *necessarily* open, lest it be snuffed out. Physically external relations are essential to the flame's ability to maintain itself, an ability that is a genuinely *emergent* causal power.

The same applies in the case of any stable far-from-equilibrium process systems, such as biological systems—from plants to humans. And this is why it has been so easy for philosophers to become confused. For at any one time, a biological system does seem to have a stable set of constituents: the cells and complexes of cells that constitute its body. Nevertheless, the relevant principle is the same as in the case of the flame. In general, a biological system turns over the cells that constitute its body a number of times during its lifetime. (Humans generally replace almost all the molecules in their body parts many times during their lifetime.)

Philosophers who have taken note of this have tended to assimilate it to the case, much discussed as a puzzle about identity, of the boat whose planks are replaced one by one while remaining afloat. To think about the turnover of the cells in our bodies in the terms of that famous example, however, is to miss the most significant aspect: the significance, the *necessity*, of the external interactions of the body. In this respect, biological systems are like candle flames; as complex organizations of processes, they persist *only so long as* they are able to maintain appropriate

[18] The example comes from Paul Teller (1992).

interactions with their environment, by which to sustain their existence. They eat and drink (and in many cases, breathe)—and perspire and excrete. If they stop those activities, they die.

More deeply, it is not just that ontologically open systems are *dependent* on interactions with their environments, they are *constituted* by, realized in, those interactive processes. They are constituted as twists or knots in the topologies of process flow, not in the mereological bases of particles or other particulars. In traditional metaphysical language, since those interactions are necessary to their very existence, they belong to the 'essence' of such systems.

So, we have to dismiss the argument that if there were emergent entities and properties, they could have no causal powers of themselves which (in part) reorganize their own sub-processes. This argument takes lower-level emergence bases to be causal factors in a way that belongs with the particle metaphysics rejected in Sect. 5. It ignores the crucial role of organization in the emergence of properties and powers from systems of lower-level processes. It also ignores the essential role of physically external relations in the self-maintenance of stable far-from-equilibrium process systems, from flames to human beings. Yet stable far-from-equilibrium process systems are the ones that have given rise to the most interesting and intriguing of emergent causal powers: life, consciousness, and self-consciousness.

To recognize this is to recognize the need for a different ontology, one that takes fields-in-process as ontologically primitive. Processes are not particulars; their logic is different. While processes of various scales and complexity occur in particular spatiotemporal regions, unlike particulars, they can occur in different regions, and can re-occur. That is, while raining must occur somewhere in order to exist at all, raining can occur both in Canberra and Bethlehem, both yesterday and on Tuesday last week, whereas yesterday's raining in Canberra cannot occur in Bethlehem and cannot have occurred last week. Generic processes are dynamic features, a 'going on thusly', something that is not a particular in the traditional sense at all (Seibt 2001, 2009). Significantly, process predications give rise to nominalizations that require mass-quantifiers, not the quantifiers over the denumerable domains of standard predicate logic.[19]

Again, taking fields-in-process as ontologically primitive means that the physicalist claim that higher-order properties and powers are all 'micro-based' collapses, since there are no 'basic particulars' upon whose properties and powers they could supervene. Consequently, the attempt to broaden the concept of the 'physical' to cover everything (except perhaps qualia) also fails. So what should count as 'physical'? We cannot make any sense of "physical" other than "what is described using only the language of physics". Quantum field processes are so described, and larger-scale entities—which we would characterize as relatively stable, relatively cohesive *systems* of quantum processes—that do have certain physical properties, such as mass. But intentionality is not part of that language.

[19] This correspondence between process predication and mass nouns is explored in an illuminating way by Barry Taylor (1977), by Alexander Mourelatos (1978), and by Peter Roeper (1987).

Mass is an aggregative property, but intentionality, precisely because it is not aggregative, has to be classified as an emergent property.[20]

To make our position clear, we are not arguing that replacing a 'particle' physics with a physics of quantum fields, of itself, resolves all the philosophical issues of intentionality. We *are* arguing that this development in physics draws attention to the need to develop an ontology of fields-in-process, an ontology that recognizes the irreducible role of organization in generating novel properties and causal powers. Recognizing the fundamental causal role of organization in itself blocks mereologically based arguments, and opens the way for the possibility of modelling causally efficacious emergent mental processes. At the level of human beings, at least, such an ontology would have to reckon with emergent characteristics beyond those dealt with in physics.

The fundamental point, which must be taken utterly seriously, is that there is nothing to a field without its organization—which does not preclude the possibility that the kind of organization which manifests intentionality can have as constituents quantum field processes—and molecules, and cells, and … etc. (at different 'levels'). However, it is not their micro-constituents which generate the distinctive properties and causal powers of intentional systems—and of the properties and causal powers of chemical and biological systems, too—but how in each case they are organized. And macro-organizations *alter* the organization of the micro-constituents which enter into that macro-organization.

Generic processes are what scientific theories describe. What is physically significant about them is how they are configured and organized; their configurations and organizations are what ground their causal powers, not the properties borne by particular things which might be supposed to constitute them. Working through the details of this alternative ontology would take us beyond the scope of this paper. But enough has been said to show that there is a conceivable metaphysics which avoids both dualism and physicalism, and which offers a way of understanding the genuine emergence of properties and causal powers. Within the process-based metaphysics sketched here, the possibility of emergence is no longer problematic. The hard work before us is in creating theoretically rigorous and empirically well-confirmed models of actual emergents.

References

Alexander S (1927) Space, time and deity, vol 2. Macmillan, London

Bickhard MH (1993) Representational content in humans and machines. J Exp Theor Artif Intell 5:285–333

Bickhard M (2009) The interactivist model. Synthese 166:547–591

[20] The definitive discussion of 'aggregative properties' is in William C. Wimsatt (1986) He proposes that different kinds of emergent properties correspond to the failure of different kinds of aggregativity. The only extra feature which a whole could have, over and above those which could result from an aggregation of its parts, is *how the parts are organized*. So in 1997 he turned this point into a positive definition of the concept of emergence: *an emergent property is—roughly—a system property which is dependent upon the mode of organization of the system's parts*.

Bickhard M, Campbell DT (2000) Emergence. In: Andersen PB, Emmeche C, Finnemann NO, Christiansen PV (eds) Downward causation. Aarhus University Press, Aarhus, pp 322–348
Brown HR, Harré R (1988) Philosophical foundations of quantum field theory. Oxford University Press, Oxford
Campbell DT (1974) Evolutionary epistemology. In: Schilpp PA (ed) The philosophy of Karl Popper. Open Court, LaSalle, pp 413–463
Campbell R (2009) A processed-based model for an interactive ontology. Synthese 166:453–477
Cao TY (1999) Introduction: conceptual issues in quantum field theory. In: Cao TY (ed) Conceptual foundations of quantum field theory. Cambridge University Press, Cambridge, pp 1–27
Davies P (1984) Particles do not exist. In: Christensen S (ed) Quantum theory of gravity. Adam Hilger, Bristol, pp 66–77
Hansen CM (2000) Between a rock and a hard place: mental causation and the mind-body problem. Inquiry 43(4):451–491
Kim J (1991) Epiphenomenal and supervenient causation. In: Rosenthal DM (ed) The nature of mind, pp 257–265
Kim J (1998) Mind in a physical world. MIT Press, Cambridge
Kim J (2005) Physicalism or something near enough. Princeton University Press, Princeton
Kim J (2006) Emergence: core ideas and issues. Synthese 151:547–559
Marras A (2006) Emergence and reduction: reply to Kim. Synthese 151:561–569
McClelland D (1987) Human motivation. Cambridge University Press, Cambridge
Moreno A, Umerez J (2000) Downward causation at the core of living organisms. In: Andersen PB, Emmeche C, Finnemann NO, Christiansen PV (eds) Downward causation. Aarhus University Press, Aarhus, pp 99–117
Morgan CL (1923) Emergent evolution. Williams and Norgate, London
Mourelatos APD (1978) Events, processes and states. Linguist Philos 2:415–434
Roeper P (1987) Principles of abstraction for events and processes. J Philos Log 16:273–307
Saunders S, Brown HR (eds) (1991) The philosophy of vacuum. Oxford University Press, Oxford
Seibt J (2001) Formal process ontology. FOIS-2001 Proceedings, ACM Digital Publications, 2001
Seibt J (2009) Forms of emergent interaction in general process theory. Synthese 166:479–512
Taylor B (1977) Tense and continuity. Linguist Philos 1:199–220
Teller P (1992) A contemporary look at emergence. In: Beckermann A, Flohr H, Kim J (eds) Emergence or reduction? Essays on the prospects of nonreductive physicalism Walter de Gruyter, Berlin, pp 139–153
Weinberg S (1977) The search for unity—notes for a history of quantum field theory. Daedalus 106:17–35
Weinberg S (1995) The quantum theory of fields, vol I, foundations. Cambridge University Press, Cambridge
Weinberg S (1996) The quantum theory of fields, vol 2, modern applications. Cambridge University Press, Cambridge
Wimsatt WC (1986) Forms of aggregativity. In: Donagan A, Perovich A, Wedin M (eds) Human nature and human knowledge. Reidel, Dordrecht, pp 259–291
Wimsatt WC (1997) Aggregativity: reductive heuristics for finding emergence. Philos Sci 64:S372–S384. Reprinted in Bedau Mark A, Humphreys P (eds) Emergence—contemporary readings in philosophy and science. MIT Press, Cambridge, MA, pp 99–110