Acta Biotheor (2007) 55:243–267
DOI 10.1007/s10441-007-9014-9

REGULAR ARTICLE

The Limitations of Kim's Reductive Physicalism in Accounting for Living Systems and an Alternative Nonreductionist Ontology

Slobodan Perovic

Received: 13 July 2006 / Accepted: 5 June 2007 / Published online: 11 July 2007

Abstract Jaegwon Kim's exclusion argument is a general ontological argument, applicable to any properties deemed supervenient on a microproperty basis, including biological properties. It implies that the causal power of any higher-level property must be reducible to the subset of the causal powers of its lower-level properties. Moreover, as Kim's recent version of the argument indicates, a higher-level property can be causally efficient only to the extent of the efficiency of its micro-basis. In response, I argue that the ontology that aims to capture experimentally based explanations of metabolic control systems and morphogenetic systems must involve causally relevant contextual properties. Such an ontology challenges the exclusiveness of micro-based causal efficiency that grounds Kim's reductionism, since configurations themselves are inherently causally efficient constituents. I anticipate and respond to the reductionist's objection that the nonreductionist ontology's account of causes and inter-level causal relations is incoherent. I also argue that such an ontology is not open to Kim's overdetermination objection.

Keywords Ontology · Reductionism · Causation · Overdetermination · Metabolic control · Morphogenesis · Nonseparability · Emergence

1 Jaegwon Kim's Wholesale Argument for Reductive Physicalism and the "Life-body" Problem

According to Jaegwon Kim (1993, 1998, 1999), the Principle of Causal/Explanatory Exclusion (CEP) is the foundational principle of physicalism, of both its ontology and the causal explanations to which it is committed. The Principle rests on the

S. Perovic (✉)
Department of Philosophy, Carleton University, 3A36 Paterson Hall, 1125 Colonel By Drive, Ottawa, ON, Canada K1S 5B6
e-mail: sperovic@hotmail.com

claim that there is a sufficient and necessary cause/causal explanation of any (caused) physical event. This excludes other events as causes, given that two distinct events *m* and *p* cannot both be independent causes of an event *e*, as they cannot both be necessary for its occurrence. If two independent events both cause *e* independently, then neither is necessary for its occurrence. Accepting this possibility would presumably deny the coherence of causal explanations.

The CEP is substantiated by an ontological account, which specifies what counts as a distinct event and, thus, as a distinct cause. More specifically, the Physical Realization Thesis, which, according to Kim, should be accepted by all physicalists, states that every property is a physical property, or is realized by a physical property or a set of such properties. This implies that every (caused) physical event (i.e., an event characterized by physical properties) is caused by a necessary and sufficient *physical* cause.

A nonreductionist physicalist argues that a higher-level property *M*, realized by a physical-basis property *P* of an event *e* can cause a property *M'*, realized by the lower-level basis of an event *e'*, *independently of the causal power of P*. It is important to note that as a physicalist, the nonreductionist accepts the Physical Realization Thesis. Now, the conjunction of her argument and the Physical Realization Thesis unavoidably leads to the following destructive dilemma. Either, given that the nonreductionist does not reject that every higher-level property is instantiated by a physical property, she has to explain the causal power of *M* in terms of its basis *P*, in which case she must discard her aspiration to argue for the nonreductive feature of physicalism. Or, she must acknowledge that her explanation implies overdetermination of the effect by multiple instances of ontologically different kinds of properties, which violates the CEP (and accordingly the Physical Realization Thesis).

Kim states the overdetermination objection in somewhat looser terms: since the causal power of *M* is instantiated in *P* due to the Physical Realization Thesis, what causal work is left for *M* to do? One must acknowledge, Kim argues, that the two events *m* and *p*, presumably characterized by the properties of ontologically (causally) distinct kinds, for instance, a mental and a physical event, cannot both cause an instance of a property of an event *e* without violating the Physical Realization Thesis (as substance dualism does) or the CEP.[1]

Put otherwise, a physicalism grounded on the CEP and the Physical Realization Thesis cannot be reconciled with the view of higher-level properties as independently (or even highly autonomously) causally powerful. As every physicalist necessarily embraces the Physical Realization Thesis, our nonreductionist described above must acknowledge the primacy of the physical over the mental: she is committed to *the Supervenience Principle* as a minimal requirement of physicalism (Kim 1998, 15).[2] Thus, an advocate of a coherent physicalism must acknowledge

[1] Thus, an overall physical cause of an effect can be rather characterized by physical events *m* and *p* as its distinct features, not as ontologically independent, yet causally relevant. For example, "My pouring down of a glass of water over the burning dish caused the smoke" involves the events of "pouring down water" and "dish burning" as two events that figure as features of the overall cause.

[2] Kim's notion of supervenience is defined in terms of determination of causal powers (and causal efficiency) of supervenient properties by subvenient basis. It is not simply identified with the claim that there can be no change in supervenient properties without the change in the subvenient basis. Rather, the latter is substantiated by the former.

that every higher-level property is causally powerful only to the extent of the power of its physical basis (The Causal Inheritance Principle). She must explain (away) the causally independent nature of higher-level properties.[3]

As Kim's exclusion argument is a general ontological argument, its wider relevance goes beyond the mind-body problem. As Kim puts it:

> So if the argument works, it should work against the causal potency of all supervenient properties, and since there is good reason to believe that biological, geological, and other special-science properties are supervenient properties in the same sense in which mental properties are supervenient, the argument should work equally well against these properties. (Kim 1998, 86–87)

In short, Kim's destructive dilemma is relevant with respect to any higher-order properties, as well as the entities that instantiate them, which are claimed to supervene on a physical basis. This includes those properties which characterize living systems. Thus, the mind-body problem is only one instance of a general ontological issue, and it has a special place in contemporary philosophy because of the intricate nature of intentionality.

Although mental properties are special, we should not emphasize this at the expense of treating the relation between other higher-level properties and the so-called physical basis as instances of the same ontological issue. For example, biological properties are also special, albeit in a different sense, and this forces us to think about the implications of the exclusion argument with respect to what we might call the life-body problem.

Because of its general nature, then, Kim's argument is best approached on a level-by-level basis. Such an approach will be especially advantageous if an ontological argument that successfully challenges reductionism arises from the conceptual study of relevant empirical investigations. Ideally, such a study would inform a sound criticism of, and enrich the domain of, acceptable alternatives to reductionist principles of physicalism. Purely speculative analysis alone may not be capable of fulfilling the task.

My argument concerns the level of living systems. Specifying the context of a system in terms of the strong relatedness of its components, pertinent to their functionality and that of the entire system, has become an indispensable and irreducible aspect of biological explanations. Arguably, this is not limited to certain explanations of rare and not-so-well explored biological phenomena; rather, it is an essential component of a wide domain of biological explanations.

I will discuss examples from the theory of metabolic control and morphogenesis.[4] In the so-called systems of metabolic control, the contextual component of enzymes' activities and interactions, as well as genes responsible for their activation, contribute to the system's behavior, along with the properties of

[3] As we will find out later on, it turns out that Kim gives up the formulation of the reductionist credo in terms of reducibility of causal powers. The revised version of the credo concerns the causal efficiency of macroproperties supervenient on the causal efficiency of microproperties.

[4] I will also sketch the possible relevance of the structural context in the theory of natural selection.

individual genes and enzymes. In morphogenesis, the complex behavior of cells that results in the formation of the organs' patterns is predicated on the structuring influence of a pattern "template" on the communication among the tissue cells.

The context as a structuring component in such biological explanations challenges reductionism heuristically and epistemologically. Perhaps we simply need to invoke structurally contextual properties in our explanations of living systems in order to cope with the complexity of living systems. However, I would argue that we are justified in drawing a stronger conclusion.

My argument is divided into three inter-related parts. In Sect. 2, I will provide a conceptual argument for ontology of *contextually nonseparable systems* (CNS). I will also discuss experimentally grounded explanations of biological systems of metabolic control and morphogenesis, which fundamentally concur with the suggested ontology. In Sect. 3, I will explain the ways in which this ontology differs from and challenges the grounding principles of the ontology of reductive physicalism as Kim advocates it. Being predicated on the exceedingly narrow understanding of the causal nature of macroproperties and inter-level causal relations, the ontology of reductive physicalism (as advocated by Kim) fails to capture the nature of the above-mentioned biological systems. Finally, in Sect. 4, I will respond to the reductionists' objections, namely their criticism of the incoherence of the ontology of the CNS and the overdetermination that it supposedly presupposes.

2 Towards a Nonreductionist Ontology of Living Systems

Unlike contextually separable systems, the CNS instantiate what I will call *structuring collective properties (C)*. Before I specify such properties in more detail, it might be helpful to distinguish them from the collective properties characterizing the systems which are separable in an obvious way. For example, the redness of a sheet of paper is not an instance of *C*, although it characterizes an entity composed of multiple parts, since each piece remains red after the sheet is torn into multiple fragments (unless, of course, we decompose it into macromolecules, whose properties result in the sheet's redness). Thus, the sheet's redness is merely an epiphenomenal collective (macro) property—a result of the sheet being composed of multiple red parts (the redness of which results from relevant macromolecular properties).

Yet this example is fairly simple, and a more substantial distinction between separable and nonseparable systems is required. Although a CNS that instantiates *C* is decomposable into smaller components, it is distinguished from separable systems by meeting the following criteria:

1) *C* is *a collective property*: It cannot be instantiated in the components that result from decomposition of the system. (This is not true of the redness of a red sheet of paper, but it may be true of other properties, such as an image of a jigsaw puzzle.)

2) *C* is *a contextual property*: It is not instantiated in each individual constituent but rather in the context of the system's (individual) constituents. (This might be true of a jigsaw puzzle as well. The image is instantiated in its interrelated pieces, not in any one piece individually.)
3) *C* is *a structuring property*[5] : Neither the causal power nor the causal efficiency of *C* is reducible to the causal powers of the properties characterizing the system's individual constituents. The structuring context of a CNS acts as a causal (macro) *constituent* of the system, and is highly autonomous from the individual (micro) constituents, their properties, causal powers, and causal efficiency.[6]

This third condition is certainly the most important, as it elevates *C*'s status beyond merely collective property, that is, beyond the status of a causally irrelevant epiphenomenon. Further, it is ontologically contentious from the point of view of reductive physicalism, the topic of Sect. 3. The fact that *C* is instantiated in a system must change the way its individual constituents will interact and the way they affect other systems. More specifically, *C* must act as one of the causal determinants, over and above the individual constituents, and in some cases, as the most relevant determinant, by constraining the causal autonomy of individual constituents.

At least some living systems are contextually nonseparable. For example, the way the relevant genes affect an organism's traits depends on the context of the phenotype (e.g., protein molecules and enzymes). The phenotype, in turn, is multi-layered. For instance, in metabolic processes, there is the level of proteins coded by the genes, the level of the activity and interactions of enzymes, and finally, the most general level of phenotype involving the visible traits of the organism. But the multi-layered phenotype is context-dependent as well (e.g., dependent on the molecular structure of the environment), and as such, it is not a simple product of the genotype.

The indispensability of invoking the structuring relevance of the context becomes apparent even in very simple metabolic processes.

A metabolic pathway is a chain of biochemical reactions, involving a number of substrates whose activity results in the production of yet another substrate. If we observe a very short metabolic pathway, for example, a system of chemical reactions catalyzed by interacting enzymes,[7] the flux of the system (i.e., the level of the substrates' concentrations) that results from the initial concentrations and subsequent interactions of the enzymes may be considered its product or phenotype (Nijhout 2001, 131). Initially, as the molecular base sequence of the genes codes for their activity (i.e., coding for the enzymes' rates of concentration), the variations that result from mutations in the genes will change the corresponding enzyme's activity. The specific enzyme's contribution to the change of the flux can be

[5] I will be using the expressions "structuring context" and "structuring collective property" depending on what aspect of the essentially identical concept I want to emphasize.

[6] I am drawing the distinction between causal powers and causal efficiency, as this distinction will prove to be crucial to the reductionist argument and to my response to it.

[7] The enzymes are the catalysts of protein interactions, the synthesis of proteins being coded by genes.

measured as the sensitivity coefficient: it shows to what extent the flux will change with a given amount of change in that particular enzyme.

But the genetic variation is not the only determining factor: *the number of enzymes on the pathway also affects the sensitivity coefficient of a single enzyme*. Thus, a change in number of enzymes in the system will change the flux by affecting the genetically coded rates of concentration. This would not be the case if the change in the flux were merely a result of the addition of genetic variability in each enzyme. An activity of a single enzyme in such a case should not depend on the number of the enzymes involved and should change the flux in a linear fashion.

This does not happen, and one might well ask why. Simply stated, the so-called *summation property* of the system structures the relation between the sensitivity coefficients of all enzymes in the system in such a way that their sum is always equal to one. If the end product of the system is a flux in which one enzyme is contributing significantly more than the others (i.e., its concentration is higher than the concentration of other enzymes), then the concentrations of the enzymes (whose concentration would be high if predicted on the basis of genetic variation alone) are "toned down," until the sum of the coefficients is equal to one.

Yet one needs to invoke the structuring influence of the context in relevant explanations at an even deeper level. In fact, the structuring context (with *C* instantiated in it) is responsible for metabolic control at the level of the genetic background, where the variation in one gene affects the change in the flux.

Such explanations are not limited to the models of the systems, the status of which one could conceivably treat as convenient shorthand for the *in principle* available explanations that do not invoke the structuring context (i.e., the summation property). Rather, these explanations are also advanced on the basis of *experiments* which tackle, in a thorough manner, the nature of structuring collective property, at multiple levels of metabolic systems (Kascer and Burns 1973, 1981; Kascer and Porteous 1987; Keightley and Kascer 1987).

Thus, it is possible to trace the variation in only one gene, which exhibits the variation and codes for one enzyme. This can be done through an experiment with a system, which involves different numbers of additional enzymes whose genes do not exhibit any variation. It turns out that this single genetic variation will not code in a linear fashion for the enzyme's activity, nor for its effect on the flux. If there are two genetically co-dominant alleles (i.e., everything else being equal, they should code for two distinct physiological effects in a simple additive fashion), then, one will be suppressed by the system (i.e., by the system's property of structuring/stabilizing interactions among genes), allowing the other to become dominant.

Explaining the dominance of a certain enzyme by tracing interactions between it and other enzymes is neither sufficiently informative nor entirely precise. In fact, given relevant genetic variation and the molecular structure of the enzymes, the interactions could result in a number of different outcomes. The actual outcome is determined by a particular property, namely the summation property that *structures all the interactions*. Whether an enzyme will dominate the flux depends not only on how it interacts with other enzymes, but also on how these other enzymes interrelate. The causal impact of all these causal relations will determine the flux. Having one less enzyme will not change the outcome in a linear fashion, for which

one could account by simply subtracting its influence on the dominant enzyme, as it will also affect the outcome indirectly by changing the interactions among other enzymes.

Although nothing changes in terms of the enzyme's (molecular) structure (as in the above-described experiments), the enzymes interact differently simply because there is one less enzyme in the system. Thus, although the interactions among the constituents are an indispensable component of the system's activity, the nature of each interaction and its outcome, and thus, the behavior of individual constituents, will depend on the structuring property realized in the entire system, not *vice versa*.[8]

There is another striking example of biological systems instantiating *C* that yields a causally significant context: namely, morphogenetic systems, the systems responsible for organismal pattern formation.

Morphogenetic processes are often identical, even in very different species characterized by substantially different genetic structure. The phenotypic endpoint is reachable by different developmental pathways, and natural selection results in morphological novelty by means of new epigenetic relationships. This indicates that *epigenetic factors* play a critical role. Accordingly, the switching of the genes "for" cellular processes, which lead to pattern-formation may be only one of the underlying processes. It turns out that the mechanisms governing the organism's pattern-formation, including those characterized in terms of Mendelian analysis, cannot be tackled without acknowledging the "contextual information during the process of development" (Müller and Newman 2003, 61). And such "contextual

[8] In experimentally tackling the properties of metabolic systems and their constituents, one fascinating experiment (Boogerd et al. 2005) compares Hamiltonians (representing the function of the total energy of the system) accounting for kinetic processes as constituents of a microorganismal metabolic process, *within* and *outside* the cellular context. Without going into the details of the experiment, the authors of the experiment argue that in a metabolic process of glycolysis, the observed system exhibits certain *inherently systemic properties*, captured by a nonlinear function, which depend on the context of the system, and are not displayed by the parts, or subsystems *in isolation*.The argument is not based on computer modeling, but rather on the so-called *computer experimentation*. In fact, it is virtually impossible to perform an experiment that tests a complex, far from equilibrium, system such as the metabolic system of a microorganism. The number of components that must be accounted for and controlled is immense. Thus, the precise computer kinetic replica of the process reveals the properties of metabolic regulation in the context of the cell, and this can be compared to the properties of the units in isolation.The outcome of the experiment suggests that although inherently systemic properties are explanatory—they explain in terms of the components of the system—such properties are exhibited only in the cellular context. The state-dependent properties can be deduced from the Hamiltonians that account for the energy of the parts of the system in the cellular context, but not from the context-independent properties. The results show that these inherently systemic properties cannot be predicted from the behavior of parts of a simpler system, which is similar to the system that exhibits these inherently systemic properties. Nothing like the inherently systemic properties responsible for the control of the system is exhibited once the system ceases to exist, even though they are critical for the predictions of the system's behavior.Arguably, this may be another case that exemplifies the CNS. Although the systemic properties are measured in individual microconstituents (i.e., the microorganisms composing the system), the value of the overall energy does not reflect its simple distribution throughout the (assembled) system. This seems to suggest that the individual interactions act as intermediaries of systemic properties rather than as merely individual constituents whose energy is distributed throughout the system in a linear fashion. It is more accurate to say that they are structured by the context of the system, rather than vice versa.

information," as we will see shortly, reflects the structuring relevance of the context.

There are two kinds of epigenetic factors relevant to morphogenesis. The "physics of dynamic systems and 'condensed, excitable media'" (Maynard Smith and Szathmáry 1999, 116) constitute the first group.[9] The second group includes "the conditional responses of tissues to each other and to external forces" (Müller and Newman 2003, 8) which influence such conditional responses. Both groups are on a par with "gene sequence variation and gene expression as the primary causal agents in morphological origination" (Müller and Newman 2003, 8). The structuring context arises in the factors of the second group, and becomes critical for pattern formation.

With respect to the relevance of the first epigenetic group, the strings of amino acids, carrying the information on the linear DNA structure, fold up, thus "translating" the initial information to the information on the three-dimensional structure of proteins. Maynard Smith and Szathmáry (1999, 116) point out that "in most cases, the [DNA] string will fold up on its own" and that "folding is a self-organized dynamic process, depending on the laws of physics, which do not have to be programmed" genetically. The self-organizing processes turn out to be critical for pattern-formation as much as "genetic programs" (Meinhardt 2000; Meinhardt 1982).

In fact, the genetic information provides only rough "guidelines" to the cells; the actual pattern-formation rules emerge in the communication among the cells of the tissue which is developing into an organ, via production and inhibition of certain chemicals. Typically, there are two kinds of such chemical substances (i.e., organic tissues) playing opposing roles. The presence of the first, the so-called long ranging *inhibitor* (e.g., the tissue composing the hydra's body), slows down the self-production (autocatalysis) of the second, the so-called *activator* (e.g., the tissue composing the hydra's tentacles).

An underlying mechanism that enables inhibition and activation is *diffusion*; it regulates changes in concentrations of chemical substances, which come in contact with each other. Thus, the net exchange of the molecules between two cells of two different types of tissue (i.e., inhibitor and activator) will depend on their respective molecules' concentration difference.

Although necessary, diffusion alone is an insufficient explanation of why most processes do not simply result in the formation of a homogenous tissue, the most stable state produced by diffusion, and why particular patterns arise from the homogenous tissue instead.

Emphasizing the role of the genes that code for these complicated patterns of activation and inhibition is necessary but also insufficient: simply stated, the gamete is too small and insufficiently structured to contain the very detailed information and rules required to govern the pattern formation (Meinhardt 2000).

This is where the second group of epigenetic factors, the one involving the conditional responses of different tissue cells to each other, becomes essential. As in the case of metabolic control, the relevant explanations are not based exclusively on

[9] See also (Prigogine 1971).

the modeling of the systems at stake, but rely on substantial experimental evidence (Meinhardt 2000; Van Doren et al. 1992; Simpson-Brose et al. 1994; Bissell at al. 2003).

The rate of the activator cells' production of their own molecules (autocatalysis) is *position specific*. In other words, the activator cells increase or decrease their self-production not only with respect to the concentration difference due to diffusion, but also with respect to the positions of inhibitor cells. Each activator cell "knows" the exact position of its neighboring inhibitor cells. *But it also knows (due to the long-range inhibition to be explained shortly) the position of each inhibitor cell in the "flooded" (by inhibitor tissue) area.* Although the activator cells will continually receive the "positional information" on the inhibitor cells via mutual exchange of their molecule concentrations (diffusion), this will be the final step of the process.

Now, only because the exchange itself is structurally contextualized can the entire pattern cause changes in individual activator cells.

Thus, a *common field of competition* structures the activation of genes in individual activator cells, resulting in the formation of an organ's pattern.[10] Each cell receives multiple feedbacks on its own autocatalytic activity by means of multiple feedback mechanisms,[11] but the competition field among the activator cells causes only one of these feedbacks to remain stable in each cell. And which feedback mechanism will survive, and accordingly, which particular gene will switch on in an individual cell, ultimately depends on the entire pattern of the inhibitor "flood" which affects the common field, that is, on the positions of all the inhibitor cells. *The information about the position of each inhibitor cell in the entire area (i.e., the entire inhibitor cells pattern) embedded in the field* is transferred as a "signal" (i.e., the feedback mechanism that survived the competition) to each particular activator cell. This pattern prompts each cell to switch on a particular gene; such switches, in turn, result in a particular pattern-formation. The cell-to-cell exchange of molecules is thus only a last stage in this process, and is almost entirely determined by the outcome of the field activity. A topological collective property of the entire inhibitor pattern structures the interactions between individual constituents, at both the cellular and molecular level, rather than *vice versa*.[12]

[10] Bissell et al. (2003) talk about the "molecular matrix" as responsible for the cellular activity and the dependence of cell's reactions on "structural clues."

[11] A particular substance (a tissue or simply traces of a tissue), the so-called *repressor*, acts as a common feedback catalyst through its ability to repress the genes of both the inhibitor and the activator tissues involved resulting in the common (repressor) field.

[12] Robert Rosen (1972, 1991, 1999) recognized early that the molecular structures in living systems indicate the importance of the context in which they operate and "control" living processes. His concept of *active sites* offers a comprehensive idea of the relevance of the molecular context in such cases. He argues that "[t]his kind of site, is an example of something nonfractionable from the scaffolding that carries it. … a corollary of this nonfractionability is that we cannot get at a functional description of the site from a purely structural characterization of the scaffolding" (Rosen 1991, 54). He developed this idea with respect to the notorious problem of the so-called protein folding, the transformation of the two-dimensional stretch of the DNA molecule into three-dimensional proteins. Rosen also suggests a theoretical model that should account for the functional significance of the cell as the level distinct from the level of its molecular components. He emphasizes the inherently contextual aspect of the model, which provides a general framework for the kind of experimental results that I discuss.

Although the substantial autonomy of the structuring context may not be as firmly established in morphogenetic theory as it is in the theory of metabolic control, the available empirical evidence makes it clear that it is by far the most promising general approach.

Therefore, any satisfying causal account of the CNS must take into account and sort out the hierarchy of the causal relevance of the levels of *individual constituents* or *microconstituents* and the context of the system as a *systemic constituent or macroconstituent*.[13]

In synchronic terms, it is hard to see how the *causal power* of the CNS could be identified with, or reduced to, the causal powers of its microconstituents if we take the relevant explanations seriously. Rather, the *causal power* of the CNS is a result of causal co-determination by two distinct causal constituents, the context of contextually conjoined microconstituents, on the one hand, and the microconstituents themselves, on the other hand. In *dynamic* terms, a microconstituent's property instance does not "single-handedly" cause a change in another microconstituent. Rather, the *causal efficiency* of a microconstituent is responsible for the occurrence of a change only as it is structured by the context. Since the system is causally significant, it must itself be regarded as a causal constituent. But as my detailed analysis in the following sections will show, this view turns out to be unacceptable to the advocates of Kim's reductionism.

Dretske's (1995) distinction between *structuring* causes and *triggering* causes, which he invokes with respect to mental properties and events, may be a useful initial distinction if we are to devise a causal map in the CNS. A structuring collective property does not trigger changes externally in any one microconstituent; rather, it constrains the entire set of interactions among microconstituents. A strong contextual conjoining of all microconstituents is the necessary condition of its occurrence. The system-based causal efficiency, as opposed to the micro-based causal efficiency (of microconstituents) must be invoked in order to capture the causal structure of the CNS. Hence, in terms of their interactions, the microconstituents act as *intermediaries* of highly autonomously, causally relevant constituents, rather than as causal determinants.

It is essential to understand that this distinction between microconstituents and macroconstituents, in terms of their causal relevance is empirically significant. Not every complex system is contextually nonseparable. In fact, a majority of complex systems may be such that their causal explanations do not require us to invoke causally relevant multiple levels. In such cases, the notion of the context is simply shorthand for a simple sum of individuals in terms of either their causal powers or their behavior. But in the case of the CNS, the behavior of microconstituents (e.g., enzymes) would be different were C (e.g., the summation property) never instantiated in the system. C is indispensable as a distinctly contributing causal factor, since the counterfactual claims concerning the causal efficiency of a system which instantiates C are different from the counterfactual claims concerning the

[13] The microconstituents are not identified exclusively with microphysical particles here, as is often the case in the literature. The level of microconstituents should be defined in a minimalist way, that is, with respect to the level higher than the one they define. For example, organic molecules are microconstituents of the organism.

outcomes based solely on the activities of its microconstituents. C must be included as a distinct constituent in the overall causal account of the system.

On one hand, if the microconstituents were the same at *t* with respect to the properties they instantiate as individuals, they would behave differently if *C* were not instantiated at *t*. If they acted as a simple sum of "weakly" conjoined (i.e., conjoined, but lacking an instantiation of *C*) individuals, the outcome would be different. The structuring context of the system is thus indispensable in the structuring behavior of its microconstituents at *t*.

On the other hand, if their causal powers were different (within a certain limit) from those they actually have at *t*, the microconstituents would be constrained (structured) by *C*. This is why the context, rather than the microconstituent, is a structural invariant in the CNS, and, as such, it tolerates a certain extent of change among microconstituents.[14]

3 Does the Ontology of the CNS Challenge Reductive Physicalism?

The vast majority of both reductionists and anti-reductionists would agree that the properties for which science can, in principle, account, including collective properties, have to be instantiated in physical tokens if we understand "physical" in a very broad sense, that is, as synonymous with "natural." Yet simply stating that a token of a collective property must be identical to something (i.e., some kind of token(s)) in the natural world is too general to be interpreted as reductionism, unless we, rather presumptuously, identify reductionism with naturalism. And if we do this, we may render redundant the entire debate on reduction, including Kim's elaborate arguments for reduction.

Kim's reductive physicalism is certainly more specific than this very broad account of naturalism in that its main concerns are the relations between the causal powers and causal efficiency of micro and macroproperties and the tokens that instantiate them.

Kim's early view was that any collective property's (i.e., macroproperty's) causal power in a biological or any other macroconfiguration is reducible to the causal power of the microproperties, which characterize its microconstituents. As has been noted, Kim reformulated his view to allow the existence of a causal power of macroproperties in complex entities, which exceeds that of their microconstituents (this will be discussed in more detail in due course). His reformulated reductionist credo concerns the reduction of causal efficiency, not the causal power of (macro)tokens which instantiate macroproperties.

Arguably, a reductive physicalist, regardless of whether she follows Kim's initial or his revised argument, must be committed to something like the following view concerning the nature of the tokens which instantiate collective properties (or macroproperties, in Kim's parlance):

[14] See also Rosen (1972, 1991).

(A) A token instantiating a collective property *C* is and has to be (identical to) a configuration of fundamental physical tokens in order to be causally efficient.[15]

At first glance, as it is a minimal ontological requirement, (A) does not seem at odds with (Kim's) reductive physicalism. Nor is it clear why one should characterize the CNS as a non-reductive system unless the suggested ontology of the CNS outright denies that (A) is true, which *prima facie* seems implausible. Yet upon closer examination, things become substantially more complex.

(A) might be simply stating that the causally efficient microconstituents that constitute the configuration which instantiates *C* (i.e., the macrotoken) must be causally efficient for the macrotoken to be (indirectly) efficient. This is certainly true in the sense that the microconstituents must be able to interact with each other for the configuration to be fully functional. But is there more to the causal structure of the macrotoken and its microconstituents? And how exactly does this account accommodate the structuring properties of the CNS?

Simply stated, the answers to these questions depend on our understanding of (A). Complicating the issue is the fact that, while it may not be immediately obvious, there are two distinct understandings of (A), namely, the broad and the narrow. Although I will start my analysis with the initial formulation of (A), it will be necessary to formulate it in more precise terms along the way.

It is important to note that (A) does not necessarily exclude the possibility, essential to the ontology of the CNS, that the instantiation of *C* in a (macro)token brings about the macro-based causal efficiency. The microconstituents' causal efficiency is certainly indispensable for a fully functional macrotoken, but may not be the whole story.

On *the broad understanding of (A)*, which I discuss in detail, the token instantiating *C* must be a configuration of the causally efficient microtokens, but *it* is a causal *macroconstituent*, by virtue of its inherent causal power (which exceeds that of its microconstituents) and its macro-based causal efficiency. *The macrotoken is identical to a configuration of its microconstituents in a weak sense, merely with respect to their spatial properties.* The macro/micro identity in terms of spatial properties alone, does not determine whether the microproperties have causal primacy over the macroproperties or *vice versa*. The spatial identity between the macrotoken and the constituent microtokens, or more precisely *the fact that the macrotoken is made up of certain microtokens, and that the microtokens comprise the macrotoken in a certain way, does not simply imply reducibility of the causal power or the efficiency of the macrotoken's properties to that of its microconstituents.* (As we will see, Kim's more recent work clearly indicates that he agrees with this view regarding the causal powers of macroproperties and thus rejects the strong macro/micro identity that he advocated earlier).

Now, the microconstituents might be fundamental, in the sense that their microproperties make them causally efficient, but, as I argue shortly, given the weak macro/micro identity, nothing prevents the macrotoken from being fundamental in the same sense; more specifically, the macrotoken's macroproperty might be causally efficient *as such*, not merely by virtue of the causal efficiency of its

[15] This particular formulation was suggested by one of the reviewers of the paper.

microconstituents. Moreover, as far as the structuring collective properties go, the nature of the macrotoken cannot be captured accurately if its causal function is reduced to a causal efficiency realized exclusively by the microconstituents, as this leaves out its structuring influence, that is, the distinctness and *primacy of its causal (structuring) efficiency over the microconstituents' efficiency*. The causal efficiency of the latter is mediating, while that of the former is determining. (And there is an underlying hierarchy of the properties' and tokens' causal significance in this case, and possibly of the corresponding hierarchy of the relevant laws of nature.)

One could, however, subscribe to a more *narrow understanding of (A)*, one which indicates that the macrotoken cannot be causally efficient *per se*. In other words, if a (property of a) macrotoken is to be causally powerful and/or causally efficient, it can be so only by virtue of the causal efficiency of (a property, or a set of properties, of) its microconstituents. Hence, the macrotoken is identical to/reducible to the microconstituents in terms of both its spatial and causal properties.

I will argue that Kim's physicalism in both its earlier and later phase subscribes to an insufficiently broad understanding of (A) that cannot accommodate the structuring properties indispensable in the ontology of the CNS. There might be other accounts of reductive physicalism that do not commit to the narrow understanding of (A), but Kim's is not one. (The ontology of the CNS that I suggested, which does not commit to the narrow understanding of (A), is better characterized as a nonreductionist ontology, as neither the causal power nor the efficiency of the collective properties of macrotokens reduce to those of microconstituents in any substantial sense.)

Now, whether the broad or the narrow understanding of (A) is presupposed by a given ontology, such as Kim's, that is, whether it presupposes the weak or the strong view of macro/micro identity and reduction, depends on the ontology's account of fundamental physical tokens. Thus, if we wish to assess whether nonreductionists arguing for the ontology of the CNS or reductionists like Kim agree with the narrow or broad understanding of (A), and under what conditions, we need to know more about the structure of the tokens with which the token instantiating *C* is identified. I argue that the *claim made by Kim in his later work on the topic, namely that macroproperties of macrotokens are micro-based (a view and terminology to be explained shortly), commits him to the narrow understanding of (a more precise formulation of) (A), an understanding which is far too narrow for a satisfying ontology of the CNS*.

The second part of (A) explicates the ontological status of the macrotoken at stake by defining it in terms of a configuration. Even more specifically, it is understood to be a configuration of *fundamental physical tokens*. This implies that the fundamental physical tokens (whatever they are) must be *fundamental* in some sense, and, as such, they are constitutive of the configuration.[16]

[16] One could argue that atoms are fundamental tokens of a molecule as a configuration or that elementary particles are fundamental tokens of an atom. Yet we should be able to define a fundamental physical token in a general sense, without referring to representative cases. Once we clarify the concept of a fundamental physical token, and turn to the actual cases, however, it turns out we do not need to appeal to the atomic level or to the level of elementary particles in order to substantially question (Kim's) reductionism. Also, it should be noted that Kim's reductionism stands little chance of being reconciled with the ontology of microphysical particles, as they involve quantum theory (Humphreys 1996, 1997; Silberstein 2001).

Very broadly speaking, the tokens can be physically fundamental in at least two different senses: first, with respect to the spatial properties which they instantiate, and second, with respect to their causal powers and their efficiency by virtue of their properties.

Let us assume that the configuration's constituents are fundamental physical tokens simply in the sense of the spatial properties they instantiate. As this means nothing more than that the location of the configuration is reducible to the spatial locations of its constituents, the only sense of reduction concerns the fact that the spatial properties of the configuration are reducible to the spatial properties of its constituents. But such a notion of reduction is weak, making (A) a trivial claim and showing why reductionists typically think of fundamental physical tokens in a stronger sense, in terms of their causal powers and the efficiency of their tokens.[17]

Typically, reductionists (such as Kim in 1993) identify fundamental physical tokens as the tokens instantiating fundamental *causal powers*. Hence, "a configuration of fundamental physical tokens" in (A) reads as "a configuration of the tokens (the properties of) which have basic causal powers." This substantiates Kim's reductionist credo: "[The Causal Inheritance Principle] If *M* is instantiated on a given occasion by being realized by *P*, then the causal powers of *this instance* of *M* are identical with (perhaps, a subset of) the causal powers of *P*" (Kim 1993, 355). The portion of (A) which characterizes the macrotoken reads as "...a configuration of *the constituents that instantiate fundamental causal powers*." Such an understanding of the tokens as physically fundamental certainly provides a more substantial concept of reduction than the one concerning spatial properties. Thus, (A) reads as "The token instantiating *C* is (identical to) the configuration composed of tokens that instantiate fundamental causal powers." This is to say that the configuration does not instantiate any fundamental causal powers other than those instantiated in its microconstituents. In other words, its causal power does not exceed that of its microconstituents. Both the spatial and causal macro-to-micro reductions (identifications) are retained.

But as already mentioned, Kim abandoned this rather typical reductionist understanding of fundamental physical tokens. Perhaps compelled by recent forceful nonreductionist arguments (Kim 2005, 157–8n), he acknowledges that the "*macroproperties can, and in general do, have their own power, the powers that go beyond the causal powers of their constituents*" (Kim 1998, 85). Further, the "complex systems" which instantiate these macroproperties "obviously bring new causal powers into the world, powers that cannot be identified with causal powers of the more basic simpler systems" (Kim 1999, 36n). Apparently, stating that a token instantiating *C* is a configuration of the fundamental physical tokens, if by

[17] This weak sense of reduction may not be based on anything more than merely an unjustified ontological attitude. Are we really justified in treating the constituents of a configuration as fundamental even in terms of spatial properties? Aren't we assuming too much when stating that the location of the configuration is necessarily spatially reducible to the locations of its constituents, when perhaps it is only coexistent with them?

fundamental we mean "to have (i.e., to instantiate) basic causal powers," and thereby implying that the second-order causal status of the configuration (and *C* that it instantiates), is now unacceptable to Kim. Moreover, the causal power of configuration is, in some sense, superior to (i.e., "it goes over and beyond") that of its microconstituents, which implies that the strong macro/micro identity (reduction) no longer applies. This is an odd statement, coming as it does from the most distinguished proponent of reductive physicalism.

If Kim is right, however, (A) is deficient if it fails to acknowledge that the configuration is not merely a configuration of the tokens which instantiate basic causal powers and if it sticks to the strong understanding of macro/micro identity. Rather, *it* turns out to be a fundamental token (albeit not a microphysical token) in terms of its causal power. In other words, it is a causal (macro, rather than micro) constituent itself. Thus, in order for a token instantiating *C* to be causally efficient, it has to be, or at least it may be (identical to) a configuration of microphysical tokens *as a fundamental (macro-physical) token* (i.e., a causally powerful token), rather than merely (identical to) a configuration of fundamental (micro) physical tokens (i.e., microconstituents).[18]

At this point, we need to substantially reformulate (A) so that it accommodates Kim's (revised) view of the causal powers of macroproperties and the tokens that instantiate them. We can do this in the following way:

> (A)' A token instantiating a collective property *C* is and has to be (identical to) a *fundamental* (macro) physical configuration composed of microphysical tokens (where "composed of" does not diminish the causal significance of the configuration itself) in order to be causally efficient.

Although (A)' seems both plausible and sufficiently general, it is not *prima facie* clear that all reductionists would accept it, since Kim's earlier straightforward concept of the reduction of the causal power of the macroproperty to the properties of the microconstituents, as instantiated by the configuration, has now vanished. Even if we take (A)' to involve reduction of some kind (e.g., since configurations turn out to be basic physical tokens, the causal power of their constituents could be treated as reducible to the causal power of the configurations themselves), it is still far removed from early Kim. And the second part of (A) certainly does not imply such a kind of reduction if it is understood to disqualify the configuration as causally powerful by committing to the strong view of macro/micro identity.

The question remains: if we accept Kim's later view of configurations' causal powers, is there an interpretation of fundamental physical tokens that provides a substantial sense of reduction?

Analyzing the limitations of the view that the causal power of the configuration exceeds the causal powers of its microconstituents will help us answer this question. According to Kim, the causal powers of macroproperties are "the causal powers of microstructural, or micro-based properties, of a complex system" (Kim 1999, 36).

[18] See endnote 13.

By this, Kim means that the causal powers of macroproperties, although irreducible to the powers of microconstituents in the above-stated sense, are nevertheless *realized by these microproperties* (hence, they are *micro-based* properties) (Kim 1998, 84). While he admits that the causal power of macroproperties is novel and cannot be identified with the causal powers of the microconstituents, he argues that the causal power of a macroproperty is "anchored in" or "based in" (Kim 1998) the microproperties of its constituents. This means that any *exhibiting* of this novel causal power of the whole is realized by (and thus reducible to) *the exhibiting* of the causal powers of the microconstituents.

Thus, the constituents of the configuration are fundamental physical tokens in the sense that they, not the configuration as it instantiates the causal power of C, are responsible for causing events/property instances. A macroproperty, instantiated in a configuration, is causally efficient only by virtue of the causal efficiency of the microproperties instantiated in its microconstituents. It appears that we can retain a substantial sense of reduction, while rejecting the strong macro/micro identity: every macro-event, which exhibits the configuration's causal power is reducible to/explainable in terms of microinteractions (i.e., interactions among microconstituents).

Kim's initial reductionist credo (The Causal Inheritance Principle), then, must be reformulated in terms of causal efficiency so that the *causal efficiency* of a macroproperty, realized by the *causal efficiency* of the instances of (a subset of) the microproperties, is identical to them. Let us call this a revised credo in which Kim attempts to establish macroproperties as a kind of "causal epiphenomena": although they have causal powers that exceed the power of the microconstituents, they owe their causal efficiency to them,[19] and the configuration is not a causal constituent in the full-fledged sense.[20]

Thus, (A) reformulated in order to accommodate the understanding of fundamental physical tokens *in terms of causal efficiency* must read along the lines of Kim's reformulated credo. One can attempt the following reformulation:

[19] After all, the goal of Kim's argument has been to undermine nonreductionism, not epiphenomenalism. In his earlier writings, in the face of the supposed failure of nonreductionist physicalism, he considered epiphenomenalism a viable, albeit not very satisfying, option (Kim 1993).

[20] But, is this a tenable position? A machine for producing sausages has the power to make a sausage: none of its constituents alone does. A *simple* sum of its constituents (e.g., a pile of the parts of the disassembled machine) does not have this power. But does *a configuration* of the constituents (i.e., a properly assembled machine), which are "at rest", have this power? The answer depends on how we cash in the extra causal power of the configuration. What does the distinction between causal power and causal efficiency add to our understanding of the system, if anything?One way to answer these questions, while adhering to Kim's intention, may be to say that only if the constituents are "at work," in other words, if their causal powers to interact with each other in certain ways are *realized* (i.e., they are not mere powers any more, but rather realized powers), can the machine, as a whole, produce sausages. But even so, could we not conclude that the causal power of the whole is identical with or at least reducible to the causal powers of the constituents *vis-à-vis their mutual interactions*? What else is there to the causal power of the whole? The reduction, or the lack of it, in terms of causal powers concerns the synchronic aspect of the entities and of the properties they instantiate. The causal efficiency, however, concerns the way these are exhibited, which is a *dynamic* aspect of entities and their properties. Whether it could be successfully argued that synchronically irreducible systems can be diachronically reducible (which seems to be the consequence of Kim's position) remains unclear, but we will leave this question unresolved and return to our central issue.

(A)'' A token instantiating a collective property *C* is and has to be (identical to) a configuration of *causally efficient tokens* in order to be causally efficient.

But even this may not be specific enough, as the narrow understanding of (A) which accords with Kim's reformulated credo must eliminate the possibility that the macrotoken instantiating *C* is inherently and micro-independently causally efficient, and the possibility that the microconstituents have a "subdued" (mediating) causal efficiency (i.e., it must avoid the broad understanding of (A)'').

Regardless of whether or not he accepts that configurations have causal powers exceeding those of their constituents, Kim argues that the constraining power as the manifestation of the causal efficiency of the whole is based on (depends on/supervenes on/is anchored in) the causal efficiency of the microproperties of constituents *(i.e, the microproperties are always "micro-based" or "microstructured"*!*).* A sausage-producing machine is causally efficient because its causal efficiency is identical with (and thus reducible to) the causal efficiency of its constituents: the machine's property of being a sausage-producer is causally efficient solely because of the causal efficiency of each of the constituents as they interact. Thus, any exhibiting of the causal power of a configuration is done via interactions among microconstituents. The latter are the distributors of the causal power of the former.

Given that the causal efficiency of the properties of the latter is instantiated in that of the former, the exclusive kind of interaction, which can occur according to revised Kim's credo is microinteraction *simpliciter*, an interaction taking place among microconstituents of a configuration. Here, the microconstituents exhibit the causal power of the configuration, realizing *C*'s causal power by means of another property, say *P*, that they instantiate. *P* serves as a "causal vehicle" of *C*, and talk of an independent or even highly autonomous causal efficiency of the configuration's *C* in terms of constraining or structuring the constituents becomes merely metaphorical. In the course of interactions, the appearance of each microconstituent "adjusting" to the configuration's structure is reduced to its immediate causal interactions with its neighbor microconstituents. It is not surprising, then, that Kim declares the Principle of Supervenience to be the broadest possible criterion of physicalism (Kim 1998, 15). And this is why only the narrow reading of (A)'' accords with revised Kim's credo: namely, a configuration is causally efficient *only by virtue of the causal efficiency of its microconstituents*.

According to Sperry's well-known argument (Sperry 1969), the motion of the molecules in a wheel is constrained by the wheel as a whole: this property of the wheel cannot be reduced to the properties of its constituent molecules. This is an example of an emergent property, Sperry argues. Kim (1999, 30–31) points out that he (and physicalists like him) can accept that the structure of the wheel constrains the motion of its molecules. But this is acceptable only because it does not deny that the whole's causal efficiency (regardless of whether its causal power exceeds that of its constituents) is supervenient on the interactions among its microconstituents. Kim insists that "emergent downward causation should not *simply be identified* with causation from properties of the whole to properties of its own parts" (Kim 1999,

31; emphasis added). Nothing suggests that the properties of the whole, although perhaps exceeding the causal power of properties realized in the microconstituents, are not rendered causally effective by being realized by the causally efficient microproperties of the microconstituents.

The instances of emergent downward causation involving independently causally efficient macroproperties must be substantiated by more than cases like Sperry's wheel: it is essential to demonstrate that an identification of emergent downward causation with whole-to-(micro)constituents causation challenges the narrow reading of (A)''.

Sperry's wheel is nothing more than an example of the microinteractions *simpliciter*, at least not in any obvious way, and invoking it to argue for emergent downward causation is an unsatisfying "simple identification." Yet not all configurations are like this one. The nature of living configurations, at least those captured by the ontology of the CNS, and the nature of microinteractions characterizing them, are fundamentally different. (For the purposes of clarity, I will call such configurations *systems* in order to distinguish them from configurations like Sperry's wheel.)

A microconstituent's adjustment to a CNS is structured by the system and intermediated by the microconstituent's causal efficiency. A macroproperty's exhibiting of causal power is not necessarily a single-layered process as Kim's revised credo suggests. The system's *C* structuring of a microconstituent's behavior is not exhaustively explained by its causal power being realized through the (immediate neighbors of) the microconstituent, because *C*'s being realized in a neighboring microconstituent is only one component of the overall causally relevant structure.

The behavior of any microconstituent depends on whether or not C is instantiated in all other or at least in the vast majority of the system's microconstituents. Thus, since *C* exhibits a structuring influence over the microconstituents which are doing the mediating causal work, the assumption that the causal influence of *C* is simply a metaphor and can be explained in terms of the causal actions of microconstituents is misleading. In short, we need an additional and quite distinct contextual layer of causal influence within the system.

Admittedly, in the case of cell signaling, the cell's change of behavior will be due, in part, to the "information" (i.e., the molecular concentration level) passed on by its neighboring cells. But such immediate signals are structured by the information concerning the entire pattern (i.e., positions and concentrations) of inhibitor cells: the activator cell receives information about the extent of the presence of multiple inhibitor cells at multiple positions (embedded in the common (repressor) field and passed to individual cells—see Sect. 2), and this causes an individual cell to act in a certain way. Furthermore, this "collective information" is passed on to individual activator cells, resulting in a new pattern (of an organ).

This signaling pattern embedded in the common (repressor) field and distributed to individual cells through intermediating individual cell activity cannot be left out of our analysis. Indeed, the configuration of inhibitor cells is a

fundamental macroconstituent, as far as the exhibiting of fundamental causal powers goes, because it affects the activator cells' ability to change their concentration.[21]

Similarly, in the case of metabolic control systems, whether an enzyme dominates the flux depends not only on its interactions with other enzymes, but also on both (the regulative influence of) the mutual interactions among these other enzymes and the number of enzymes in the system.

One could try to interpret the collective properties as micro-based, that is, as relational properties instantiated in the microconstituents which determine how exactly the constituents will relate to each other. In pursuing this strategy, it is not enough to define a microconstituent *e* in terms of its basic causal power and its causal efficiency. Rather, this must be amended by considering the way in which the constituent will interact with another constituent of the same kind *e'* by taking into account the "nature" of *e'* (i.e., its capacity to interact with *e*). Although *C* maps potential relations between microconstituents, it may be *micro-based* after all: it is causally efficient by virtue of being realized in a microconstituent.

Given this, one could perhaps argue that micro-based *C*, instantiated in *e*, somehow maps the interactions of the constituents of the entire system. If this were so, however, all microconstituents would have to instantiate the same "map" of microinteractions as they react to the change in the system in a coordinated manner (the lack of one enzyme affects all other enzymes, thereby affecting the enzyme that would otherwise be dominant).

But this seems a roundabout way of stating that the causal efficiency of *C is* partially *system-based*, thus affecting the causal efficiency of microconstituents' properties and causing *them* to be system-based. More precisely, the causal efficiency of *e* is system-based since the constituent is causally efficient by virtue of *C* being instantiated *both in* e *and in every other constituent of the system*. Put otherwise, its most immediate interactions with other constituents are a result of the "history" of the interactions among all other constituents. This "history," however, is really *the structure (defined by C) of all these previous interactions that is embedded in (distributed to) these immediate interactions*. Thus, these interactions are only causal microintermediaries of the system's structure (i.e., all of the interactions as they are structured by *C*—e.g., the summation property in metabolic systems of *n* enzymes), not its determinants.

Nor is simply listing all interactions sufficiently informative in terms of understanding the system's nature: we need to add an explanation concerning the transfer of such information to particular constituents. This must involve *C* as the structuring-contextual property of the system, as it determines the course of the

[21] Kitcher (1984) argues that the indispensability of the level of explanation of topological properties that supplements the genetic level supplied by molecular biology, in morphogenetic explanations, constitutes an argument against reductionism. He argues against Nagel's theory-reduction model. Yet, appealing to multiple level explanations alone does not substantiate a sufficient response to Kim's argument. One must demonstrate and specify *the causal relevance* of these epigenetic levels of morphogenesis.

(listed) microinteractions.[22] As they interact, the constituents are the intermediaries of *C*, and, thus, the configuration depends on their causal efficiency, but this is not the whole story: the system exhibits a causal influence that does not (passively) supervene on the causal efficiency of the constituents but rather (actively) structures it. Both types of interactions happen simultaneously.

All this may be true to some extent of any structurally constrained system and may be inherent to mechanisms characterizing them. But one should be careful not to predicate one's assessment of the ontology of the CNS on analogies between living systems and simple systems such as Sperry's wheel or a brick wall. The movement of a brick in the brick wall is determined by the constitution of the wall through its interactions with the bricks attached to it. But these immediate interactions are simply intermediaries of the (topological, in this case) structure of the interactions of other microconstituents. In such cases, one could say that the constraints imposed by the configuration are nothing but the boundary conditions of the environment limiting the system (e.g., the brick wall is constrained by the iron frame of the house).

But in the case of the metabolism of morphogenetic systems or the systems of metabolic control, the level of complexity is much higher, thereby enabling the system to (continuously) create its own internal constraining condition(s). For example, the summation property imposes constraints on the interrelations of all the microconstituents (enzymes); a constraint is not externally imposed on the boundary microconstituents (e.g., the bricks in contact with the iron frame) and then transferred by microinteractions throughout the system. Similarly, the constraints in morphological organization are defined by intra-systemic connections rather than external pressures (Rasskin-Gutman 2003, 306).

Entities of a certain structure and complexity are required if these internal systemic constraints are to emerge.

Systemic internal constraints affect the way enzymes interact, while constraints imposed by the boundary conditions define the way in which the bricks in the wall interact. Living systems are defined, in part, by the refinement and complexity of such structuring properties, and this is why they should be distinguished in ontological terms (i.e., in terms of their causal power and efficiency) from the systems of such entities as bricks.[23] In fact, as far as the CNS and similar systems go, it may be more advantageous to talk about the *causal relevance* (or significance) of the properties and the entities they instantiate rather than the causal efficiency of

[22] Although, we usually do not know exactly how and in what order the interactions among *n* constituents take place due to the sheer complexity of the CNS, we know that they embed the structural/contextual property of the system, not only at one, but at multiple levels. The collective property might be micro-mediated by constituents' interactions in a number of ways. For instance, they can develop various shortcuts in the system. The interactions among microconstituents might also embed a "random structure" or decontextualized property. But in our examples, they are embedding a very specific structure, a very specific contextual property invariant across the interactions among constituents.

[23] See also (Thompson 2007).

entities, as the former is a broader term that may accommodate both structural and triggering causes.[24]

This cannot be the case if (A)'' implies the reducibility of the causal efficiency of macroproperties to the causal efficiency of the properties of microconstituents. Only the broad understanding of (A)'' can accommodate the CNS, prompting the breakdown of reduction, not only in terms of causal powers but, more importantly, in terms of causal efficiency.[25] This broad reading of (A)'', however, which is the only reading that can accommodate the CNS, is ultimately unacceptable to Kim's type of physicalist reductionist, since it disposes of the revised Kim's credo and the exclusively micro-based causal efficiency of macroproperties. Moreover, it pushes the reductionist to the last line of defense, the accusation of CNS incoherence.

4 Is the Ontology of the CNS Coherent, and Does it Avoid Overdetermination?

A reductionist might argue that in subscribing to the broad interpretation of (A)'', the nonreductionist assumes an incoherent concept of downward causation and/or runs into difficulty with overdetermination. Unless one is willing to advocate an unviable form of dualism, the reductionist argues, *C*'s causal efficiency must be regarded as autonomous for practical purposes only.

With respect to the first objection, Kim argues that downward causation, as defined by a causally efficient collective property, must be incoherent:

> At a certain time *t*, a whole *W*, has emergent property *M*, where *M* emerges from the following configuration of conditions: *W* has a complete decomposition into parts *a1*, ..., an; each *ai* has property *Pi*; and relation *R* holds for the sequence *a1*, ..., *an*. For some *aj*, *W*'s having *M* at *t* causes *aj* to have *Pj* at *t*. (Kim 1999, 28)

Kim finds this simultaneous upward (*aj*'s having *Pj* resulting in *W* having *M*) and downward (*aj*'s having *Pj* as an effect of *W* having *M*) determination unacceptable and even incoherent (Kim, ibid.).

[24] An even stronger claim is that the concept of *interactions* is not a particularly helpful way of characterizing the processes taking place in the CNS as it usually goes hand in hand with Kim's revised credo. In such cases one should talk of *state transitions* or *transformations of systems* in such cases. For an insightful discussion of the limitations of the concept of interactions in biology, see (Machamer 2000).

[25] The research I have discussed so far only indirectly concerns our understanding of evolutionary mechanisms. The metabolic and morphogenetic systems do not depend on any (long-term) time-evolution properties. However, as these systems might be subjects of natural selection, the structural context might be indirectly relevant to evolution. But it might be relevant in a more direct way as well. Providing that natural selection is a population phenomenon (Millstein 2006), the causal account of natural selection may require us to introduce the so-called proximate and ultimate causes. Thus, for instance, as sexual selection causes modifications of the share of traits in populations, not of individual traits, tracing the modifications of a trait from ancestor to ancestor does not provide insight into selection. The ultimate cause responsible for selection, therefore, is instantiated at the level of population, and as such, it determines the outcome at the level of individuals (changed directly by proximate causes) by selecting them (Sober 1984; Lewontin 1983; Wouters 2005).

In response, Kim defines both downward and upward determination in terms of the micro-micro interactions, leaving out the possibility that downward causal determination is characterized by a structuring cause and does not affect microproperty instances in the same way as these instances affect each other. But it is presumptuous to define downward causation in such a way, since there might well be an empirically significant difference between interactions in the CNS and microinteractions *simpliciter*. Instead of attempting to *a priori* dismiss the possibility of a coherent ontology that can capture such processes, we should rather seek an alternative to the narrow account of causation on which reductionism is predicated. This becomes even more obvious with respect to the overdetermination objection.

Does the ontology of the CNS suffer from the overdetemination problem as Kim defines it?[26]

First, if we follow Kim's view that he advocates in his later work, we should understand the overdetermination argument as a concern with the *exhibiting* of causal powers (or in other words, with the causal efficiency of entities and their properties), rather than a concern with their causal powers *per se*.[27]

Second, the argument works well against strong dualism, where two events, while they are both candidates for causing another property instantiated in the basis, belong to distinctly different ontological categories. But can we say that the two events in question, which are of *the same ontological kind*, say two physical events, can act as distinct causes in order to produce an instance of *P'*? If we respond in the spirit of Kim's argument, we do not see them as distinct: they are parts of the same physical cause and do not violate the CEP (the Causal Exclusion Principle). For instance, when a car hits another car, the fact that the slippery road and the car's speed are equally responsible for the crash is not a good enough reason to designate either as a distinct cause. Any introduction of distinct causes of this kind is redundant and points to an unacceptable dualism.

This is why the reductionist would insist that the postulation of two distinct, causally efficient levels of a configuration (namely macro- and micro-levels), which act simultaneously in order to produce an effect, should not be understood literally. Rather, they represent an epistemological shorthand for a system whose causal efficiency is predicated on microinteractions. Thus, a single cell can never be acted upon by both the configurations' *C* and another property *P* of another cell, as distinct causally efficient instances: in a final analysis, they must be part of the same causally efficient event of the *P*-kind, realizing *C*'s causal potential through *P*'s action.

Such causal analysis does not challenge the CEP, being predicated on a very specific way of determining which kind of events can count as ontologically/causally distinct, and leaves out the possibility of *the macro-micro interactions* exhibited in the CNS. We could try to interpret all causal events in a way that

[26] See Sect. 1.

[27] Again, we will leave aside a possible problem with the discrepancy between irreducible causal powers of at least some configurations on the one hand, and reducibility of their causal efficiency to that of their constituents on the other, as it does not seem to affect the overdetermination argument.

preserves the CEP and sticks to the narrow understanding of (A)''. But that would be putting the cart before the horse, as some causal events, at least those related to the CNS, might require causal explanations that involve *the macro-micro interactions*, which are not allowed by Kim's reductionism predicated on the CEP. Or they might square with the CEP, if it is understood in a broad sense where "physical cause" is synonymous with "natural cause." In such instances, the CEP would allow that every (caused) physical event has a set of physical causes. This, in turn, agrees with the broad understanding of (A)'', which, however, disqualifies both early and revised Kim's reductionism.

We should not limit our ontology to an armchair analysis that produces a narrow *a priori* understanding of what causation must be and try to force our science into such framework. Instead, we should inform our ontology (i.e., the view of entities and their properties) by our knowledge of various ways in which the causal powers are exhibited in the world and only then devise a suitable account of causation and ontology.

Our account of causal powers should be derived from our empirically informed account of the structure of the causal efficiency of natural systems, rather than limited by the assumptions concerning it. Sticking to the narrow understanding of (A)'' prevents us from accounting for configurations in a way that may be indispensable to biological explanations. We need to introduce two ontologically distinct ways in which entities in the CNS exhibit their causal powers. The reductionist hopes that this indispensability is epistemological in nature, simply shorthand for an *in principle* viable reductionist account (that sticks to the narrow understanding of (A)''). But it is necessary to hold on to this hope, albeit an empirically suspect and faint hope, if one subscribes to a narrow account of causation.

Limited by Kim's reductionism and by the narrow understanding of (A)'', the CEP disregards both the causal relevance of the contextual component of the system and the exact nature of the microconstituents' causal relevance. The causal efficiency (activity) of wholes and their parts is not necessarily structured in terms of the supervenience of the former on the latter. Any satisfying causal account of the CNS will acknowledge this and sort out the (hierarchy of) causal relevance of multiple entity and property levels. These turns out to be irrelevant in terms of the causal account predicated on the CEP as it is incorporated in reductionist physicalism.

Suitable causal accounts may already be available, however. For example, Spohn (2006) suggests a Humean account of causation that might satisfy such requirements. And I have already mentioned that Dretske's distinction between structuring and triggering causes may be particularly helpful in pointing out the key aspects of causal analysis overlooked in the reductionist analysis.

As the causal structure of interactions defined in terms of such reciprocal causal relevance of multiple levels is not predicated on the acceptance of the CEP (wedded to reductionism), it avoids the overdetermination problem. Such explanations should not be suspected of any devastating kind of dualism, as they talk of the same kind of entities and properties that they instantiate, namely, natural entities and properties in the broad sense (not physical entities and properties in a narrow

reductive physicalist sense), which are nevertheless distinct in terms of the ways in which they exhibit their causal powers.[28] Further, such non-CEP accounts of causation allow for the multiplicity of ways in which the causal powers of entities are exhibited, and account for the rich fabric of the natural world. They do not become a metaphysical liability.

Acknowledgments I would like to thank Evan Thompson, Jagdish Hattiangadi and Ljiljana Radenovic for discussing their ideas on the topic of the paper with me. I am also grateful to three anonymous reviewers and Thomas Reydon, the Managing Editor of the journal, for their very helpful comments and patience, and Elizabeth Thompson for her help with proofreading the manuscript.

References

Bissell M, Mian IS, Radisky D, Turley E (2003) Tissue specificity: structural cues allow diverse phenotypes from a constant genotype. In: Müller GB, Newman SA (eds) Origination of organismal form. MIT Press, Cambridge, MA, pp 103–119

Boogerd F, Bruggeman FJ, Richardson RC, Stephan A, Westerhoff HV (2005) Emergence and its place in nature: a case study of biochemical networks. Synthese 145(1):131–164

Dretske F (1995) Mental events as structuring causes of behavior. In: Heil J, Meele A (eds) Mental causation. Clarendon Press, Oxford

Healey RA (1991) Holism and nonseparability. J Philos 88:393–421

Humphreys P (1996) Aspects of emergence. Philos Topics 24(1):53–70

Humphreys P (1997) How properties emerge. Philos Sci 64:1–17

Humphreys P (2007) Emergence of patterns. Erkenntnis, forthcoming

Kascer H, Burns JA (1973) The control of flux. Symposia Soc Exp Biol 27:65–104

Kascer H, Burns JA (1981) The molecular basis of dominance. Genetics 97:639–666

Kascer H, Porteous JW (1987) Control of metabolism: What do we have to measure? Trends Biochem Sci 12:5–14

Keightley PD, Kascer H (1987) Dominance, pleiotropy and metabolic structure. Genetics 117:319–329

Kitcher P (1984) 1953 And all that: a tale of two sciences. Philos Rev 93:335–373

Kim J (1993) Supervenience and mind. Cambridge University Press, Cambridge, MA

Kim J (1998) Mind in a physical world. MIT Press, Cambridge, MA

Kim J (1999) Making sense of emergence. Philos Stud 95:3–6

Kim J (2005) Physicalism, or something near enough. Princeton University Press, Princeton and Oxford

Lewontin R (1983) Darwin's revolution. NY Rev Books 30:21–72

Maynard Smith J, Szathmáry E (1999) The origins of life. Oxford University Press, New York

Machamer P, Darden L, Craver CF (2000) Thinking about mechanisms. Philos Sci 67:1–25

Meinhardt H (1982) Models of biological pattern formation. Academic Press, New York

Meinhardt H (2000) Biological pattern formation as a complex dynamic phenomenon. In: Carbone A et al (eds) Pattern formation in biology, vision and dynamics. World Scientific, NJ

Millstein RL (2006) Natural selection as a population-level causal process. Br J Philos Sci 57(4):627–653

Müller GB, Newman SA (2003) Origination of organismal form: the forgotten cause in evolutionary theory. In: Müller GB, Newman SA (eds) Origination of organismal form. The MIT Press, pp 3–11

Nijhout HF (2001) The ontogeny of phenotypes. In: Oyama S, Griffiths PE, Gray RD (eds) Cycles of contingency. MIT Press, Cambridge, MA, pp 131–140

Prigogine I (1971) Unity of physical laws and levels of description. In: Grene M (ed) Interpretations of life and mind. Humanities Press, New York

Rasskin-Gutman D (2003) Boundary constraints for the emergence of form. In: Müller GB, Newman SA (eds) Origination of organismal form. MIT Press, Cambridge, MA, pp 305–322

Rosen R (1972) Some relational cell models: the metabolism-repair system. In: Rosen R (ed) Foundations of mathematical biology. Academic Press, New York

Rosen R (1991) Life itself. Columbia University Press, New York

[28] See also (Humphreys 2007, forthcoming) and (Healey 1991).

Rosen R (1999) Essays on life itself. Columbia University Press
Silberstein M (2001) Converging on emergence. J Consciousness Stud 8:61–98
Simpson-Brose M, Treisman J, Desplan C (1994) Synergy between the *hunchback* and *biocid* moprhogens is required for anterior patterning in *Drosophila*. Cell 78:855–865
Sober E (1984) The nature of selection. MIT Press, Cambridge, MA
Spohn W (2006) Causation an alternative. Br J Philos Sci 57:93–119
Sperry R (1969) A modified concept of consciousness. Psychol Rev 76:532–536
Thompson E (2007) Mind in life: biology, phenomenology and the sciences of mind. Harvard University Press, Boston, MA
Van Doren M, Powell PA, Pasternak D, Singson A, Posakony JW (1992) Spatial regulation of proneural gene activity: auto- and cross-activation of *achaete* is antagonized by *exctramacrochaetae*. Genes Dev 6:2592–2605
Wouters A (2005) The functional perspective of organismal biology. In: Reydon TAC, Hemerik L (eds) Current themes in theoretical biology. Springer, Dordrecht, pp 33–69