Levels, Emergence, and Three Versions of Downward Causation

by Claus Emmeche*, Simo Køppe** and Frederik Stjernfelt***

published pp. 13-34 in: Peter Bøgh Andersen, Claus Emmeche, Niels Ole Finnemann and Peder Voetmann Christiansen, eds. (2000): *Downward Causation. Minds, Bodies and Matter*. Århus: Aarhus University Press. (contents of this book, see here). Page shifts in the published version is indicated below by the sign //p.x/.

* Niels Bohr Institute, Blegdamsvej 17, DK-2100 Copenhagen Ø, Denmark.
** Psychological Laboratory, Njalsgade 94, DK-2300 Copenhagen S, Denmark.
*** Department for Comparative Literature, Njalsgade 80, DK-2300 Copenhagen S, Denmark.

//p. 13/

Levels, Emergence, and Three Versions of Downward Causation

Claus Emmeche, Simo Køppe and Frederik Stjernfelt

Abstract

The idea of a higher level phenomenon having a downward causal influence on a lower level process or entity has taken a variety of forms. In order to discuss the relation between emergence and downward causation, the specific variety of the thesis of downward causation (DC) must be identified. Based on some ontological theses about inter-level relations, types of causation and the possibility of reduction, three versions of DC are distinguished. Of these, the `Strong' form of DC is held to be in conflict with contemporary science; the `Medium' version of DC may for instance describe thoughts constraining neurophysiological states, while the `Weak' form of DC is physically acceptable but may not in practice be a feasible description of the mind/brain or the cell/molecule relation. All forms have their specific problems, but the Medium and the Weak version seems to be most promising.

Introduction

The concept of Downward Causation (DC) has a rather diffuse origin and it is difficult to ascertain the first use of the concept.[1] In any case, downward causation must presuppose the assumption that several levels of reality coexist, be it merely as levels of description or as levels of description as well as of ontology. Together with level theories, the concept of emergence is very often used as a designation for the relation between the new or unpredictable property on the higher level //p. 14/ and its basis in the lower level. As a kind of immediate extrapolation of this idea, downward causation is used as a designation for an alleged downward effect which emanates from the emergentically defined higher level onto its constituents in the lower level. Thus, as Jaegwon Kim has put it, "... downward causation is much of the point of the emergentist program" (Kim 1993, p. 350). Thus, to maintain a theory of ontological levels and emergence, a rational interpretation of the concept of Downward

Causation (DC) must be given. This shall be our aim in this paper. The sections 1 and 2 will contain some remarks on the concepts of level, causation, form, and substance, which we shall use for discussing the distinct versions of the downward causation concept.

Inclusivity of levels

Several attempts have been made to formulate emergence theories of levels of organization, description, systems, or reality. In many ontological versions of these theories, the entities of each level are considered as being irreducible to those at the level below. The vitalism/reductionism debate in the life sciences shows that a concept of emergence as something in principle inexplicable will often be falsified by the history of science. Nevertheless, the concept of emergence keeps reappearing in various sciences (most recently in complex systems research and artificial life), and it seems that it cannot easily be dispensed with.

We have elsewhere[2] argued for an ontological non-reductionist theory of levels of reality which includes a concept of emergence, and which can support an evolutionary account of the origin of levels. Though the interpretation of *emergence* as `the creation of new properties' involves several philosophical problems (such as specification of the vague terms `new' and `creation'), the intuition involved in the basic idea does not refer to anything mysterious. The concept can be defined formally[3] so as to conform to a more careful standard interpretation of emergence, according to which "a property of a complex system is said to be `emergent' just in case, although it arises out of the properties and relations characterizing simpler constituents, it is neither predictable from, nor reducible to, these lower-level characteristics."[4] The ontology of levels we attempted to give was framed in a materialist and evolutionary perspective that implied that the relation between levels was considered to be *inclusive*, permitting the `local' existence of different ontologies, all included within the physical level and non-violating physical laws. We identified, as a working hypothesis, four primary levels -- the physical, the biological, //p. 15/ the psychic and the social level, these having nonhomomorphic inter-level relations. *Inclusiveness* entails the two theses that (a) the evolutionary emergence of a new higher level from the physical (all-inclusive) level does not violate or change any physical laws in spite of the appearance of a new irreducible level of organization (and generally, that lower-level principles of organization are not altered by the emergence of higher levels); (b) the biological ontology is *local* to the extent that different biologies, different organizing principles of life, may emerge on other planets (who knows if life universally takes shape as the natural selection of DNA-coded genotypes?).[5] Furthermore, the inter-level relations are *nonhomomorphic* in the sense that the emergence of the biological from the physical level does not have the same complex of inter-level relations of dependence as the emergence of the social and psychic levels from the biological one due to the continuous mutual conditioning and interdependence between emergent psyche and sociality. The processes that lead to the first-time emergence of the biological level differ not only materially but also in a formal ontological way from the processes that constitute the psychological and the sociological level: for the latter two, involving the emergence of self-consciousness and institutions, these level-constituting processes are interwoven and depend on both intersubjectivity and language; while for the biological level, they depend upon specific conditions at one single level, the physical one (leading to the evolution of the first cells). But the different cases of emergent phenomena must share some formal features that distinguish them as emergent in comparison with non-emergent phenomena.

One can argue at length about the number of (and demarcations between) the primary levels. Our choice of the four levels mentioned was in part pragmatical (thus, multicellular life and non-self-conscious psyche are serious candidates for further primary levels), but what is ontologically important is that such levels of reality can in fact be rationally distinguished. Our

methods for making such distinctions are of course dependent on the historical development of scientific theories and disciplines; thus, one may conceive of arriving at an even more fine-grained theory of levels. By scientific we mean the natural sciences, which deal primarily with phenomena at the physical and biological level, as well as the humanities and the social sciences, which deal with phenomena at the other levels.

Recent research in self-organizing non-linear dynamical systems represents a revival of the scientific study of emergence, and it can be argued that these developments are the final `devitalisation' or demystification of emergence and thus may also help to clarify the concept of downward causation.

//p. 16/

Form, substance, causality

We shall try to discern three different conceptions of downward causation and exemplify them by concrete descriptions of alleged emergence, that is, biochemical system --> (emergence of) cell, and nervous system --> (emergence of) psychic system, two of the most classic level borders in the various level theories. Before doing so, we shall formulate two postulates or hypotheses, each of which is given in its inverse variant, to function as an axis in the discussion.

In level theories, the concept of *entity* is used as a designation for the unit that is constitutive for a given level. A level is thus characterized by a certain primary entity possessing the emergent property defining the level. Hence, the specific conception of this entity is crucial as regards the kind of level theory for which one opts.

1 a *Constitutive reductionism*

> Ontologically or materially, a higher level entity (for instance a biological cell) *consists of* entities belonging to the lower level (the cell consists of molecules). These lower level entities are constituents of the higher level and are organized in a certain way that yields the higher level entity (the cell). This does not mean that the higher level can be reduced to the lower (in which case no levels would be relevant), but that the higher level does not add any substance to the entities of the lower level.

1 b *Constitutive irreductionism*

> Ontologically or materially, a higher level entity is constituted by the lower level, but even if the lower level entities are a necessary condition for the higher level, this higher level cannot be reduced to the form or organization of the constituents. Thus, the higher level must be said to constitute its own *substance* and not merely to consist of its lower level constituents.

2 a *Formal realism of levels*

> The structure, organisation or form of an entity is an objectively existent and irreducible feature of it. The specific form characterizing a higher level entity (organizing its lower level constituents) cannot be reduced to lower level forms or substances.

2 b *Substantial realism of levels*

> A higher level entity is defined by a substantial difference from lower level entities. The morphological or organizational aspect is a necessary but not sufficient condition of a higher level entity. Through emergence, an ontological change in substance takes place.

These hypotheses are related in such a way that 1a and 2a are often connected in a given argument, as are 1b and 2b. Most theories of downward causation can be //p. 17/ placed under one of these two headlines, depending on whether it is the first or the second set of assumptions which (most often implicitly) founded the theory.

Finally, we shall use a reinterpretation of the traditional four Aristotelian types of causality (which of course are not exclusive), and thus distinguish between the following:

(i) *Efficient causality*

> is a temporal cause-effect relation involving an interactional exchange of energy pertaining to the entities of a given level. It results in a temporal sequence of states being causally interrelated. In everyday language it is often described in terms like "implies," "effects," "entails," "causes," "inflicts," "bring about," etc.

(ii) *Material causality*

> refers to immanent properties in the entities of a given level (which may themselves be composed of the entities of a lower level). Material causality is often described by concepts such as "consisting of," "made of," etc.[6]

(iii) *Formal causality*

> refers to the form of a given entity or process insofar as it is not reducible to effective or material causality. It is often described by concepts like "the structure of," "organizes," etc.

(iv) *Functional causality*

> refers to the role played by a part within an integrated processual whole, or the purpose of a behaviour seen from the perspective of a system's chance of remaining stable (or `surviving') over time. Terms such as "govern," "control," "regulate," "role" and, of course, "functions" are applied here.

Our reason for using the Aristotelian framework will appear in detail below, but can be briefly stated here: there is a place for a rational concept of downward causation (in some version) in science and philosophy, but only within a broader framework of causal explanation. Very often `causality' is implicitly equated with the usual notion of efficient causality, but if downward causation is regarded as an instance of efficient causality it will form a "strong version" of the concept, which, as we shall see, is not a plausible one. The notion of causality should therefore be enlarged to make sense of downward causation. Even in everyday language the dominance of the `efficient' sense of causality is well known, as can be noticed, for instance, in the apparent oddity of such expressions as "the wood is the cause of the table," which refers to the material cause and makes perfect sense when translated to the proposition "the table is made of wood." The notion of functional causality, our reinterpretation of the Aristotelian 'final cause' (often misunderstood as implying the paradox of a future state (a goal) influencing a present state) does not play any substantial role in the present argument and is only included here to emphasize the general necessity of a set of multiple causal explanatory //p. 18/ tools in science, and to posit that this notion can be interpreted in a scientifically legitimate way as well (regardless of whether it is reinterpretation as functional causality, as above, or as intentional

causality, which is an alternative possibility).

Now, our idea is that theories of downward causation fall in three types: strong, medium, and weak downward causation, respectively. Before discussing these three types, it is necessary to point out the relation of downward causation to what may be termed *upward causation*. The assumption of downward causation most often rests on the idea of `bottom-up' processes involving upward causation, and it is possible that, in turn, upward causation always rests on a certain interpretation of the concept of emergence.[7]

The idea behind *upward causation* can be termed as follows: the emergence of a higher entity from a lower one is characterized by a certain causal process leading from the lower level entities to the higher level ones, so that the lower level can be seen as the cause and the higher level as the effect. Now, this interpretation of upward causation -- to which we shall return in our discussion of the three alternatives -- has a tendency to leave the higher level as a fairly impotent construction, seriously threatened to be but an epiphenomenon of the lower one. Thus, supplying this first assumption of upward causation with a complementary one of a downward causation type, thereby restoring the higher level's ontological prominence, seems unavoidable. The idea is that once constituted, the higher level is equipped with causal powers of its own, so that it is then able to inflict effects on the lower level having caused it. Consequently, the biological cell is able to control the single molecules of which it consists, or thought processes are able to guide their neurophysiological substrate in the brain. In this way, every time downward causation is assumed, this is a supplementary gesture to counterweight the threatening reductionism inherent in a previous upward causation. Of course, once both are at work, they work in exactly opposite directions, which is rather hard to imagine; sometimes the notion of "dialectics" is invoked to picture this process.

Now, given the fact that some kind of downward causation is a necessary supplement to every supposition of an upward causation, the downward causations tend to group in the following three categories.

Strong downward causation

The idea of strong downward causality may be briefly described as follows: *a given entity or process on a given level may causally inflict changes or effects on entities or processes on a lower level*.

//p. 19/
This idea requires that the levels in question are sharply distinguished and autonomous, and it can thus be seen as associated with the theses 1b and 2b above. In the history of science, representatives of this theory may be found in the classical vitalists of early biology, who supposed the existence of a creative or formative power outside the range of scientific description. When the vital power has done its work and created the higher level entity, this entity functions autonomously and independently of the lower level. The best examples of such theories are probably found in psychology and philosophy among the classical dualists, who assumed the existence of an immaterial soul that inhabits the body and is able to control it due to its special causal powers.

The theory of strong downward causation is thus based on a constitutive irreductionism (1b): higher level entities do not (only) consist of lower level entities; they possess a substantial existence qualitatively different from lower level entities.

The main arguments against strong downward causation theories are the following. (a) Faced with vital power or the powers of the soul, scientific description is by definition ruled out. Entities like these are simply outside the realm of science (an idea which betrays the

theological heritage in this position). (b) Even if a possible and sufficient scientific description of vital power and the soul were imagined, the theory would still be unacceptable because it entails a direct change in the laws of the lower level (or at least a change in lawful regularities at this level) effected from above. If, for instance, gravity could be influenced by secret means belonging to the soul -- and a higher level phenomenon thus could inflict a direct causal influence on the lawfulness of processes at a lower level -- then the hypothesis of *the inclusivity of levels* does not hold (that is, the idea that higher levels are based on certain complicated subsets from the lower levels and do not violate lower level laws).

The theory of strong downward causation is based on hypotheses 1b and 2b, it introduces a non-scientific, that is, irrational principle, and violates the assumption of the inclusivity of levels.By considering a cell as an emergent entity on the biological level and its physical basis, this criticism of strong downward causation may be exemplified. In describing this emergent entity, we are very often tempted to use downward causation-like concepts in the strong sense; as if, for instance, the emergence of the cell as a living substance efficiently causes changes in the molecules, making them somehow specifically `biological', i.e. substantially different from molecules of the non-living world,[8] or, alternatively, as if the cell as such //p. 20/ (efficiently) causes changes in the biochemical reactions among its constituents. But if we imagine a microscopic view of this alleged causal process, we will be unable to find any effective causality in the scenario. First, the process does not take place in time; second, the two events in question do not even possess the ability of causing each other. Of course, it is evident that the biological cell "governs" or "influences" the biochemical processes taking place in it -- but at the same time the cell remains in itself a biochemical construct. So on the biochemical level we see nothing but individual biochemical reactions causing one another. There is simply no identifiable process through which the cell ("as such," i.e., non-biochemically conceived) inflicts a cause on biochemistry. The cell consists of biochemical processes, we could say, but this is a non-temporal (mereological) relation and therefore non-causal in the efficient-causality use of the word. So even the idea of an upward efficient cause (or "strong" upward causation) from biochemistry to cell is wrong because of this; what we could say instead is that the molecules and the biochemical relations in question *constitute* the cell, that is, they are the material and formal causes of the cell. (Therefore, the cell is -- to anticipate the notion of weak DC -- like a stable (pseudo)-cyclical attractor in a biochemical phase space and thus it attracts the trajectories for the biochemical processes, but this "attraction" is a tendency in phase space, a property in the biochemical system in question and not a cause in the efficient use of the word. This attractor description is explained in the section on weak DC).

The concept of strong downward causation may be the result of too radical and substantial a divorce between levels, making it impossible for them to influence each other at all. When I kick a ball, it is the physics of my body which by means of efficient causality influences another physical object; it is not an autonomous psychical poseur incarnating itself in physics. On the other hand, the physics of my body is of a certain, very complicated kind, yielding emergent properties of a biological as well as of a psychical kind. The physics possessing these complicated psychical properties is able to kick the ball.[9]

//p. 21/
For this reason, strong downward causation seems to be in danger of re-importing ideas about causation that pose the same unsolvable problems as vitalism did in biology. Before going on to -- what we consider -- the two more plausible versions of downward causation, it might be worthwhile to attempt to understand the reason behind the recurrent and faulty assumption of strong downward causation. What in fact is the intuition behind the very idea of a (strong) upward causation?

The naive scenario leading to this idea is probably the following: "first" we gather a range of complicated organic molecules on the lower, physical level and "then" this structure "causes" a

biological organism to "be created." "When first" constituted, its metabolism functions as a stable and tightly regulated system, so that as a whole it emits "causes downwards" onto the chemical-physical level, which consequently is now governed by a downward cause. (The example of physics/biology may easily be substituted by brain/psyche or any other prominent distinction between levels).

The fault in this scenario or thought experiment is first of all the temporal, processual rhetoric structuring it -- here emphazised by quotation marks -- making it appear as if the process of the thought experiment corresponds to a real, existent chain of efficient causes, a regularity in time involving the exchange of energy. But contrary to what seems to be demonstrated by the experiment, the upward- and downward causes are not temporally distinct (the lower level does not cease to make up the higher one while this is assumed to "cause back"). Given this situation, the same phenomena are at the same time cause and effect for each other (which, it will be recalled, is Kant's definition of the organism in *Kritik der Urteilskraft* [10]). It is not easy (within the restricted framework of efficient causality) to understand how this should be possible except in cases where the interchangeable causes and effects do in fact belong to the same level -- unless one //p. 22/ tries to wrap the problem up in pseudo-explanations like "dialectics".[11] An additional fault in the experiment is that it is even interpreted as a real process. The biological system is not "first" realized physically-chemically, causing in turn the appearance of the biological system. The biological system *is constituted* by a certain constellation of the physical-chemical level. Of course, the coming into being of this constellation, arrangement, organisation, or whatever is the reason why the physical system is now also biological. But since the system remains physical *at the same time* as being biological (it is not *first* physical and *then* biological; it is our language from the thought experiment which erroneously leads us to introduce this temporality), it is mysticism to say that the physical level exerts an upward causality on the biological level. We may say that *when* we have a physical-chemical system in a certain arrangement, *then* it is also biological, but the words "when" and "then" in this context refer to a logical sequence, not a temporal one.

Even if we regard talking about physical causes as unproblematic, talking about *upward causes* in a strong sense is already problematic, as it is unclear what the ramifications are of assuming that a physical cause could have an effect which was not physical (and, moreover, with which it is simultaneous, in contrast to the conception of efficient cause as a temporal regularity). A phenomenological experience which leads us to accept this fallacious idea is probably an intuition like the following: if we inflict a small physical cause on our biological organism, then it might be amplified and cause severe biological changes, in some cases death, that is, effects on the biological level. But what fools us here is the fact that the "little physical cause" is only a releasing cause and not the whole physical cause of the development in question; this cause or complex network of causes is constituted by the physical description of the whole organism in addition to the small spectacular change, which just seems peculiarly cause-like seen from our point of view because of our possibility to inflict it or remove it (whilst we cannot remove crucial parts of our body ...).

Thus, the naive assumption of even strong upward causality must be discarded. Consequently, even more problematic, if possible, is the idea of a strong downward causation -- the idea that the biological level *as such* should be able to inflict purely physical effects. A physical effect may only be the result of a physical cause which eo ipso does not exist on the higher level, unless the higher level is interpreted as a peculiarly complex and stable organisation of entities of the lower level. But then the cause is no longer biological in any strict, level use of the word. This whole difficulty is in our opinion purely conceptual and has its background //p. 23/ in the artificial isolation of physics, and the assumption that physics as we know it in its present state is already a complete and mainly microphysical science. When this is assumed, higher organizations of physical matter may of course seem mystical and strange. But the fact that physical systems of far more complicated and self-organizing kinds exist -- organisms, for

instance -- may suggest that this complexity is in itself a property in certain arrangements of matter and as such a physical property. Earthly DNA-life is, as is well-known, deeply dependent on carbon chemistry -- without it being the case that carbon should be the cause and life its effect. The fact that carbon life is possible is rather a physical property in carbon, even if it may not be deduced from the observation of a single carbon atom.

Quite a different question from the one raised in our -- logical, not temporal -- thought experiment is the truly temporal question of how the actual origin of biology in its earthly version came about, that is, what complicated and extraordinary but still purely physical processes took place in the primordial soup leading to life. This historical and exciting question of efficient causality must not be mixed up with the purely logical question of which levels are constituted -- not caused -- by processes and entities on other levels. It simply does not make sense to say that any efficient cause should go from the lower towards the higher level or vice versa. The levels are not objects for causes, they are not real in the sense of the saying, all is real that has causal powers. In our view, the very source of the vitalist fallacy is the identification between this temporal chain and the constitutive relation.

Now, given the rejection of the vitalism (and substance dualism) inherent in strong downward causation, two more possibilities exist, according to whether we ascribe the purely fomal cause on the higher level any power to constrain lower level processes or not -- these we will coin medium downward causation and weak downward causation, respectively.

Medium downward causation

The distinctive feature of medium DC in contrast to strong DC is that it does not allow higher level phenomena to have a direct influence on lower level laws. The medium DC defends 1b (in which a higher level entity, such as a cell or a psyche, is a real substantial phenomenon in its own right), but can exist with both 2a and 2b.

As an example of this position, the neuropsychologist Roger W. Sperry (1980, 1986) and his "emergent interactionism" may be mentioned. For most of his professional life Sperry was reductionist in a way we do not discuss here -- everything could be reduced to physics, also the psychic. It was rather late that he //p. 24/ started defending so-called emergent interactionism -- after his investigation of split-brain patients and the duality of the brain.

One of the central examples given by Sperry (1969) is quite simple: a wheel running downhill. None of the single molecules constituting the wheel or gravity's pull on them are sufficient to explain the rolling movement. To explain this one must recur to the higher level at which the form of the wheel becomes conceivable. On the one hand, Sperry rejects any law-changing strong downward causation as well as the idea of a substantial dualism (or pluralism, for that matter). On the other, he maintains -- as is evident from the concept of "interactionism" -- that the higher level performs a function irreducible to the lower. If we take a neuropsychological example and try to explain the origin of some state of consciousness from the state of the nervous system, then the only consistent interpretation of Sperry's point of view is that a given state of consciousness is *chosen* among a series of states of consciousness made possible by the nervous system at a certain moment. The decisive point in this *choice* of state of consciousness lies on the higher level, in this case the psychical level. It is the previous states of consciousness which determine or select which one of the possible states of consciousness should be realized. Hence the interactionism; the interaction between the manifested states of consciousness decides which possibility is to be realized.

Medium downward causation can be defined as follows: an entity on a higher level comes into being through a realization of one amongst several possible states on the lower level -- with the previous states of the higher level as the factor of selection. This idea can be made more precise

with the aid of an interpretation of the concept of "boundary condition."

This concept is primarily used in physics and mathematics. Mathematically, the boundary condition is the set of selection criteria by which one can choose one among several solutions to a set of differential equations describing the dynamics of a system.[12] In classical mechanics, a system's initial conditions are defined as the set of parameters describing the starting point of a system at a certain moment and which -- measured with sufficient precision -- may form the basis for the calculation of an, in principle, unlimited predictionof the system's behaviour. In complex physical phenomena it is supposed that certain changes in initial conditions make central properties in the dynamics change; these are named boundary conditions because they delimit the set of initial conditions within which the properties in questions will be found. In this context the concept does not entail the assumption of levels.

//p. 25/
In relation to level theories, boundary conditions are conceived as the conditions which select and delimit various types of the system's several possible developments. The realization of the system implies that one of these typical developments is selected, and the set of initial conditions yielding the type of possibility chosen are thus a certain type of boundary condition which has been called *constraining conditions*. They only exist in complex multi-level phenomena on a level higher than the focal level, and are the conditions by which entities on a high level constrain the activity on the lower focal level.[13]

On this basis, *medium downward causation* can be reformulated as follows: *higher level entities are constraining conditions for the emergent activity of lower levels*. And -- hence the Sperry example -- in a process, the already realized higher level states are constraining conditions for the coming states.

How are we to understand the nature of this constraint? One interpretation is to say that the higher level is characterised by *organizational principles* -- lawlike regularities -- that have an effect ("downward," as it were) on the distribution of lower level events and substances. Thus, if, for instance, evolution by natural selection is such a lawlike regularity, we can only understand the physical distribution of energy and matter in a ecosystem if we consider the effect of natural selection on frequencies of genotypes, and thus on the phenotypes of the various existing organisms, which themselves influence the cycles of matter and energy in the system. This interpretation of medium DC is close to the view of Campbell (1974).

In contrast to weak downward causation, medium downward causation is characterised by this claim; even if no law-breaking influence top down is admitted, the higher level *constrains* which higher level phenomenon will result from a given lower level state. Thus, the radical forms of dualism or vitalism of strong downward causation is avoided at the expense of a less radical idea that the same lower level constituents may correspond to a series of different higher level phenomena.

In contrast to strong DC, medium DC does not involve the idea of a strict "efficient" temporal causality from an independent higher level to a lower one, rather, the entities at various levels may enter part-whole relations (e.g., mental phenomena control their component neural and biophysical sub-elements), in which the control of the part by the whole can be seen as a kind of functional (teleological) causation, which is based on efficient, material as well as formal causation in a multinested system of constraints. The kind of determinative relation between part and whole is not quite clear, and the term "interaction" is according //p. 26/ to Sperry (1987) not the best for the kind of relationship envisaged.[14] Thus, "Mind is conceived to move matter in the brain and to govern, rule, and direct neural and chemical events without interacting with the components at the component level, just as an organism may move and govern the time-space course of its atoms and tissues without interacting with them" (Sperry

1987).

We have to differentiate between the following two assumptions. (a) Higher level entities function as criteria for the selection of lower level emergent processes. The higher level entities constrain the development of lower level processes in accordance with the history of the level. (b) One set of entities at a lower level can be the starting point for different entities at the higher level. This is a sort of inverse supervenience. One can, for the sake of the argument, assume that two organisms consist of the same amount of different substance -- but are very different organisms. This conclusion rests on the premise that the levels already exist -- they cannot be used to describe or explain the development of levels.

Weak downward causation

The theory of weak downward causation admits neither of the two claims just mentioned and instead interprets the concept in the light of theses 1a and 2a. Thus, the higher level is conceived as an organizational level, characterized by the organization, the whole, the pattern, the structure, in short *the form* into which the constituents are arranged. Even if it subscribes to the constitutive reductionism of 1a, it must not be identified with physical or mechanical reductionism; the forms of the higher level are supposed to be non-reducible. In contrast to medium downward causation it is characterized by not admitting the special interpretation of boundary conditions as constraining conditions, and hence it does not allow the possibility that several higher level phenomena correspond to one and the same lower level phenomenon.[15] One possible way of describing weak downward causation is by using the phase-space terminology invented by qualitative dynamics.[16] Phase space maps all the possible states of a system into a space defined by a //p. 27/ set of dimensions, each of them corresponding to a parameter of the system. Through a continuous change in these parameters, any change in the system will be modelled by a trajectory in the phase space. Classical conservative mechanical systems will result in one distinct trajectory through the phase space, but various dampened, thermodynamic systems lose energy all along and may approach the same behaviour as systems with other initial conditions. An *attractor* is the name of a set of points in the phase space in which trajectories with many different initial conditions end. Attractors may vary in kind from points (corresponding to no change in the system), to orbits (corresponding to cyclically recurrent states), to pseudo-cycles (corresponding to overall but not precisely recurrent behaviour), and the strange attractors of chaos theory (with unpredictable behaviour due to exponentially divergent trajectories from nearby points -- but still with pattern properties). Attractors are of course not unique to emergent behaviour (unless all thermodynamic micro-macro distinctions involve emergent behaviour), but it seems to be the case that emergent higher levels are regulated by stable and complicated attractors for the dynamics of the lower level, often characterized by cyclical mechanisms of regulation.

Hence, in the biological case, organisms can analogically be regarded as consisting of highly complicated attractors for the behaviour of organic molecules in a biochemical space -- an attractor with stable part-cycles (metabolism, reproduction, etc.). Given the relevant organic molecules, these attractors exist in a certain (Platonic) sense before the particular living organism. As argued in detail by the theoretical biologists Kauffman and Goodwin, the fact that a biological species consists of stable organisms is neither a wonder nor solely a product of selection, as traditionally held by neo-Darwinism.[17] The stability is the result of internal, formal properties in the organisation of the organism, and the job of natural selection is only to sort the possible stable organisms and find those most fit for the given milieu; in this sense, the genes selected by natural selection set the parameters that specify the initial conditions for emergent development.

As a formal tool the attractor description may be applied to various cases of biological

complexity. A relatively simple example is provided by our body, which consists of about 250 different *cell types*, each cell (except red blood cells) has approximately 75,000 different genes. Nearly all contain the same genes, but they differ in their type-specific set of genes that is `turned on' or `off.' Thus a liver cell may be modelled as a stable point in the whole state space of the human cells characterized by the configuration of *n* active and 75,000-*n* inactive genes (*n* may typically be in the range of 10,000 - 15,000). Though an eye cell and a liver cell might have the same *n*, the eye cell has genes (e.g. for pigment) turned on that are inactive in the liver, etc. Now this description allows us to see developmental cell //p. 28/ differentiation as the establishment of a historical tree of trajectories (representing cell divisions and cell diversification during embryogenesis) moving towards about 250 different basins of attraction, where each attractor represents a stable cell type in the adult organism, and where a point on that attractor represents, for example, a possible configuration of the active and inactive genes in which a liver cell can be. If it is a point attractor, there is only one such configuration, but cells usually have many genetic states and may thus cycle along a cyclic attractor (see Kauffman (1993) for a detailed model).

An attractor conception of the higher level as characterized by formal causes of the self-organization of constituents on a lower level does not yet constitute a detailed explanatory theory, but merely gives a framework for description. Yet it makes clearer some crucial implications of the idea of levels. The fact that the attractor resulting in the higher level entities is highly complicated and presupposes a specific historical process of generation means that it is localised in a strongly delimited part of the biochemical phase space. The boundary conditions for its stability (and origin) are many and specific (even if they may occur generically in certain realms of the phase space). Thus it is rather small and delimited areas in the phase space which constitute the basin of attraction of this attractor; and the feeling that life is "rare" in relation to other parts of physics confers the idea that systems must pass through a very narrow bottleneck for higher levels to show up (Gell-Mann 1994). On the other hand there are two local properties which might counterweight the overall rarity of this attractor: first it can be imagined that areas exist (still relatively rare, globally conceived) where the occurrence of stable attractors is generic, that is, where a movement of the trajectory outside the basin of attraction does not lead to a non-species (or, a non-higher level entity), but instead to *another* species (or higher level entity). Second, any attractor is the centre of a basin of attraction, that is, even if it is very rare on the global level, it is locally generic; stable and insensitive to perturbations. This insensibility might be the reason for the assumption of downward causation: even if you give an organism a disease, a small push, a little distress, in a large number of cases it will be able to push off the change, that is, stay inside the basin of attraction and hence return to the stable attractor after having rejected the perturbation for some time. The relative stability of the attractor is, then, identical to the "governing" of the behavior of the entity, which is so easy to interpret as a case of downward causation: the physical perturbation is regulated by the biological attractor. The attractor functions as a "whole" at a higher level than the processes that constitute it -- hence the frequent talk of holism in connection with the discussions of level and emergence.[18]

//p. 29/
This gestalt holism is natural, insofar as the bundle of processes described by the attractor constitutes a pattern which necessarily comprises a plurality of entities from the lower level -- but the attractor model is deprived of any vitalism whatsoever. The attractor does not "take care" to sustain itself in any spiritual use of the word, and any attractor of the kind described possesses its own limit of tolerance for perturbations. Beyond this limit, the dynamics enters another basin of attraction and the organism ceases to exist.[19]

The attractor also functions, however, as a whole in another sense of the word. Because the attractor is the drain of a basin of different initial conditions (which is small globally but very large locally in comparison with the attractor itself), the attractor automatically becomes a sign

for all the initial conditions attracted by it: the attractor thus subsumes these states. This implies that the attractor is a general *type*, of which the single phase-space points in its basin will be *tokens*. Put in another way: the entity on the higher level may be instantiated by a long range of (tightly related) constellations of entities from the lower level. The entity on the higher level is thus necessarily *more general* than the particular constellation of entities from the lower one. This general "governing" of particular lower level constellations is phenomenologically very striking and one more reason for the widespread interpretations of it as teleological and too strongly downwardly causal -- seen from a weak downward causation point of view. A tiger is a tiger, even if it may be so in a very wide range of physical ways and may appear in a lot of different states; a thought is a thought even if it may be bio-physically instantiated differently in different brains.

Consequently, the organizing pattern which the weak downward causation point of view may interpret in the attractor language of qualitative dynamics has the character of potentiality. In a certain sense,[20] it existed before it was realized, just as water -- given the properties of oxygen and hydrogen -- existed as a potentiality, as an attractor, before the single empirical water molecule was realized. It is possible to try to use Occam's razor against potentialities of this type, arguing that the phase space in which these attractors dwell is so enormous that we have no practical or even theoretical possibility to map it. It must contain attractors (in the biochemical case) not only for any single species, extinct or alive, but also for any possible species in any possible world (with our biochemistry). What is lost by an Occam cut here, however, is *the conception of higher level entities as attractors for the dynamics of lower levels*. In that case, the stronger down-//p. 30/ward causation versions threaten to replace these potentialities with actualities, with certain yet unknown causal powers, and this very easily borders on seeing the formation of higher level entities as a mystical and para-religious act of creation.[21] Even `weaker' points of view than weak downward causation, on the other hand, will try to explain away higher level entities, either as mere contingent facts of specific phenomena, or as targets for physicalist reductionist programs not yet, if ever, possible. Nobody, of course, will shave so close with Occam's razor that also rather well-described cases of higher level entities as attractors for lower level dynamics disappear -- as, for instance, the water molecule. But where should the limit be fixed, how close must the shave be? It would in this respect be appropriate to recall that already at the pre-biological level of organic macromolecules, the amount of possible chemical compounds is completely unknown and in principle probably ungraspable. That is, even if we restrict the content of our biochemical phase space to only organic macromolecules, we have no precise idea of the geography of the attractors of this phase space. But this fact does not logically force us to accept the notion of strong downward causation in this case, or to believe in claims that these molecules are just "created" in some mystical process.

On the other hand, accepting the pattern/attractor model of weak downward causation invariably entails certain important consequences. First off, we will find these phenomena in a much larger variety of cases than in the few famed emergent transitions between large levels. *Patterns* are already crucial in macrophysics, solid state physics, hydrodynamics, cosmology and many other branches of physics, just as we will find attractors in any dampened system, that is, any system in which thermodynamics plays some role. A crucial question is, then: if patterns and attractors are necessary descriptive tools for a *level theory* of weak downward causation, what is the *sufficient* condition for the constitution of a level? Our guess is that the answer provided by consistent weak downward causation must be that within the theoretical framework of qualitative dynamics there is no sharp boundary distinguishing the large cases of emergence (physics/biology, etc.) from, for instance, micro/macrophysical emergences, or from intra-biological emergences (singe cell/multicellular). There is a continuum from simple patterns to complicated patterns involving a large range of different, specific interactions between parts. But this implies that levels are not necessarily sharply delimited; however, it does not imply that levels do not possess any ontological status. Even if levels have a gestalt

quality, they are not mere epistemological constructs, since the gestalts governing them are objective. In the weak downward causation opinion, then, //p. 31/ emergence and downward causation must be formal ontological concepts in a complex mereology[22].

We should stress that claiming the existence of weak downward causation is different from claiming that emergent properties are just epiphenomenal. In contrast to Jaegwon Kim, who holds that "if emergent properties have no downward causal powers, they can have no causal powers at all,"[23] we have argued that the attractor description could count as a good case for the existence of an emergent property (i.e. belonging to an attractor in state space)[24] without the very notion of "causal power" having to be relevant at all. To us it seems that the ontology of abstract objects -- forms, shapes, mathematical and topological relations -- indicates that entities do not have to have causal powers in order to exist.

To sum up the position of weak downward causation: downward causation cannot be interpreted as any kind of efficient causation. Downward causation must be interpreted as a case of formal causation, an organizing principle.[25]

Form, causality, supervenience

It may be difficult to decide whether the medium or the weak version of downward causation is the most promising, and even if the strong version can easily be discarded, we, the authors of this article, do not hide the fact that we disagree with the choice between the latter two. The point of departure in both cases is the assumption of formal causality. As higher level entities (e.g. a cell) supervene on lower order entities (molecules), formal causality on the higher level supervenes on the efficient causality of the lower level. This can be interpreted as the selection -- from a very large set of possible (efficient) interactions -- of a small set of realizable (efficient) interactions on the lower level, on which the higher level then //p. 32/ (formally) supervenes. In any case, in our view this is the only non-contradictory version of downward causation possible.

As regards weak downward causation, we interpreted the formal cause in terms of the attractor concept. This description of phase space evidently gives downward causation a legitimate scientific status without involving any untenable metaphysical idea of a temporal causal process from a higher level to a lower one (or vice versa, for that matter). What medium DC will add to this conception is the claim that *the very process of emergence will necessarily change the local appearance of phase space*, so that the higher level attractors are only created in the process of emergence -- leading to the distinction between boundary conditions and constraining conditions. This idea leads medium downward causation to pay special attention to "first time emergences" where the creation of these allegedly new attractors takes place;[26] while weak downward causation with its more structural approach does not consider these attractors any product of creation or contingent historical processes, and, consequently, only regards the question of the first-time generation of an emergent entity as a problem of how to manoeuvre a system into the right corner of a pre-existing phase space where the interesting attractors lie.[27]

A precondition for a strict mathematical application of the dynamical phase-space description of any material system is that the very phase space can be precisely defined and its dynamical equations known before the study of specific trajectories that gives the system (i.e. its description) the characteristic forms of movement. In other words: the phase space must, as it were, be an abstract pre-formed structure with fixed parameters and boundary conditions within which one may follow the dynamical development of the trajectories (as revealed in a computer simulation) of the individual states of the system (represented by points in the phase space), which may be attracted to, for instance, chaotic attractors. Thus, if we want to apply such a

description to the emergence of life, we face the problem that this very emergence will change not only the states of "the system" (which is the big `primordial soup'), but the very parameters that are important for the description of the system. The emergence of life seems to change the very phase space, so that it can no longer be considered a fixed structure; rather, new higher order attractors will appear. To say so, however, is to apply dynamical //p. 33/ descriptive language (qualitative dynamics, non-linear systems) in a somewhat metaphorical sense.

Notes

[in the printed version distributed in the foot of the main text].

[1] In his 1923 treatise, Lloyd Morgan used the term supervenience (which later acquired a different meaning) for the return action of emergents upon the lower level events from which they arose. Campbell (1974) uses the term explicitly, as does Sperry (see below). For comments, see Blitz (1992).

[2] Emmeche, Køppe & Stjernfelt (1997). Our argument in the present paper is a separate one, but the notions of local ontology and inclusiveness are discussed in greater detail in the previous paper .

[3] Baas (1994); Baas & Emmeche (1997)

[4] Kim (1995). Of course, there are problems with this interpretation too. See also Berckermann, Flohr & Kim (1992).

[5] If truly different local ontologies exist in the universe, inclusiveness and locality imply that the historically contingent form of a given, say, biological, level may influence the further emergence of higher levels. Thus, if self-consciousness and institutions are characteristic for the primary entities at the psychological and sociological level locally known on Earth, we cannot be sure that these characteristics are universally found. Though this is speculative, the general theory of levels conforms to the principles of contemporary science.

[6] Depending on the specific entity and frame of description, the properties may include potential energy, or specific energy states of particles, field strength, etc.

[7] cf. Küppers (1992).

[8] As if `organic chemistry' needed other laws of chemistry than inorganic chemistry, as once believed. (By the early 1840s, organic chemistry had become a chemistry of carbon compounds, not a chemistry of living systems as such). Of course, proteins in living cells are indeed characterized by their bio-functional specificity, but this is exactly related to the their functional role in the metabolism of the cell and the property of being (partly) specified by the sequence information in the DNA and produced by the cell's complex molecular machinery. However, this feature of being produced this way and having such and such a functional role does not change the chemical properties of a given molecule. A biological protein such as a cytochrome C enzyme is not qualitatively different from other non-natural (artificially produced) proteins in its substance, and differs only in chemical properties (e.g. its reaction kinetics) from such non-natural proteins for the same reason that two biological proteins (with different three-dimensional structures) differ in their chemical properties. In other words: the biological specificity of naturally occurring proteins is their ability to catalyze a particular chemical reaction that is functional for the living cell; the fact that the cell contains DNA that specify (or `code') the primary structure of these proteins (a tiny amount out of a huge space of combinatorically possible proteins of a given size) is a consequence of the whole evolutionary history of the cell. In this sence, the chemical properties of biological proteins are perfectly

chemically explainable, but their functional role in the cell and their evolutionary origin involve kinds of explanation (functional-cybernetic and evolutionary) different from pure biochemistry.

[9] Pattee (in a comment on a version of this paper) remarked that this is a good illustration of why "the universal cause" (see Pattee's paper in this volume) does not explain anything, as we seem to be "begging the question of what makes a good model of measurement and control processes (...). Why is not the psychical state of John's brain that ardently desires a goal as efficient a cause (or useful or acceptable an explanation) as the physics of collisions?" Pattee suggest that we should "try telling a soccer player who had just made the winning kick that it was not he, an excited `psychical poseur', who caused the goal, but only the most probable sum-over-histories of collisions of innumerable quarks and gluons. It is just because the latter is an *in principle* universal cause-less model for all conceivable events that it is not in practice useful, appropriate, or explanatory for events where there is a simple causal model." We agree with Pattee's emphasis on the pragmatics of explanation in any given situation, as well as with his realism concerning mental causation (so the desire of the player to score is in this sense an efficient cause). We do not claim to have answered the ontological question about the causal relations between the different modes of existence of various parts of the physical universe. The intention of a person may be enough as explanation in some contexts, but when physical, biological, psychological and social science enters the scene, each with different kinds of explanations of the same act, the ontological question remains. When Pattee (this volume) writes that "concepts of causation are subjective *in so far as* they cannot be separated from the observer's choice of observables and the choice of measuring devises" (our emphasis) he is quite right, but causation is objective *in so far as* the causal powers of nature exist even though our conceptual representations of these powers are observer-relative.

[10] Kant, 1790 [1951], SS66, second part: "This principle, which is at the same time a definition, is as follows: *An organized product of nature is one in which every part is reciprocally purpose [end] and means*. In it nothing is vain, without purpose, or to be ascribed to a blind mechanism of nature." One can easily interpret this as an instance of our modified Aristotelian concept of functional causality.

[11] Within a broader framework, however, one could interpret "being cause and effect for each other" as an instance of functional causation, and then give some standard account of 1) the occurrence of natural functional relations in organisms by referring to the operation of evolutionary adaptation by Darwinian natural selection, or 2) the occurrence of artificially created functional relations in tools and machines by referring to intentional construction by human beings.

[12] E.g., if the solution contains *r* arbitrary constants, these constants may be eliminated to give a unique solution to a problem if there are *r* given *conditions* that the solution must satisfy. *Boundary* conditions, which may be for the function and/or its derivatives at certain boundary points, may be used to obtain a solution which is valid for the region specified by the conditions.

[13] In Salthe's (1985) triadic structure, an entity must be considered at its own (focal) level, the level above and the level below. Lower level constraints act on the emergent process as possibilities or *initiating conditions*, higher level constraints or *constraining conditions* may act as the boundary conditions of the environment or other such constraints. Salthe has since (1993) modified his model.

[14] "mental phenomena are described as primarily supervening rather than intervening, in the physiological process" (Sperry 1987). For further discussion of Sperry's interactionism, see Kim's paper, this volume.

[15] Of course, it is easy to imagine, for instance, that a brain cell in a certain state may partake in several different thought or emotion processes. But advocates of weak downward causation will argue that the range of observation is mistakenly chosen here: constitutive reductionism requires the lower level process to be envisaged in the right proportions, that is, the same proportions as the higher level process (the thought) in question. On the level of whole connected series of thoughts, then, it cannot be the case that identical lower level (neurological) processes give rise to different higher level (psychological) processes. It is the stable forms of the higher levels which determine the scope of observation.

[16] A phase space is a multiple space in which the coordinates represent the variable required to specify the states of the system (e.g. a six-dimensional space incorporating three dimensions of positions and three of the momentum of a single particle, or a 6*N* state space in which a single point characterizes the positions and momentums of a gas of *N* particles). For a popular introduction to this concept and qualitative dynamics, see Gleick (1987).

[17] Kauffman (1993); Goodwin (1994).

[18] One should remember that the attractor is a set of points in a phase space, and a single point represents a complete (micro) state description of the entire system.

[19] This presupposes that the organism can be specified as a set of states [within such an attractor-or within a cluster of "healthy" as well as "diseased" attractors] amenable to a (physical) dynamical systems description. For some systems that are `self-modifying' this presupposition may not be met (Kampis 1991).

[20] At least in the mathematical sense of some real forms (a mathematical description of an attractor) that may or may not correspond to actual `empirical,' existent things. According to modern cosmology, no molecules and very few kinds of elements existed in the very early universe, except (according to a potentialistic interpretation of natural laws) as a mere possibility of the organization of elementary particles.

[21] It is interesting to note that weak DC may equally be regarded as an idealistic, religious aberrance in the perspective of stronger DC versions. Weak DC's insistence on potentiality can be considered an insistence that everything is determined beforehand, the landscape of attractors is always already there, as if arranged during Genesis by a deistic God. On the other hand, the theology of the stronger version is not deistic, but rather theistic or even mystical: in any case of emergence, a creative force is at work, as if God had not finished by the sixth day.

[22] One could argue that emergence and DC must be formal ontological concepts. If, on the one hand, they are mere epistemological concepts in a subjectivist sense of the word, they would cease to possess scientific interest and just become consequences of our present lack of knowledge. On the other, if they are material ontological concepts, they would pertain to specific material cases and there would be no conceptual interest in comparing the emergence going on between physics and biology with the one taking place between brain and psyche. The very concept of emergence (and hence of DC) presupposes that these processes share some crucial formal properties-which then in the weak DC interpretation are the gestalt structural organization of entities from the lower level.

[23] Kim, this volume

[24] It could be any kind of attractor: point, cyclic, quasi-cyclic or strange. This is a central point where our interpretation of emergence and weak downward causation differs from that of Newman (1996), who requires the attractor to be strange (chaotic) in order to meet the requirement of unpredictability for an emergent phenomenon. We would argue that even though complex systems like living organisms can be modelled by non-linear dynamics as

being deterministic chaotic (and thus non-predictable) systems, they are non-predictable at a deeper level (than represented by non-predictable chaos) due to their historicity, i.e., their genesis is dependent on mutations (that ultimately may be quantum-indeterminate) and complex contingent environmental fluctuations.

[25] A prominent forerunner of such a position is Ernst Cassirer, who attacked the latter-day vitalists of his age and instead invoked a "nicht-stoffliche Ordnung" as crucial in biology.

[26] That is, the very space of possible states (as characterized by a given set of parameters) undergoes change, and often new parameters or observables become relevant to the system description. E.g., describing the "state space" of living cells after the emergence of life, a description of the system as a $6N$-dimensional state space of particles with $3N$ position coordinates and $3N$ momentum coordinates, will not be feasible. Rather, a genetic sequence space of possible genetic states (characterized by DNA sequences) may in a lot of cases be a more interesting model, compare Küppers (1992) and Eigen (1992).

[27] In this respect, the dissension between the medium and weak DC positions may be seen in the historical context of the preformation versus epigenesis debate, as discussed by Moreno & Umerez's paper in this volume.

References

Baas, N.A. 1994. "Emergence, Hierarchies, and Hyperstructures," pp. 515-537, in C. G. Langton, ed. *A life III, Santa Fe Studies in the Sciences of Complexity, Proc. Volume XVII*, Redwood City, CA: Addison-Wesley.

Baas, N.A. & C. Emmeche. 1997. "Emergence and explanation," *SFI Working Paper* 97-02-008. Santa Fe Institute, New Mexico. (submittet to *Intellectica*).

Beckermann, A., H. Flohr and J. Kim, eds. 1992. *Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism*. Berlin & New York: Walter de Gruyter.

Blitz, D. 1992. *Emergent Evolution. Qualitative Novelty and the Levels of Reality*. Dordrecht: Kluwer Academic Publishers.

Campbell, D.T. 1974. "`Downward causation' in hierarchically organised biological systems," pp.179-186, in F. Ayala & T. Dobzhansky, eds. *Studies in the Philosophy of Biology*. Berkeley: University of California Press.

Cassirer, E. 1991. *Das Erkenntnisproblem in der Philosophie und Wissenschaft der neueren Zeit*. Vol. 4. Hildesheim/Zürich/New York: Georg Olms Verlag.

Eigen, M. (with R. Winkler-Oswatitsch).1992. *Steps towards Life. A perspective on Evolution*. Oxford: Oxford University Press.

Emmeche, C., Køppe, S. & Stjernfelt, F. 1997. "Explaining emergence-towards an ontology of levels," *Journal for General Philosophy of Science,* 28: 83-119.

Gell-Man, M. 1994. *The Quark and the Jaguar*. New York: W.H. Freeman.

Gleick, J. 1987. *Chaos. Making a new science*. New York: Viking Penguin.

Goodwin, B. 1994. *How the leopard changed its spots. The evolution of complexity*. New York: Charles Scribner's Sons.

Kampis, G. 1991. *Self-modifying Systems in Biology and Cognitive Science*. New York: Pergamon Press.

Kant, I.1790. *Kritik der Urteilskraft* (*Critique of Judgement*, trans. J. H. Bernard. New York: Hafner Publishing Company, 1951).

Kauffman, S. 1993. *The Origins of Order. Self-organization and selection in evolution*. Oxford: Oxford University Press.

Kim, J. 1993. *Supervenience and Mind*. Cambridge: Cambridge University Press.

Kim, J. 1995. "Emergent properties," p. 224, in T. Honderich, ed. *The Oxford Companion to Philosophy*. Oxford: Oxford University Press.

Küppers, B.-O. 1992. "Understanding complexity," pp. 241-256, in Beckermann, Flohr & Kim, eds. 1992.

Køppe, Simo. 1990. *Virkelighedens Niveauer*. Copenhagen: Gyldendal [treatise in Danish on `The Levels of Reality'].

Morgan, C. Lloyd. 1923. *Emergent Evolution*. London: Williams and Norgate.

Newman, D.V. 1996. "Emergence and strange attractors," *Philosophy of Science* 63: 245-261.

Salthe, S.N. 1985. *Evolving hierarchical systems*. New York: Columbia University Press.

Salthe, S.N. 1993. *Development and evolution*. Cambridge, Massachusetts: The MIT Press.

Sperry, R.W. 1969. "A modified concept of consciousness," *Psychol. Rev.* 76: 532-536.

Sperry, R.W. 1980. "Mind-brain interaction: mentalism, yes; dualism, no," *Neuroscience* 5: 195-206.

Sperry, R.W. 1986. "Macro- versus micro-determination," *Philosophy of Science* 53: 265-275.

Sperry, R.W. 1987. "Consciousness and causality," pp.164-166, in R. L. Gregory, ed. *The Oxford Companion to the Mind*. Oxford: Oxford University Press.

*