On the Reality of Emergents

CHARBEL NIÑO EL-HANI
Federal University of Bahia

Abstract

The controversy over the notion of emergence has recently re-emerged. But a rigorous debate concerning how it might be explained or defined is often lacking. Emergence is discussed here under two strict conditions: (i) emergents can be predictable from the knowledge about a system's parts, (ii) emergents can be regarded as dependent on, and determined by, the system's micro-structure. O'Connor's definition of an emergent property is taken as a starting-point for a new definition, incorporating Emmeche and colleagues' analysis of downward causation and Baas' treatment of emergence. It is not necessary to assume that this definition might provide the solution to the problem of emergence. Rather, theoretical pluralism regarding different pragmatically-workable notions of emergence is welcome. The reality of emergents is discussed here from the standpoint of Dennet's mild realism.

1. Introduction

The debate about emergence has recently re-emerged (Kim 1999, Cunningham 2001, Pihlstrom 2002). This concept has been increasingly employed in fields like Artificial Life and neuroscience, often without a rigorous debate concerning how it might be explained or defined. But we must keep the meaning of the term 'emergence' clear, inasmuch as it has carried for a long time a burdensome load of confusion about many of its aspects. I advocate here an understanding of emergence grounded on two strict conditions: (i) an emergent property[1] can be shown to be predictable from the knowledge about a system's parts, and, even so, be properly characterized as 'emergent', (ii) an emergent property should be dependent on, and determined

© *Principia*, 6(1) (2002), pp. 51–87. Published by NEL — Epistemology and Logic Research Group, Federal University of Santa Catarina (UFSC), Brazil.

by, a system's microstructure. It is worth trying to define emergence under these strict conditions, as the resulting definition will be sufficiently strong to overcome a number of criticisms of this notion. And, if either of these conditions does not hold (e.g., emergent properties are shown to be unpredictable), this will only make the concept stronger. The opposite will be true if we start regarding emergent properties as unpredictable from the knowledge of the parts.

The unpredictability of emergents is indeed a fundamental claim of most emergentist philosophers. Lloyd Morgan maintained that emergents would be unpredictable, even if we had complete knowledge of the antecedent events, component parts, and relevant laws. In his words, one cannot ever predict "the emergent expression of some new kind of relatedness among pre-existent events" (Lloyd Morgan 1923: 6). Notwithstanding its popularity among emergentists, a number of authors consider that unpredictability, as an epistemological issue, does not fit properly in the core definition of emergence, as an ontological concept (Bunge 1977a,b; Blitz 1992). We should neither define emergence in terms of explanation or prediction nor think that the explanation and/or prediction of an emergent can eliminate it. If one sticks to the unpredictability of emergents, any contender will be in a position to argue, say, that a given property qualified as an emergent one is not really so. The fact that we cannot currently predict it from the knowledge of the parts might simply result from a shortcoming of the current state of knowledge and/or our cognitive systems.[2]

A possible solution is to transform the epistemological argument about the unpredictability of emergents into an ontological argument about their indeterminacy. In a well-known thought experiment, one can imagine a Laplacian demon who would have a complete knowledge of the fundamental laws of nature and the total distribution of matter at a certain point in the evolution of the universe. One may claim, then, that if determinism is *not* true, that demon will be incapable of predicting the emergence of a given property, no matter its complete knowledge. But, as the debates about determinism are quite complex and controversial, and seem to be far from any generally accepted solution, this would be too heavy a burden for a concept of emergence to bear. An account of emergence that does not rest on such a major assumption about the nature of the universe is clearly preferable

I also argue that emergents are dependent on, and determined by, their underlying microstructure This clearly follows if we treat emergents as species of supervenients Although there are emergentists who disagree with this position (e g , Humphreys 1997), it offers a persuasive account of an idea that many emergentist philosophies emphasize a given emergent property appears when, and only when, certain appropriate basal conditions are satisfied by a system's microstructure This makes it possible to explain emergence, and, following Bunge (1977a,b, 1979), I intend to accept a philosophy that combines an acknowledgement of emergence with the thesis that emergence is explainable and predictable within bounds

2. O'Connor's account of emergence

O'Connor (1994) offers a good starting-point to understand the relations between supervenience and emergence [3] He claims that the nature of an emergent property's dependence on the lower-level properties can be grasped if we think of emergence as a species of Kim's (1993 65) 'strong supervenience' As emergent properties are characterized in relation to the properties of an object's parts, the definition of strong supervenience should be couched as follows

(SS) A-properties of objects *strongly supervene* on B-properties of their parts $=_{df}$ Necessarily, for any object *x* and A-property *a*, if *x* has *a*, then there are B-properties *b*, *c*, *d* (including relational properties) such that (i) some proper parts of *x* have (variously) *b*, *c*, *d* and (ii) necessarily, for any things collectively having *b*, *c*, *d* there is an object of which they are parts that has *a* (O'Connor 1994 96)

The concept of supervenience is usually associated with two basic ideas concerning the relations between sets of properties *dependence* and *determination* If a set of properties A strongly supervenes on a set of properties B, this will mean, first, that any A-property is *dependent on* some B-property If anything instantiates a given A-property, it or some part (or parts) of it instantiates a given *B*-property or set of *B*-properties which is necessarily sufficient for that *A*-property The supervenience relation also entails that *A*-properties are *determined by B*-properties, i e , nothing can be just like a given thing as regards its or its parts' *B*-properties (including relational properties) without also being just like it concerning its *A*-properties

These features of dependence and determination apply both to emergent and resultant properties O'Connor (1994 97) advocates an account of property emergence, fitting together the notions of *supervenience*, *non-structurality*, and *novel causal influence*, which seems to offer, at first, an appropriate distinction between emergents and resultants Only supervenience would be common to both kinds of properties Emergents would be demarcated from resultants because they are *not* structural properties and have a determinative influence over the behavior of a system's parts

O'Connor (1994 98) defines an emergent property as follows

(EP) Property *P* is an emergent property of a (mereologically-complex) object O iff
  (1) *P* supervenes on properties of the parts of O,
  (2) *P* is not had by any of the parts of O,
  (3) *P* is distinct from any structural property of O, and
  (4) *P* has direct ('downward') determinative influence on the pattern of behavior involving O's parts

Non-structurality entails that an emergent property should be (a) potentially had only by objects of some complexity, (b) not had by any of the object's parts, (c) distinct from any structural property of the object (O'Connor 1994 97) The third item involves the definition of 'structural property'

(SP) A property, *S*, is structural $=_{df}$ If and only if proper parts of particulars having *S* have some property or properties not identical with *S*, and this state of affairs is, in part at least, constitutive of the state of affairs of the particular's having *S* (O'Connor 1994 93)

O'Connor states that this notion amounts to the idea of 'configurational pattern' in Alexander's account of emergence This account, albeit influential, would fail to capture a sufficiently strong understanding of emergence (O'Connor 1994 92–4) Alexander ([1920]1979, vol 2, book III, 47) identifies emergent properties with configurational patterns O'Connor claims that, once an emergent is identified with a structural property, one may say that it is not anything 'over and above' the having of all the various microphysical properties and relations by an object's parts Alexander's view would not establish a fundamental difference between resultants and emergents

O'Connor proposes the notion of 'non-structurality' to differentiate emergent properties from configurational patterns Nonetheless, it is quite difficult to understand properly what it is meant by the idea that emergents are 'non-structural' or 'non-configurational' properties (El-Hani & Emmeche 2000, El-Hani 2000) It is hard to reconcile the notions of 'non-structurality' and 'strong supervenience' When Kim examines supervenience in the context of level theories, he states that it naturally turns into "the thesis that properties of a whole are determined by the properties and relations that characterize its parts" (Kim 1997 278) And 'determination' means that "what higher-level properties a given entity has are totally fixed by the lower-level properties and relations characterizing its parts" (Kim 1996 222) Thus, when we characterize emergents as species of supervenients, the very notion of 'emergence' seems to be at risk It is not an easy task to explain how the claim that emergents are dependent on, and determined by, the microstructure from which they emerge can be reconciled with the thesis of their irreducibility

When we put the notions of 'non-structurality' and 'supervenience' together, emergent properties show a seemingly paradoxical relation with the properties and relations of a system's parts If an emergent property supervenes on these properties and relations, we can conclude that it is *dependent on*, and *determined by*, the latter Nonetheless, if that same emergent property is also non-structural, it *cannot*, by definition, *be constituted* or *totally fixed* by the microstructure There is a critical tension between the acceptance of strong supervenience as a way of characterizing the dependence relation between an emergent and its microstructure and the idea that emergents are 'non-structural' (El-Hani & Emmeche 2000)

Maybe we should try to understand, with Alexander, emergence as a phenomenon related to particular, special-case configurational patterns observed in entities like organisms or minds, instead of appealing to the arguably obscure notion of 'non-structurality' Nevertheless, before jumping to such a conclusion, we should consider the possibility that O'Connor's account of the emergents' causal influence over a system's parts provides a way of reconciling supervenience with non-structurality O'Connor understands 'novel causal influence' as a term intended to capture

> a very strong sense in which an emergent's causal influence is irreducible to that of the micro-properties on which it supervenes it bears its influence in a direct, 'downward' fashion, in contrast to the operation of a simple structural macro-property, whose causal influence occurs *via* the activity of the micro-properties that constitute it (O'Connor 1994 97–8)

This novel causal influence (in Campbell's (1974) terms, 'downward causation' [DC]) might explain how an emergent property can be non-structural, notwithstanding its determination by the subvening properties and relations But consider, first, that the characterization of DC as the emergent's *'direct' determinative* influence on the pattern of behavior involving the object's parts, *independently of the microproperties' activity*, must be properly explained, if one wishes to understand this notion in a physicalist framework (El-Hani & Pereira 2000) O'Connor (1994 102) construes the idea of downward macrodetermination as the claim "that the emergent *structurally* determines [ ] the systems' relational structure" 'Structural determination', by its turn, is conceived as "a species of causation distinct from ordinary efficient causation through time" (O'Connor 1994 103) It is not clear what other causal mode O'Connor has in mind and, moreover, whether DC, when interpreted according to it, could be ascribed to the emergent properties themselves (El-Hani & Emmeche 2000) Another problem is that downward macrodetermination entails the failure of microdeterminism It is necessary to explain in what sense emergence can be thought of as a species of strong supervenience, while microdeterminism, which follows from mereological supervenience, does not hold As I understand it, O'Connor's account of the emergents' causal influence does not yield a satisfactory explanation of this paradox It may be worth exploring Alexander's idea that emergent properties can be identified with configurational patterns

3. Downward causation

A proper explanation of DC may give us a rational basis for understanding emergence Nonetheless, DC is a hopeless problem for a picture allowing only strict efficient causation It is quite difficult to understand the causal agency of a higher-level event, property or entity over the lower-level components as an efficient cause, since the part-whole relation is simultaneous, not sequential Besides, an interpretation of DC as an efficient cause is incompatible with a physicalist picture (Emmeche et al 2000, El-Hani & Pereira 1999, 2000, El-Hani & Emmeche 2000, El-Hani & Videira 2001) [4]

Emmeche and coworkers (2000) suggest that an Aristotelian understanding of causality (see Ross [1923]1995, Lear 1988) may help us grasp the nature of the causal influence of the whole over its parts An emergentist framework would demand a revaluation of classical causal notions, resulting in a sort of neo-Aristotelian approach Other attempts to have Aristotelian causal notions inspire biological thought are found in Salthe (1985, 1993), Riedl (1997), Van de Vijver et al (1998), El-Hani & Pereira (1999, 2000), El-Hani & Emmeche (2000), El-Hani & Videira (2001), Vinci & Robert (2001) When examining this approach, one should not forget that Aristotle is only a reference point for inspiration The idea is not to preserve Aristotle's analysis of causality or Aristotle's philosophy in its entirety, but to re-interpret it under the light of contemporary problems and frames of reference It is natural that, in this re-interpretation, new notions of causality are substituted for the original meaning of Aristotle's categories But one has still to consider that the Greek word translated as 'cause' in Aristotle's works does not mean 'cause' in the modern sense (Ross [1923]1995 75, Lear 1988 15) For Aristotle, a 'cause' is not only an antecedent event sufficient to produce an effect or the goal of a given action, but the basis or ground of something. He states that we understand something when we know why it is what it is, and the cause is what shows the 'why' of things being what they are (*Physics* II 3, 194b17–20). Or, in a more observer-oriented approach, when we describe, in Aristotelian terms, the cause of something, we are able to understand why it is classified as a given kind of entity in our classification schemes. This usage of the term 'cause' in a broader sense than that established in modern philosophy makes it equivalent to 'explanatory feature', so that Aristotle's four causes can then be said to correspond to four kinds of explanation (Mackie 1995).[5]

Emmeche and colleagues (2000: 17) reinterpret the Aristotelian causal modes as follows: (a) Efficient causality is a cause-effect relation involving an interactional exchange of energy pertaining to entities of a given level and resulting in a temporal sequence of causally interrelated states; (b) Material causality refers to the immanent properties in the entities of a given level; (c) Formal causality corresponds to the form or pattern into which the components of a given entity or process are arranged; (d) Functional causality amounts to the role played by a part within an integrated processual whole, or the purpose of a behavior as seen from the perspective of a system's chance of remaining stable over time.[6] They identify three versions of DC (strong, medium and weak), based on different interpretations of the causal modes at stake. Medium DC is the most interesting for our purposes. DC is *not* interpreted, in the medium version, as an ordinary cause-effect relation – this interpretation is characteristic of strong DC. In medium DC, the central idea is that DC should be understood as a kind of formal causality. The starting-point is the observation that higher-level entities come to be through the realization of a subset of the total number of possible arrangements of their components, and their behavior is always restricted to a particular region of the state space. When lower-level entities are composing a higher-level system, the set of possible relations among them is *constrained*, as the system causes its components to have a much more ordered distribution in space-time than they would have in its absence. This constraint on the components' relations results from their being part of the space-time form, or pattern, of the system's structures and processes (*formal causality*). Besides, as their relations are thus restrained, the components perform specific functions, contributing to the system's stability (*functional causality*). In short, the *modification* suffered by a complex system's parts is understood, in medium DC, as a *constraint implied by being part of a pattern* (El-Hani & Pereira 2000, El-Hani & Emmeche 2000).[7]

Aristotle claims (*Physics* II 7, 198a 25–7) that the efficient, formal, and final causal modes are all aspects of form. This can be taken to mean that the efficient causal relations observed at the micro-level depend, to be instantiated, on the context provided by the higher-level pattern of relations in which they are embedded. The efficient causal interactions observed in a system at *t* all take place at the lower level, but those very interactions that come to be at *t* were selected by the system's state (and, in open systems, also by the environment) at $t-1$. This selection of the relations instantiated by a system's parts can be interpreted as corresponding to an instance of formal causality.

Two assumptions are central in this interpretation of DC:

- (i) A higher-level entity (and its environment) *constrains* the development of lower-level processes throughout its temporal evolution by *selecting* the particular set of relations among the parts that will be instantiated in each time t_i, among all the possible sets that could be selected in that time;
- (ii) A *single* set of lower-level entities may be the starting-point for the *realization* of *different* higher-level entities.

Some important remarks on O'Connor's account of emergence stem from this neo-Aristotelian approach. First, we can try now to reconcile DC with supervenience. Medium DC allows us to interpret mereological dependence as a *symmetric* relation, in the sense that, while efficient interactions between the components *realize* the form or pattern of structures and processes observed at the higher-level entity, the form *constrains* the efficient interactions that will realize the following state instantiated by the system. The best way to understand the idea of mereological co-determination is as follows: the part-whole relation is symmetric *in virtue of the conjunction of two distinct dependence relations, both asymmetric: supervenience and DC*. As mereological dependence contains another (asymmetric) determinative relation besides supervenience, we can claim that macroproperties are not wholly determined by microproperties, despite the dependence and determination features embodied in the supervenience relation.

Secondly, DC, thus interpreted, cannot be consistently ascribed to emergent properties as such. The idea that it is the higher-level entity as a whole which has a formal causal influence over its parts seems both reasonable and natural. If this interpretation is accepted, it will follow that downward macrodetermination does not stem from emergent properties, rather, new properties emerge in evolution because of the constraining action of wholes over parts. This idea is quite different from the usual claim that emergent properties "bring into the world new causal powers of their own, and, in particular, that they have powers to influence and control the direction of lower-level processes from which they emerge" (Kim 1999: 6).

The constraining action of a system over its components can be interpreted in terms of *organizational principles* that have a downward effect on the dynamics, distribution and magnitude of lower-level events and entities (Emmeche et al. 2000: 25, Blitz 1992: 161–2). Natural selection can be interpreted so as to provide an example. Campbell (1974: 181) mentions the case of the jaws of a worker termite or ant. The laws of macromechanics are obviously obeyed by these structures. Nonetheless, these laws are only one of the explanatory principles required to understand the molecular features of the jaws:

> Selection at that level has optimised viability, and has thus optimised the form of parts of organisms, for the worker termite and ant and for their solitary ancestors. We need the laws of levers, *and organism-level selection* [...] to explain the particular distribution of proteins found in the jaw and *hence* the DNA templates guiding their production (Campbell 1974: 181).

The distribution of molecular components in the jaws of worker termites and ants depends upon the historical process of selection, and this selective process crucially depends on the global state of the organisms and their environment. Natural selection can be interpreted as a kind of formal cause, epitomizing the constraining conditions operating over organisms (considering, as Campbell, organism-level selection) as parts of the spatio-temporal form of a population of conspecifics and its environment (including other populations) Natural selection is a very strange kind of 'force' (Williams 1973), if it is a force at all (Walsh 2001) Forces typically change the characteristics of the objects over which they act, as in the alteration of the state of movement of a billiard-ball by a cue But, when we consider that natural selection often 'acts' (although not always) over individual organisms, it seems that natural selection is different from a force, as it does not change directly the characteristics of individuals, but the distribution of characteristics in a given population It seems inadequate to think of natural selection as a 'force', and a possible reason for this lies in the fact that this term suggests an efficient causal action Natural selection could be rather understood in terms of a relational pattern observed in a population and its environment that constrains the operations of a large number of efficient causes, involved in each event in an individual's history of life which is consequential for its survival and reproduction

Polanyi (1968) maintained that the notion of 'boundary conditions' is useful for characterizing the conditions that constrain the behavior of an entity at a given level He observed that machines are peculiar things, as they work by applying mechanical power according to the laws of physics but possess a structure shaped by humans in order to harness these laws to serve specific purposes They work under the control of two distinct principles

> The higher one is the principle of the machine's design, and this harnesses the lower one, which consists in the physical chemical processes on which the machine relies (Polanyi 1968 1308)

The higher principle amounts to "the imposing of *boundary conditions* on the laws of physics and chemistry" (Polanyi 1968 1308) A living organism works under the control of similar principles both its environment and its spatio-temporal form can be thought of as boundary conditions harnessing the physical-chemical processes through which its components perform functions contributing to the maintenance of its dynamical stability (Polanyi 1968 1310)

At any given time t, any system is constrained to one element of the total set of its possible states If we represent the successive states of the system in a state space, we will see that it is always confined to a particular region of that space According to medium DC, the system's dynamics, as shown by the trajectory of its representative point in the state space, depends on the selection, at each time t, of one among several possible states by the overall state of the system itself and its environment (for open systems) at $t-1$ In the temporal evolution of a system, we can detect critical turning points, in which a new kind of relatedness among pre-existent entities is established and a qualitative change in the mode of evolution takes place Property emergence is related to such critical turning points, where new patterns of organization (and, thus, constraints) are established in a given system A specific set of properties emerges in a given system *for the simple fact that it is that kind of system, constrained to a particular region of the state space* Or, else, emergent properties appear in the system *because its parts are organized in a restricted set of states of relatedness (forms)* It is not that emergent properties appear because the system is 'more than the sum of the parts', but simply because *it is that particular kind of sum* (Hofstadter 1980 333) Because each state of each system is a particular kind of sum of the parts, it is necessary to describe not only the parts but also the constraining conditions acting over them Any higher-level system is obviously constituted by its components and their relations, but *these relations are what they are in virtue of the selection of a particular state of relatedness among several possible ones, in accordance with the previous state of the higher-level system itself*

From this perspective, nothing more is required to explain property emergence but the fact that a given system always instantiates a particular subset of its possible states, and, thus, a number of properties which are not found in the parts themselves or in other regions of the state space, where different modes of organization are instantiated There is nothing mysterious about emergence It is a phenomenon that can be explained and, maybe, even predicted Nonetheless, the explanation or prediction of an emergent property does not eliminate it, since the fact remains that a system is constrained to a particular region of the state space and realizes a specific set of properties, qualitatively different from those realized in other regions

Contrary to O'Connor's approach, medium DC suggests that emergent properties can be identified with configurational patterns, as Alexander claims

> To adopt the ancient distinction of form and matter, the kind of existent from which the new quality emerges is the 'matter' which assumes a certain complexity of configuration and to this pattern or universal corresponds the new emergent quality But whereas up to the present we have been content to treat the quality as something which is correlated with a certain configuration of its basis, we can now, following the clue of the relation between mind and its body, identify the quality with its peculiar form of body (Alexander [1920]1979, vol 2, book III, 47)

4. Emergence and reduction

It is not necessary to regard emergent properties as 'non-structural' or 'non-configurational' to avoid a full-blown reduction of a higher-level theory, concerning entities instantiating emergent properties, to a base theory about the microstructure from which they emerge (El-Hani & Emmeche 2000) Consider, for instance, that a mental property, M_i, actually supervenes on a complex of interrelated neuronal groups, $S_{p(i)}$, and not on a monadic property, P_i [8] It is natural, then, to think that M_i's instantiation depends on how the neurons are organized and structured in $S_{p(i)}$ M_i can be regarded as an emergent property because, at a considerably high level of brain organization (which we call 'mind')[9], it is the higher-level equivalent of the state of relatedness among neuronal groups $S_{p(i)}$ [10]

> We thus become aware [ ] that a process with the distinctive quality of mind or consciousness is in the same place and time with a neural process [ ] We are forced, therefore, to go beyond the mere correlation of the mental with these neural processes and to identify them There is but one process which, being of a specific complexity, has the quality of consciousness [ ] (Alexander [1979]1920, vol 2, book III, 5–6)

We should never lose from sight, when discussing the relations between mental properties and their basal conditions, that the latter always comes as morphological arrangements in the brain To see how mental functions are embodied, to explain psychological emergent properties in terms of their neurophysiological bases, we should keep in mind that the major basis of brain function is morphology In this connection, we should talk about 'relational supervenience' and recognize that constraining conditions are quite important when we examine this kind of supervenience Consider a causal relation between two thoughts, M_1 and M_2 As to the reason why a given set of relations among neuronal groups S_{p1}, or, alternatively, an *n*-adic relation S_1, realizes a given thought M_1, a compelling explanation is that the way the relations within that particular set are constrained makes it realize M_1, rather than other possible thoughts And, concerning the train of thought from M_1 to M_2, we can plausibly claim that the state of relatedness among neuronal groups S_{p1} constrains the range of possibilities in the following instant in time, so that a particular pattern S_{p2}, among several possible ones, is selected, and, for this reason, M_2 is instantiated

Notice that I am conceding that a given thought, M_i, is identified with a state of relatedness at the micro-level, $S_{p(i)}$, which can be described, in Alexander's terms, as a 'configurational pattern' M_i is instantiated simply because $S_{p(i)}$ is that kind of neuronal pattern We can thus moderate the claim that mental properties (and arguably other higher-level properties) are distinct from their underlying physical bases, as Kim (1993 356) urges us to do [11]

It is not difficult to see how this can be important Consider Heil's (1998) criticism of O'Connor's account of emergence, because of its assumption of 'pure' supervenience Heil considers doubtful that the concept of supervenience can bear the explanatory weight it is often thought to provide The main point in his argument is that supervenience as standardly characterized (e g , in O'Connor's approach) is only a modal concept Nonetheless, what is more important is not that supervenience holds, but *why it holds* One has to answer what features of the way we understand the world could make it the case that a given set of properties supervenes on another set Heil considers several possible explanations of the supervenience relation, including M = P, i e , if M and P are properties, then M and P are one and the same property Here, I suggest a different kind of identity as concerns emergent properties, supplementing Heil's inventory of reasons for a supervenience relation to hold $M = S_p$ I intend to avoid a problem in O'Connor's account, which also has to do with the claim that emergents are 'non-structural'

> Unlike a structural property [ ], an object's having an emergent property is in no sense constituted by its constituent objects having the properties they have An emergent property, then, supervenes on properties of an object's parts, but this supervenience is 'pure' It is not a matter of constitution, in the way the supervenience of a structural property is a matter of parts constituting wholes (Heil 1998 152)

Heil considers the idea that there could be cases of 'pure' supervenience unfounded Supervenience claims must be justified, he argues, through some ontological relation By claiming that the supervenience of emergents on their basal conditions is *explained* by an identification of emergents with lower-level configurational patterns, I intend to be clear about the reasons why a supervenience relation holds in this case

But couldn't this be a capitulation to reductionism? I know that the stance I am advocating here may not be non-reductive enough for many philosophers, but I still think it maintains most of the fundamental gains of a non-reductive physicalist stance, while avoiding a significant shortcoming of this position, namely, property dualism, which seems to be incompatible with a physicalist position worthy of the name (Kim 1989, Bickle 1998) Maybe this philosophical position could be called, following Bickle, a 'new-wave reductionism', but I do not intend to fight over labels here It is better to discuss what kind of 'reduction' is involved in this account of emergence and DC and what are its consequences (also El-Hani & Emmeche 2000) An emergentist "conceives of reduction primarily as an *explanation*, something that renders the reduced phenomena intelligible by explaining why they occur under just those conditions in which they in fact occur" (Kim 1996 228 See also Kim 1999 12) Reduction has a role to play in emergentism, provided it can be seen as a tool for explaining why the higher-level, emergent phenomena occur *under just the conditions in which they in fact occur* And these conditions necessarily involve the organization and structure where such phenomena take place But to what kind of reduction should we refer in this case? For the sake of my arguments, I will take Bunge's distinction between 'strong' and 'weak' reduction as a starting-point Bunge defines full (or strong) reduction as follows

[SR] Let T_1 and T_2 be two theories or hypotheses Then, T_2 is fully reducible to T_1 if and only if T_1 entails T_2 (i e , T_2 follows logically from T_1) (Bunge 1977a R80)

When we take relational supervenience seriously, it becomes clear that the kind of reduction that follows from the identification of emergents with configurational patterns is not a strong one, but what Bunge calls a partial (or weak) reduction

[WR] Let T_1 and T_2 be two theories or hypotheses and let S be a nonempty set of assumptions not contained in either T_1 or T_2 Then, T_2 is partially reducible to T_1 if and only if T_1 jointly with S entails T_2 (i e , T_2 follows logically from the union of T_1 and S) (Bunge 1977a R80)

Emergence and reduction are not incompatible Bunge states that full reduction can only be achieved in the case of theories that do not refer to emergents, and partial reduction is the sole possibility when there are emergents among the predicates of the theory being reduced The additional assumptions involved in partial reduction concern the organizational principles (and, thus, the constraining conditions) at work in a system (Emmeche & El-Hani 2000) This sort of reduction provides us with explanations of the phenomena studied by the target theory through mechanisms described by a base theory Nonetheless, it does not allow us to eliminate higher-level descriptions or to decrease the number of independent assumptions about the world, since it requires a nonempty set of assumptions concerning organizational principles and preserves emergent properties It is a kind of reduction that, on the one hand, cannot provide some of the results a reductive physicalist is usually looking for, and, on the other, allows a non-reductive physicalist to obtain at least part of what she wants, e g, the maintenance of the special sciences as autonomous, albeit not entirely independent, fields of research

On the one hand, the approach advanced in this paper is consistent with an 'ontological reductionism', i e, the belief that there is but one 'world stuff' and that this is material This claim is equivalent to the notion of 'ontological physicalism', as typically understood in the philosophy of mind (e g, Hellman & Thompson 1975, Kim 1993, 1995, 1996, Boyd et al 1991) On the other hand, this approach is not compatible with either 'methodological reductionism', claiming that the best scientific strategy is always to attempt explanation in terms of ever more minute entities, or such an extreme theory reductionism as exemplified in the so-called 'Unity of Science' movement, committed to the belief that eventually all sciences will and should be reduced to one super-theory, typically taken to be a final, complete physical theory (e g, Oppenheim & Putnam [1958] 1991) This stance intends to combine the ontological claim that all entities in the world are made from the very same materials, and the epistemological and methodological claims that different modes of explanation and investigation should be used to account for phenomena at different levels of complexity

5. Defining an emergent property

Baas (1994, 1996) conceives emergence as an *explanatory strategy*, recognizing the function of the observer in establishing an emergent property as a requirement at any level [12] He analyzes emergence in terms of a series of abstract construction processes Consider a set S_1 of first-order structures By some observational mechanism $Obs_1(S_1)$, we can obtain or 'measure' their properties These structures can be subjected to a family of interactions, *Int*, from which a new kind of structure appears, $S_2 = R(S_1, Obs_1(S_1), Int)$, with R standing for the result of the construction process S_2 is a second-order structure, whose properties can be obtained through another observational mechanism, Obs_2, also capable of observing the first-order structures Baas defines P as an emergent property of S_2 if and only if P belongs to the set $Obs_2(S_2)$ but not to the set $Obs_2(S_1)$

Baas' definition is similar to other definitions of an emergent property (e g , Bunge 1977b 97, Blitz 1992 179), with the important difference that it highlights the observer's role According to this definition, resultant and emergent properties are different because the former are observed in a system's parts, and the latter, not I think it is also important, when defining emergent properties, to reveal what kind of relation is supposed to exist between these properties and the system's microstructure As in the case of O'Connor's definition, El-Hani & Emmeche (2000 272) and El-Hani & Videira (2001 323) propose a definition of an emergent property including more than the claim that such a property is only observed at the level of the whole

(EP) A property P is an emergent property of a (mereologically-complex) object O $=_{df}$ iff

(1) P supervenes on properties and relations of the parts of O,

(2) P is not observed in any of O's parts, i e , it belongs to the set $Obs_2(S_2)$, but not to the set $Obs_2(S_1)$,

(3) O has a downward formal/functional causal influence over its parts, constraining their relations in space-time so that the pattern of constraints realizes and, thus, explains P

Three important changes in relation to O'Connor's definition must be noted (1) this new definition dispenses with the notion of non-structurality, (2) it incorporates Baas' emphasis on the role of the observer in establishing an emergent property, (3) it modifies the account of DC, avoiding any expression that might hint at a too strong interpretation of this phenomenon and ascribing it to the object as a whole, not to the emergent properties themselves

This definition includes a similarity between resultants and emergents both are supervenients It highlights two differences first, emergents are not observed in a system's parts, while resultants are, secondly, to explain why an emergent is observed, we need to consider the pattern of constraints affecting the behavior of the system's parts, while this is not necessary for explaining resultants This definition (as O'Connor's) does not consider emergents to be unpredictable

It is not necessary to assume that this definition might provide *the* solution to the problem of emergence Rather, theoretical pluralism regarding different pragmatically-workable notions of emergence is welcome (El-Hani & Pihlstrom 2002)

6. On the reality of emergents

As emergent properties are *identified with* configurational patterns, one may naturally ask if there would be any cogent reason for preserving them in our descriptions and explanations of reality Emergentists are typically property dualists, claiming that, say, the relation between mental and neural properties is a matter of two distinct properties lawfully covarying with one another (Kim 1993 364) This is a central idea in many non-reductive physicalist approaches too Nonetheless, it is hard to give a proper explanation, from a physicalist standpoint, to the idea that the instantiation of a mental property is something more than the instantiation of its neural substrate, that mental properties are something 'over and above' their physical/biological bases In the absence of such an explanation, mental properties cannot be regarded as ontologically irreducible without breaching fundamental tenets of physicalism (Kim 1993, 1998, Bickle 1998) For a physicalist who rejects property dualism, some kind of identity theory[13], as regards, say, the mind-body problem, seems to be the most attractive option But if she also intends to be a non-reductive physicalist, she may wonder if this stance couldn't be defensible, not as a purely ontological doctrine, in terms of property dualism, but as a more epistemologically- and methodologically-oriented stance (El-Hani & Pihlstrom 2002)

I stated above that, although the instantiation of a mental property does not correspond to the instantiation of a monadic neural property, it is identical with the instantiation of a state of relatedness among neuronal groups But, then, why should we maintain mental (and other high-level) properties in our pictures of the world, if they can be ultimately regarded as identical with lower-level states of relatedness? I will deal with this question by discussing, first, the fundamental issue of what it is to be 'real' and, then, arguing for the 'reality' of emergents

Kim (1992 134, 1993 348, 1996 128–30) claims that *to be real is to have causal powers* If mental properties are real, they should have causal powers that are not merely a consequence of their physical bases (Kim 1996 230) In this perspective, it will be hard to maintain the reality of emergents, if they are identified with configurational patterns (cf Kim 1999 16)

But let us consider a different account of what it is to be real, based on Dennett's arguments about the reality of patterns [14] In this connection, it is important to overcome the simple opposition between realism and anti-realism, as well as the strong realist stance usually found in the emergence debate (cf El-Hani & Pihlstrom 2002)

Dennett (1991a 27) asks if there are really beliefs, or do neuroscience and psychology show that, strictly speaking, beliefs are figments of our imagination? Such ontological questions are generally regarded as admitting just two answers either beliefs exist or they do not We should be either realists or eliminative materialists Dennett challenges this dichotomy by exploring a feature shared by beliefs and mathematical objects both are abstract objects The reality of abstract objects can be discussed along 'metaphysical' or 'scientific' avenues The former concerns the reality of abstract objects in general, while the latter considers their scientific utility Dennett (1991a 30) chooses the scientific path, considering that what is generally at stake is not the ultimate metaphysical status of concrete or abstract objects, but whether beliefs and other mental states are *as real as* electrons or centers of gravity He claims that centers of gravity are real because they are *good* abstract objects, as they are scientifically useful (Dennett 1991a 28–9) Thus, he avoids the dichotomy between realism and eliminativism, proposing a milder realism as an attractive position regarding the status of beliefs I will argue for the reality of emergents from the standpoint of this mild realism [15]

Considering the success of the predictions of other people's behavior usually found in 'folk psychology', Dennett (1991a 29) claims that it depends, as in the case of any prediction, on the exploration of some order or pattern in the world But what is the pattern a pattern *of*? Are the patterns only in our minds? Or are they real entities? To analyze the reality of patterns, he considers, first, some objects generated by him ('frames' Figure 1) Frames A–F were made by printing ten rows of ninety dots, ten black dots followed by ten white dots, etc The overall effect is the printing of five equally spaced black squares in the window In each frame, a pseudo-random 'noise' interferes with the actual printing The noise ratios are *A*, 25%, *B*, 10%, *C*, 25%, *D*, 1%, *E*, 33%, *F*, 50% Consider that each frame presents a specific pattern What is this pattern? What does it mean to say that it is 'really' there?

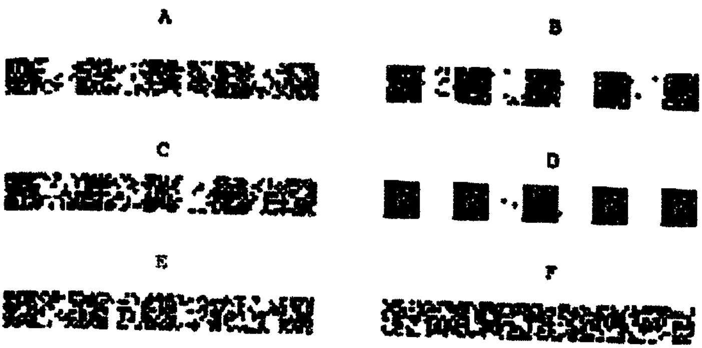

Figure 1 Dennett's frames (from Dennett 1991) [16]

Dennett (1991a 32–3) analyzes the task of transmitting information about his 'frames' How many bits of information will it take to accomplish this task? One could simply send the frame's 'bit map', identifying each dot *seriatim* This verbatim quotation is the most accurate method to transmit the information, but it is surely the least efficient But there are other possibilities, with different relations between efficiency and accuracy For instance, frame *D* could be described as "ten rows of ninety dots, ten black followed by ten white, etc , with the following exceptions dots 57, 88, etc " This expression, suitably encoded, will be much shorter than 900 bits long (the number of bits required to send the frame's bit map) The comparable expressions for transmitting the other frames will include more exceptions, and, thus, be proportionally longer

Chaitin's (1975) definition of 'mathematical randomness' allows us to grasp the idea of a real pattern A series (of dots, numbers, etc) is random if and only if the information required to describe the series is *incompressible*, i e , nothing shorter than the verbatim bit map will preserve it (Dennett 1991a 32) We can deduce, then, that a series is *not* random, showing a *real pattern*, if and only if there is a more efficient way of describing it

> A pattern exists in some [set of] data — is real — if *there is* a description of the data that is more efficient than the bit map [ ] (Dennett 1991a 34)

Conway's Game of Life offers an example midway between the world of dot frames and the world of folk psychology Dennett discerns two levels in the Game of Life, a 'physical' level, where individual cells and their patterns of change according to a set of simple rules (the 'physics' of the Life world) are described, and a higher level, where we find a series of distinct configurations

> there are the eaters, the puffer trains, and space rakes, and a host of other aptly named denizens of the Life world that emerge in the ontology of a new level This level has its own language, a transparent foreshortening of the tedious descriptions one could give at the physical level [ ] Note that there has been a distinct ontological shift as we move between levels, whereas at the physical level there is no motion, and the only individuals, cells, are defined by their fixed spatial location, at this design level we have the motion of persisting objects [ ] (Dennett 1991a 39)

Those two levels are *different descriptions of the same set of data*, albeit an ontological shift can be perceived when we move from one descriptive level to another Further, when we ascend to the higher level in the Life world, adopting its ontology, we can predict the behavior of configurations or even systems of configurations without having to bother to compute the physical level Those higher-level configurations can be regarded, from a mild realist standpoint, as *real patterns* they are more efficient than the bit map as concerns the transmission of information in the Life world

The same reasoning can be applied to folk-psychological prediction (Dennett 1991a 46–9) Consider the statement "that man staring at the lady's purse is probably thinking about stealing it" From the folk-psychological standpoint, this is not a difficult prediction Nevertheless, it is and probably will always be an intractable problem when framed at the level of the physical language Simply imagine the task of describing the scene in terms of elementary particles or, else, the thoughts involved in the prediction in terms of the physical-chemical events taking place in each neuron in the eyewitness' brain Besides, whereas the meaning of the event is readily discernible in the folk-psychological pattern, it is very likely that the same will not be true in the physical description For a mild realist, the patterns of folk psychology are quite real, as the intentional stance allows a huge compression of the information needed to describe the phenomena at stake The scale of compression in the physical language is certainly much smaller Both patterns, despite their obvious differences, can be regarded as *real*, but their reality is dependent on different contexts, and they serve different purposes

These arguments can help us understand the relation between matter and form A system in the focal level (Salthe 1985) is *matter* relative to the immediately higher-level systems and *form* relative to the immediately lower-level systems

> Any whole is a form composed of material elements, but each of these material elements considered in turn can only be described by looking at a lower level of form arrangements of smaller elements in space Thus matter and form are in this view opposing but not contradictory points of view of the same reality seen 'from above' a given phenomenon is form which is secondarily composed of material elements, seen 'from below' a given phenomenon is matter which is secondarily moulded into some form A common materialist mistake now amounts to see the first of these views as superficial or subjectivist, making questions of form impossible to grasp for science (Emmeche et al 1997 106)

When we describe a mereological system 'from above' or 'from below', we obtain *different descriptions of the same thing* Both can be regarded as *real*, given that their utility is well-defined When the system is seen 'from above', form is emphasized, while the material components recede to the background When it is seen 'from below', matter is accentuated while form is in the background In a strong realism, it is harder to maintain that both descriptions are equally real The search for the 'real nature of nature' encourages a commitment to quite strong, radically reductionist varieties of materialism, lured by the ultimate descriptive levels of reality Reductive approaches are often related to value-judgements concerning what is more 'real' or more 'fundamental' (Midgley 1995 135–7) The idea that explanations framed in the vocabulary of physics are to be preferred frequently stems from the thought that this science stands nearer to an ultimate account of matter Many physicists no longer believe that physics deals with the 'ultimate building-blocks' of matter, but lay people and also some scientists and philosophers still seem to feel, on these grounds, that physics will someday provide the only proper explanation of everything

A mild realist is in a better position to claim that we should investigate how can we combine multilevel descriptions of the world in a single explanatory picture, instead of trying to reduce all descriptions to the smallest chunks of the world This stance allows us to drift towards the perspective that different ways of talking about the phenomena are philosophically more interesting than a single reductionist description (El-Hani & Pereira 2000) This is another route to the unity of science, and, in my view, a better one At last, what is needed for science as a whole to be a unified study of nature is only that the different sciences form a continuous chain and all of them be subjected to certain general requirements regarded as necessary for scientific investigation as, for instance, advancing propositions that can be empirically tested

We can argue for the reality of beliefs (and other mental items) on the following grounds a belief is a real pattern to the extent that it provides a more efficient way of transmitting information about a person's brain state than a description of each neuron's location, set of connections and activation state (the neuroanatomical and neurophysiological 'bit map') Eliminative materialists propose that the neurosciences will offer in the future better descriptions than the mentalist or intentional stance But, as Dennett (1991a 50–1) argues, even if the intentional language were translated into the neuro biological jargon, the fact could remain that a pattern described from the intentional stance is, for a number of purposes, more efficient and useful as a description of a brain state corresponding to a given intentional category than the neuroanatomical and neurophysiological bit map Thus, we could still advocate its reality Dennett (1987 1) claims that, even though the brain is, after all, the mind, many things need saying that cannot be said in the restricted languages of neuroanatomy, neurophysiology, and behavioral psychology

Similarly, the reality of an emergent property can be advocated whenever it amounts to a more efficient description (for some purpose) of the configurational pattern with which it is identified than a micro-level description of that same configuration From this perspective, to identify emergent properties with configurational patterns does not entail that they should be eliminated from our descriptions of the world Even if such an identity is held to obtain, it can be still clear that many things need saying that cannot be said in a language that never refers to higher-level, emergent properties Certainly, this is not so strong as property dualism but a milder stance, to the effect that different theoretical objects are scientifically useful for dealing with different levels of complexity

As it identifies emergent properties with configurational patterns, this account of emergence may be interpreted in the sense that I would be giving up an emergent property *E* as a 'genuine' property and only recognizing the expression '*E*' or the concept *E* It could amount, thus, to a kind of eliminative reduction, recommending "the elimination of *E* as a property, retaining only the concept *E* (*which may play a practically indispensable role in our discourse, both ordinary and scientific*)" (Kim 1999 17 Emphasis added) But can we easily distinguish 'genuine' properties from 'mere' concepts? To advocate such an easy distinction may be tantamount to assuming a pre-Kantian or pre-Wittgensteinian metaphysical stance (Pihlstrom 2002, El-Hani & Pihlstrom 2002) 'Pre-Kantian' in the sense that, briefly, Kant's account of the nature and limits of human understanding presents a watershed for metaphysics that cannot be ignored, conceiving 'things-in-themselves' as entities in a reality which must be assumed to underlie and in some way be responsible for experience, though remaining unknowable Kant's metaphysics can be thought of as a sort of 'metaphysics of experience', a substitute for traditional metaphysics which makes it difficult to support a convincing distinction between something that is a 'genuine' property and something that is 'merely' a concept 'Pre-Wittgensteinian', very briefly, because it conceives properties which are not, at the same time, concepts, or, in other words, which are not always conceptualized Mainly for the later Wittgenstein, there is nothing that escapes the linguistic domain In short, it is hard to maintain such a duality between properties and concepts in a philosophical position taking due account of the consequences of Kant's and Wittgenstein's ideas Once we assume that the experienced, cognized world is always conceptualized, insofar as it is a world of human experience, it becomes problematic to draw a sharp distinction between 'real' properties and 'mere' concepts

Dennett's mild realism suggests that a property *E* can be retained in our pictures of the world exactly if it plays an indispensable role in our discourse It is not simply the case of arguing for the reality of emergents in *metaphysical* terms, rather, one can advocate the *scientific* utility of emergence as a modeling/explanatory tool Instead of simply worrying about "what emergent properties [ ] can *do* – that is, how they are able to make their special contributions to the ongoing processes of the world" (Kim 1999 22), we can ask what emergents can do in our theories about the world

Kim (1999 25) claims that emergent properties are "supposed to represent novel additions to the ontology of the world, and this could be so only if they bring with them *genuinely new* causal powers" Nonetheless, new modes of organization of a system's parts, instantiating new properties that can be described as lower-level configurational patterns but represent important additions to our set of useful concepts, are more relevant as novel additions to the ontology of the world And, moreover, these new modes of organization bring with them, according to medium DC, genuinely new instances of formal causation

This account of emergence and DC is compatible with the idea that "we may try to salvage downward causation by giving it a *conceptual* interpretation That is, we interpret the hierarchical levels as levels of concepts and descriptions [ ], rather than levels of properties and phenomena in the world We can then speak of downward causation when a cause is described in terms of higher-level concepts, or in a higher-level language, higher in relation to the concepts in which its effect is represented [ ] The conceptual approach may not save real downward causation [ ], however, it may be a good enough way of saving *downward causal explanation*, and perhaps this is all we need or should care about" (Kim 1999 33) From a mild realist standpoint, the discrimination between levels of concepts and descriptions and levels of properties and phenomena is not as absolute as Kim suggests Anyway, to save downward causal explanation is, in my view, indeed all we need, inasmuch as the notion of 'real' DC brings with it the very issue discussed here, 'What does it mean to say that a pattern, a property, a causal influence is *real*?'

7. Concluding remarks

The account of emergence developed here is, in some respects, different from influential emergentist philosophies I argue that, even if an emergent property is shown to be predictable from the knowledge about a system's parts, it can still be characterized as 'emergent' And the same holds, I claim, for the explanation of an emergent property on the grounds of a system's microstructure It is true that the concepts of explanation and prediction figure prominently in several emergentist philosophies (cf Kim 1999 6), but, even from a milder realist standpoint, in which the distinction between epistemology and ontology is blurred, one should not overlook that epistemological procedures such as explanation and prediction cannot eliminate entities in the ontology of a given science If we explain, say, the properties of a metallic body through the principles of quantum mechanics and solid-state physics, this will not make it lose its distinctive properties *in our descriptions* There is a fundamental difference between the *explanation* of the reasons why an entity shows a given property and the *description* of that entity as possessing that property As there are different kinds of statements within a paradigm, the distinction between explanation, prediction and description still holds when one assumes that epistemology and ontology are closely intertwined Predictions, explanations and descriptions are different kinds of narratives produced by distinct rules of a science-specific game of language

I also argue that it is necessary to moderate the commitment to property dualism found in most emergentist philosophies, proposing weaker accounts of emergence [17] These accounts can be compatible with a weak epistemological reduction (Bunge 1977a), and this mode of reduction, by its turn, can lead to a new kind of unity of science, without leveling By taking Dennett's mild realism as a metaphilosophical stance, I intend to eschew the idea that only states constituted by physical particulars having physical properties, or by a number of physical particulars related by a certain physical relation, should be accepted as 'real' and figure in our laws and theories Rather, any 'real' entity amounts, from that perspective, to a suitable way of conceptualizing the world, and, although the physical domain includes all other domains of phenomena in most current theories of levels, it is not the case that physical descriptions always provide the best patterns for our attempts to understand the world This is a kind of epistemological and methodological plea for the irreducibility of higher-level theories which will not be enough for many thinkers, who advocate a stronger ontologically-based irreducibility, but it fulfills the reasons why I have been looking for a stance which avoids a number of (in my view) undesirable consequences of classical reductionism

Many terms employed in this paper have a strong realist flavor How can we make them compatible with a milder realist approach such as Dennett's? As regards this problem, I will simply say that I do not intend to avoid the tension between the kind of realism typical of the emergence debate and the milder form of realism assumed here Rather, I wish to embrace this very tension, in order to highlight the nature of the problems we have to deal with when we put into doubt the sort of realism presupposed by most emergentists and also by their critics Anyway, I would like to offer in this last section some initial ideas regarding the compatibility of my arguments with a mild realist approach (for a more detailed discussion, see El-Hani & Pihlstrom 2002)

Consider, first, that we never have in our descriptions of the world anything like the 'bit map' in Dennett's arguments We do not have access to something that could be characterized as a 'verbatim quotation' of the external reality So, we have always to deal with different kinds of description, with different relations between accuracy and efficiency, and we are never in a position to say that some description is, *for all purposes*, the most adequate one This means that any object in our descriptions, including the basic physical particles as characterized by our best physics, neurons, cells, molecules, etc should be regarded as 'real patterns', in Dennett's sense Thus, any metaphysical discourse is a discourse inside the metaphysical component of scientific paradigms or some similar structure that, in non-scientific cultures, guides the way a community of human beings investigates and understands the world

For Aristotle, matter was unknowable in itself (*Metaphysics* VII 10, 1036a8), i e , it could not be understood in the absence of any form (*Physics* III 6, 207a24–32) As our inquiry delved into matter, our understanding would never encounter anything but form (Lear 1988 47) This can be interpreted in the sense that, at each level, what we can come to know is the principle of organization (Lear 1988 27) Any knowledge about matter is, in fact, knowledge about form one level below that one which is the focus of a particular inquiry As our inquiries delve into matter, we are incessantly devising and exploring different real patterns (*sensu* Dennett)

This work is to a great extent an endeavor in the ontology of science, dealing with theoretical entities recognized by sciences as psychology and biology Surely, there is an undeniable connection between the ontology of science and more general metaphysical questions But when these questions are located *within* the ontology of science, they become dependent on the perspective assumed by a given paradigm as regards its ontology Epistemology and ontology are inextricably intertwined when we discuss theoretical entities created from the perspective of a certain scientific paradigm When one examines the ontology of science, it is not the case of dealing with a metaphysics that has to do with the features of ultimate reality, rather, one talks about metaphysics in a more restricted sense, concerning the theoretical objects of science It makes perfect sense to opt, following Dennett, for a scientific rather than a metaphysical avenue for discussing the reality of such objects This avenue is consistent with an understanding of metaphysics as the theory of *concepts* and their relations (Harré 1985) As regards scientific concepts, a modern metaphysician (*sensu* Harré) works in the metaphysical component of a disciplinary matrix, dealing with the most general kinds of scientific objects, which are part of the ontology of science (e g , life, mind, matter, consciousness, space, time, etc)

As ontological claims are not the only kind of statements in a paradigm, it is still possible to distinguish between objects, events, properties, causes, etc (as elements of an ontological discourse) and explanations, theories, hypotheses, etc (as elements of an epistemological discourse) It is simply that this distinction cannot be taken to its extremes, justifying the idea that a discourse about causes, events, properties could be understood as a discourse free from paradigmatic assumptions

In conclusion, I would like to say that I do not intend to argue that Dennett's mild realism is necessarily the best avenue for discussing the reality of emergents My contention is only that this milder realist approach may help us advance in the emergence debate, which has been traditionally characterized by stronger forms of realism. It is stimulating to see the central problems in emergentist philosophy from such a radically different point of view Surely, when we assume a milder realist approach, we continue to fight to overcome our strong realistic commitments, often embodied in the very terms in which the emergence debate has been formulated throughout the years Nonetheless, it is only by pursuing this approach that we will be able to fully criticize such terms and commitments [18]

References

Alexander, S ([1920] 1979) *Space, Time, and Deity* Gloucester Peter Smith, 2 vols

Aristotle (1984) *The Complete Works of Aristotle* Edited by J Barnes Princeton Princeton University Press, 2 vols

Baas, N A (1994) Emergence, hierarchies, and hyperstructures In Langton, C G (ed) *Artificial Life III, Santa Fe Studies in the Sciences of Com-*

plexity, Proc Volume XVII Redwood City Addison-Wesley, pp 515–37
—— (1996) A framework for higher-order cognition and consciousness In Hameroff, S R , Kaszniak, A W, Scott, A C (eds) *Toward a Science of Consciousness* Cambridge, MA MIT Press, pp 633–48
Bedau, M (2002) Downward causation and autonomy of weak emergence, this volume (paper delivered at ISHPSSB meeting, Quinnipiac University)
Bickle, J (1998) *Psychoneural Reduction The New Wave* Cambridge, MA MIT Press
Blitz, D (1992) *Emergent Evolution Qualitative Novelty and the Levels of Reality* Dordrecht Kluwer
Boyd, R , Gasper, P & Trout, J D 1991 *The Philosophy of Science* Cambridge, MA MIT Press
Bunge, M (1977a) Levels and reduction *American Journal of Physiology* 233(3) R75–R82
—— (1977b) *Treatise on Basic Philosophy*, vol 3, Ontology I The Furniture of the World Dordrecht D Reidel
—— (1979) *Treatise on Basic Philosophy*, vol 4, Ontology II A World of Systems Dordrecht D Reidel
Campbell, D T (1974) 'Downward causation' in hierarchically organised biological systems In Ayala, F, Dobzhansky, Th (eds) *Studies in the Philosophy of Biology Reduction and Related Problems* Berkeley University of California Press, pp 179–86
Chaitin, G (1975) Randomness and mathematical proof *Scientific American* **232** 47–52
Cunningham, B (2001) The Reemergence of 'Emergence' *Philosophy of Science* **68** (PSA 2000 Proceedings) S62–S75
Dennett, D C (1987) *The Intentional Stance* Cambridge, MA MIT Press
—— (1991a) Real patterns *The Journal of Philosophy* **88**(1) 27–51
—— (1991b) *Consciousness Explained* Boston Little, Brown and Company
El-Hani, C N (2000) *Níveis da Ciência, Níveis da Realidade* São Paulo FE–USP PhD thesis
El-Hani, C N , Emmeche, C (2000) On some theoretical grounds for an organism-centered biology property emergence, supervenience, and downward causation *Theory in Biosciences* **119**(3–4) 234–75
El-Hani, C N , Pereira, A M (1999) Understanding biological causation In Hardcastle, V G (ed) *Where Biology Meets Psychology Philosophical Essays* Cambridge, MA MIT Press, pp 333–56
—— (2000) Higher-level descriptions why should we preserve them? In Andersen, P B , Emmeche, C , Finnemann, N O , Christiansen, P V

(eds) *Downward Causation. Minds, Bodies and Matter* Aarhus Aarhus University Press, pp 118–42
El-Hani, C N, Pihlstrom, S (2002) Emergence theories and pragmatic realism *Essays in Philosophy* 3(2) (http //www humboldt edu/~essays/pihlstrom html)
El-Hani, C N, Videira, A A P (2001) Causação descendente, emergência de propriedades e modos causais aristotélicos *Theoria* **16**(2) 301–29
Emmeche, C (1997) Defining life, explaining emergence On-line paper http //www nbi.dk/~emmeche/ (Published in two parts as Emmeche, C (1997) Autopoietic systems, replicators, and the search for a meaningful biologic definition of life *Ultimate Reality and Meaning* **20**(4) 244–64, Emmeche, C (1998) Defining life as a semiotic phenomenon *Cybernetics & Human Knowing* 5(1) 3–17)
Emmeche, C, Køppe, S, Stjernfelt, F (1997) Explaining emergence towards an ontology of levels *Journal for General Philosophy of Science* **28** 83–119
—— (2000) Levels, emergence and three versions of downward causation In Andersen, P B, Emmeche, C, Finnemann, N O, Christiansen, P V (eds) *Downward Causation Minds, Bodies and Matter* Aarhus Aarhus University Press, pp 13–34
Harré, R (1985) *The Philosophies of Science* Oxford Oxford University Press
Heil, J (1998) Supervenience deconstructed *European Journal of Philosophy* **6** 146–55
Hellman, G & Thompson, F 1975 Physicalism ontology, determination, and reduction *Journal of Philosophy* **72** 551–64
Hofstadter, D (1980) *Godel, Escher, Bach* Harmondsworth Penguin Books
Humphreys, P (1997) Emergence, not supervenience *Philosophy of Science* **64** (proceedings) S337–S345
Kim, J (1989) The myth of nonreductive materialism *Proceedings and Addresses of the American Philosophical Association* **63** 31–47
—— (1992) "Downward causation" in emergentism and nonreductive physicalism In Beckermann, A, Flohr, H, Kim, J (eds) *Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism* Berlin Walter de Gruyter, pp 119–38
—— (1993) *Supervenience and Mind* New York Cambridge University Press
—— (1995) Physicalism in the philosophy of mind In Honderich, T (ed) *The Oxford Companion to Philosophy* Oxford Oxford University Press, pp 679–80
—— (1996) *Philosophy of Mind*. Boulder Westview Press

—— (1997) Supervenience, emergence, and realization in the philosophy of mind In Carrier, M , Machamer, P K Mindscape *Philosophy, Science, and the Mind* Pittsburgh University of Pittsburgh Press, pp 271–93
—— (1998) *Mind in a Physical World* Cambridge, MA MIT Press
—— (1999) Making sense of emergence *Philosophical Studies* **95** 3–36
Lear, J (1988) *Aristotle The Desire to Understand* New York Cambridge University Press
Lloyd Morgan, C (1923) *Emergent Evolution* London Williams and Norgate
Mackie, P J (1995) Causality In Honderich, T (ed) *The Oxford Companion to Philosophy* Oxford Oxford University Press, pp 126–8
Midgley, M (1995) Reductive megalomania In Cornwell, J (ed) *Nature's Imagination. the Frontiers of Scientific Vision* Oxford Oxford University Press, pp 133–47
O'Connor, T (1994) Emergent properties *American Philosophical Quarterly* **31**(2) 91–104
Oppenheim, P & Putnam, H (1958) Unity of science as a working hypothesis In Boyd, R , Gasper, P & Trout, J D (1991) *The Philosophy of Science* Cambridge, MA MIT Press, pp 405–27
Pattee, H H (1973) The physical basis and origin of hierarchical control In Pattee, H H (ed) *Hierarchy Theory The Challenge of Complex Systems* New York Braziller
—— (2000) Causation, control, and the evolution of complexity In Andersen, P B , Emmeche, C , Finnemann, N O , Christiansen, P V (eds) *Downward Causation Minds, Bodies and Matter* Aarhus Aarhus University Press, pp 63–77
Pihlström, S (2002) The re-emergence of the emergence debate, this volume (paper delivered at ISHPSSB meeting, Quinnipiac University)
Polanyi, M (1968) Life's irreducible structure *Science* **160** 1308–12
Putnam, H (1967) The nature of mental states In Morton, P A (1997) *A Historical Introduction to the Philosophy of Mind Readings with Commentary* Peterborough Broadview Press, pp 320–7
—— (1994) *Words and Life*, ed by J Conant Cambridge, MA Harvard University Press
—— (2000) Aristotle's mind and the contemporary mind In Sfendoni-Mentzou, D (ed) *Aristotle and Contemporary Science* New York Peter Lang, pp 7–28
Riedl, R (1997) From Four Forces Back to Four Causes *Evolution and Cognition* 3(2) 148–58
Ross, D ([1923]1995) *Aristotle* London Routledge

Salthe, S N (1985) *Evolving Hierarchical Systems Their Structure and Representation* New York Columbia University Press
—— (1993) *Development and Evolution Complexity and Change in Biology* Cambridge, MA The MIT Press,
Smart, J J C (1959) Sensations and brain processes In Morton, P A (1997) *A Historical Introduction to the Philosophy of Mind Readings with Commentary* Peterborough Broadview Press, pp 233–42
Sosa, E (1995) Problems of metaphysics In Honderich, T (ed) *The Oxford Companion to Philosophy* Oxford Oxford University Press, pp 559–63
Van de Vijver, G , Salthe, S N , Delpos, M (eds) (1998) *Evolutionary Systems* Dordrecht Kluwer
Vinci, T & Robert, J S (2001) Aristotle and modern genetics Submitted manuscript
Walsh, D (2001) The unity of biology Natural selection and development (Paper delivered at ISHPSSB meeting, Quinnipiac University)
Williams, M (1973) The logical status of natural selection and other evolutionary controversies In *Philosophy of Science Symposium* Dordrecht D Reidel

Keywords
Emergence, downward causation, realism

Institute of Biology
Federal University of Bahia
Rua Barão do Geremoabo, 147
Campus de Ondina, Ondina
40170-290 Salvador, BA Brazil
charbel@ufba br

Notes

[1] The term 'emergence' and its derivatives are usually applied to a variety of categories, including properties, entities, objects, behaviors, phenomena, laws Nonetheless, all these different uses can be easily related and it is straightforward to trace them all back to the notions of 'emergent property' (see Bedau 2002) or 'emergent structure' I will refer mainly to emergent 'properties' in this paper

2 These arguments apply to theoretical predictability, not to inductive predictability (cf Kim 1999 8)

3 El-Hani & Emmeche (2000) discuss the difficulties faced by supervenience physicalism as a putative form of non-reductive physicalism and argue for a combination of supervenience and emergence as a way to propose an interlevel relationship meeting the double requirements of dependence and nonreducibility (see also Pihlstrom 2002)

4 What is at stake here is what Kim (1999 26) calls 'reflexive downward causation', observed when some activity or event involving a whole is a cause of, or has a causal influence on, the events involving its own microconstituents

5 In this connection, we can refer to Putnam's (1994, 2000) claim that a return to Aristotle can be a fruitful approach in the philosophy of mind, as it may be understood as partly a reaction to Aristotle's way of treating the different modes of causation on the grounds of the different ways the word 'cause' is used in different contexts As these usages correspond to different causal explanations, a pragmatic reading of Aristotelian distinctions between various modes of causation would mix ontological and epistemological issues We must, however, be careful in attributing pragmatist (or any modern) views to Aristotle (see El Hani & Pihlstrom 2002)

6 The Aristotelian notion of final causality is reinterpreted in this account as functional causality As El-Hani & Emmeche (2000 261) emphasize, this does not mean that 'function' should be identified with the Aristotelian 'final cause', nor that the notion of finality should be used at all It is simply that one should use a set of causal concepts rich enough to deal with the complexity of living beings, including a concept of function close to that usually employed in biological explanations

7 Pattee (1973, 2000) and Salthe (1985) examine the nature of constraints in complex systems

8 Surely, a complex of interrelated neuronal groups can be described, at a high level of abstraction, as a property (cf Kim 1999 6–7) I am only emphasizing that the supervenience base of a mental property is a lower-level state of relatedness among neuronal groups because this will make a difference to the arguments that follow

9 I am aware of Kim's (1998 80–7) comments about the importance of considering the distinction between 'levels' and 'orders' when discussing the idea of a hierarchy of properties in the brain I will develop my arguments here using the concept of 'levels', not 'orders' I will deal elsewhere with the consequences of the above distinction to the approach to emergence I am proposing here, in particular, as regards possible differences between

the relations of mental properties to their physical/biological bases and the micro-macro relation that obtains in the case of most biological systems. I have been treating these two cases as equivalent, but important dissimilarities could be unveiled by considering Kim's arguments about levels and orders.

[10] This is basically consistent with Kim's approach to the relation between macro-properties and their basic micro-constituents (see Kim 1999: 6–7), with the difference that my arguments intend to explore a feature whose consequences I think Kim does not sufficiently explore, namely, that the supervenience base of a higher-level property is a lower level state of relatedness among neuronal groups.

[11] It is certainly necessary to analyze whether the approach developed here entails the troublesome possibility that important features of our discourses such as values and freedom of action or of the will be regarded as mere illusions. For the moment, I will just suggest that certain current theories about brain structure and function, as Edelman's theory of neuronal group selection, might make room for free will, as they claim that the very neural bases of mental functions are largely conditioned by an individual's previous experience.

[12] I will not deal here with the technical details of Baas' approach. The notation was simplified following Emmeche (1997). It is not Baas' original one.

[13] It is important to say some words about 'identity theories'. First, they are opposed to any account of minds as non-physical objects, a 'non-physical' object being understood as an object which is not included in the physical domain (*contra* the hypothesis of the inclusivity of levels. See Emmeche et al. 1997, 2000, El-Hani & Emmeche 2000), and, thus, can violate physical laws or organizational principles. Second, this requirement for our understanding of the mind is captured in the claim that minds can be identified with brains, at least in the forms of life as we know it in its earthly instances. Mental events are, in this sense, identical with physical-biological processes in the brain. Third, an identity theory of mind states that each *instance* of a mental property has a neural correlate with which it is to be identified. Nonetheless, this claim cannot take the classical form of the mind-brain identity theory (e.g., Smart [1959]1997), simply identifying psychological types (properties, kinds) with physical types, as we have to take into account multiple realizability (Putnam [1967]1997). Nevertheless, we also have to deal with Kim's multiple-type physicalism (see Kim 1993, 1996, 1998). On the grounds of a tacit assumption in Putnam's argument, namely, that a physical state realizing a mental event is at least nomologically sufficient

for it, Kim concludes that, while multiple realizability offers a compelling argument against global mind-body reduction, it entails local reductions (restricted to specific domains, such as particular biological species or kinds of physical structures) from psychological predicates and theories to their biological/physical bases But we should never lose from sight that the neu ral correlates with which a mental property is identified are always complex sets of interrelated neuronal groups or neural maps, so that a huge amount of information is likely to be lost if we simply describe such biological bases as monadic properties

[14] Dennett (1991a) is the main source here, but see also Dennett (1987, 1991b)

[15] Dennett's mild realism has a number of connections with pragmatism, but some pragmatists worry about the confinement of ontological commitments to their scientific efficacy and fruitfulness As Pihlstrom (2002) argues, the exclusive emphasis on scientific usefulness is unnecessarily narrow I have to say that I agree with him, and believe that we should take into account not only a criterion of scientific utility, but also the pragmatic work done by emergent properties in our wider attempts to understand the world we live in The confinement of my arguments to scientific usefulness should be understood as simply a consequence of the fact that I am mainly interested in the theoretical entities used by the natural sciences

[16] Reproduced by kind permission of Dr Daniel Dennett

[17] Pihlström (2002) comments that, as they assume a strong scientific or metaphysical realism as a metaphilosophical basis, strong emergentists try to demonstrate that there *really* are emergent properties in the basic structure of the world itself, while weaker emergentists and non emergentists attempt to show the opposite As I give up that strong realistic premise, I am not trying to say that there really are no emergent properties in the structure of the world itself, but rather that emergents are good modeling/explanatory tools and, for being scientifically useful, should be regarded as real

[18] I am indebted to Jason Scott Robert, Sami Pihlström, Paulo Abrantes, Olival Freire Jr., João Carlos Salles and Osvaldo Pessoa Jr. for their comments Research supported by CAPES, FAPEX and CADCT