Synthese
DOI 10.1007/s11229-010-9723-5

Downward causation without foundations

Michel Bitbol

Received: 21 December 2009 / Accepted: 2 February 2010

Abstract Emergence is interpreted in a non-dualist framework of thought. No metaphysical distinction between the higher and basic levels of organization is supposed, but only a duality of modes of access. Moreover, these modes of access are not construed as mere ways of revealing intrinsic patterns of organization: They are supposed to be *constitutive* of them, in Kant's sense. The emergent levels of organization, and the inter-level causations as well, are therefore neither illusory nor ontologically real: They are *objective* in the sense of transcendental epistemology. This neo-Kantian approach defuses several paradoxes associated with the concept of downward causation, and enables one to make good sense of it independently of any prejudice about the existence (or inexistence) of a hierarchy of levels of being.

Keywords Emergence · Inter-level relations · Anti-foundationalism · Transcendentalism · Objectivity · Anti-realism · Quantum mechanics

1 Introduction

Downward causation looks impossible as a concept, but is well established as a fact. A top-down flux of causation going from an emergent level of processes (say biological or mental) to the fundamental level of processes (say physical) that is supposed to *underpin* it, sounds like a paradox. Yet this paradox seems to be forced upon us when we want to make sense of many phenomena ranging from psychosomatics to the mutual interaction between organismic and cellular life; or when we want to express certain facts of experience in which a change in our mental state has an immediate impact on our physiological state. My conviction is that what is likely to be

M. Bitbol (✉)
CREA, CNRS/Ecole Polytechnique, ENSTA/CREA 32, Bd. Victor, Paris 75015, France
e-mail: michel.bitbol@polytechnique.edu

Published online: 27 February 2010

paradoxical here is neither experience nor even its verbal expression in terms of downward causation, but the standard *metaphysical interpretations* that are superimposed onto them.

The metaphysical conception to be dispelled when downward causation is at stake is a *dualist* and *foundationalist* picture: a picture according to which downward causation only makes sense if there are at least *two* (ontologically) distinct levels of processes, with their own specific entities and properties able to influence each other by way of efficient causality. This is also a picture in which, conversely, if there *exists* only *one* fundamental level of being and process, then the very concept of inter-level causation is pointless. To counterweight this popular picture, I will advocate in this article the idea that the phrases 'downward causation' and 'upward causation' are loose ways of describing two modalities of *action* exerted on a process, which nothing prevents one from construing as non-dual. Acting globally (or coarsely, at a large scale) yields consequences that can be detected by experiments bearing on local levels. Conversely, intervening at the smallest accessible scale yields consequences that can be detected by experiments bearing on the larger scale. On the basis of this account, I will conclude that:

(i) Downward and upward causation are *neither* illusory *nor* inherently existent, but rather *indexed* by a certain level of intervention. Accordingly, downward and upward causation are not substantial concepts, but rather relational concepts throughout.

(ii) The relevant relations do not take place between otherwise self-subsistent entities. They contribute to the very *definition* of their terms. These defining relations are primarily the relations between the agents and the targets of their actions; and secondarily the relations between various domains of acting and experimenting.

Point (i) amounts to replacing ontological dualism with functional duality of domains of intervention. As for point (ii), it dispels foundationalism from the outset. Instead of the usual scheme of intrinsically "basic" and "emergent" levels of being (or levels of organization), one is left with flexible levels of (experiential and experimental) access, levels of action, levels of analysis, and levels of theorization.

This strongly anti-foundationalist approach does not fit well with the dominant scientific realist program of research in philosophy of science. But it makes perfectly good sense in the framework of an alternative program of research[1] inspired by Kant's "transcendental idealism," Husserl's phenomenology, and pragmatism. After all, in this kind of framework, "the access to an object partakes of the being of this object" (Levinas 2001, p. 161).

2 Paradoxes and difficulties of downward causation

The central difficulty, expressed by Kim (1999), is the threat of vicious circularity. Is it coherent, Kim asks, to assume that the presence of a certain lower-level process

[1] For applications of this research program to the philosophy of quantum mechanics, see Bitbol (1997, 1998, 2000, 2001).

is "responsible" for the presence of a higher-level process, and yet the higher-level process somehow exercises a causal influence on the lower-level process? Since the higher level entirely arises from the lower level, the idea of downward causation sounds either contradictory or redundant. It contradicts the rest of science if it really makes a difference, by violating the micro-laws that rule the low-level. And it is redundant if it merely restates in a different language, a coarse-grained language, what could, at least in principle, be couched in the fine-grained language of the detailed micro-processes.

Let us dig a little deeper below the surface of the two latter sharp statements. The first statement is meant to dismiss the idea of strong emergence, according to which the high-level processes are endowed with autonomous causal powers, and with ability to alter the low-level processes. It does so by assuming that for high-level processes to count as causal powers in the fullest sense, and to be able to alter anything significant in the lower level, they must induce a deviation in the *laws* of the micro-processes. But if this were the case, two common presuppositions of the scientific picture of the world would be denied: (a) the presupposition of nomological closure of the lower micro-*physical* level, and (b) the presupposition of *causal fundamentalism*, according to which "macro causal powers supervene on and are determined by micro causal powers" (Bedau 2002, p. 10). Strong emergence thus apparently amounts to an indefensible variety of ontological dualism.

Now, what about the second statement, which explores the other end of the spectrum of positions about emergence? It says that if the higher level makes no real difference, to wit (in this framework of thought) if no alteration of the micro-laws is imposed by it, then the very concept of a high level and of its causal powers is of no other significance than purely *verbal*. Instead of strong emergence, we are left with *nominal* emergence, and instead of dualism, the flattest possible version of monism: reductionism.

Kim's line of argument thus leads us towards an unacceptable alternative: either ontological dualism or ontological monism with epiphenomenalism.

Before I address these central difficulties, I wish to deepen the crisis of downward causation by three additional remarks.

(1) Since *efficient* causality seems to be at stake, one is inclined to examine the consequences of one of the most popular conceptions of the cause–effect relationship. This conception states that for a causal relation to arise, there must be a *transfer* of some conserved quantity (say *energy*), from the cause to the effect (Salmon 1984). Here, 'transfer' must be taken literally as: (a) loss of energy on the one side, and (b) gain of energy on the other side. But this immediately strikes us as problematic when upward and downward causation are concerned, if one embeds them within the framework of elementary physics interpreted with substantialist presuppositions. For, if this framework is upheld uncompromisingly, a pattern of organization can hardly be said to carry energy *by itself*, but only qua being made of interacting material constituents. This being granted, energy can be transferred from one material object to another object (by way of mediators), but not from a set of material constituents to their own global organization abstracted from what it organizes.

(2) According to the same conception, causal relations necessarily involve succession in time. Cases of simultaneous causation are shown to be only *apparent*, since, in view of a basic postulate of special relativity, they involve a quick but not instantaneous transfer of information and energy (Kistler 1999, p. 57). Now, it is quite clear that,

between an emergent level of organization and its alleged physical basis, there is no time-lag [except if one considers the progressive rise of the higher level in history, or the possibility of predicting it (Stephan 1998)]. Claiming that a set of interacting molecules *first* causes the biological organism to arise (upward causation), and that this biological organism *then* causes its own pattern of chemical reactions to adopt such and such configuration (downward causation), is an artifact of our discursive and sequential language. It would be more appropriate to claim that the basic elements *make up* the higher level of organization, *at the same time* as the higher level of organization *defines* the boundaries of the constituting interactions between the basic elements (Emmeche et al. 2000). "Making up" and "defining boundaries" are simultaneous, unlike causal relations.
(3) The standard idea of emergence is that a certain *form*, *pattern*, "topology" or "configuration" (Deacon 2003) supervenes on a basis made of the intrinsic properties of the components. This model is borrowed from classical physics. For instance, in classical electrodynamics, the interaction between N particles (as well as the configuration that arises from it) is supervenient on their electric charges and their positions. Within this configurational view of emergence, a dual picture of the various levels is automatically generated. The lower level is made of substances and properties, whereas the higher level is made of pure patterns. Such a duality triggers questions about a possible influence of one layer on the other, but the answer to those questions is preempted. Indeed, "upward causation" here holds only in the trivial sense of a set of properties determining their network of interactions. As for downward causation, it is made implausible by the difference in ontological status of the two levels. Figuring out how a configuration may act *as such* on a layer of basic properties is quite problematic. It would be more acceptable, in the spirit of the substantialist construal of the basic level, to say that the properties of the constituents cooperate to circumscribe their own range of possible variations, than to claim that their global configuration *causes* their alteration. But in this latter case, no motive for supposing a duality of levels is left; speaking in terms of mutual causation between these levels becomes pointless.

At this point, one realizes that, in a framework of thought that associates a substantialist conception of the basic level and a configurational conception of the emergent levels, talk of upward or downward causation is at the same time *irresistible* and reduced to the status of a *purely verbal machinery*.

3 Partial solutions: the inheritance of Aristotle and Kant

Several ideas have been offered to solve some of these difficulties, while sticking (implicitly or explicitly) to the substantialist view of the basic level. But, as we will see, they are not sufficient to account for the types of experiences and experiments that give support to the concept of downward causation.

The common ground of these ideas consists in using alternative concepts of causality, borrowed from Aristotle's well-known tetrad. The motivation for this attempt is that the formerly listed difficulties are connected with some typical features of the concept of efficient causality. So why not make use of formal causality, material causality, and even final causality?

It is quite natural, in the standard framework of thought, to consider that the basic constituent elements are the *material causes* of the emergent levels of organization, whereas, conversely, these emergent levels act as *formal causes* which circumscribe a domain of variation for the properties of their elements. Upward causation then instantiates the species of *material causality*, and downward causation the species of *formal causality*. The former concepts of "making up" and "defining boundaries" are thereby encompassed within a broad class of causal connections (Varela 1976), in spite of their being foreign to *efficient* causality.

The concept of formal causation can be developed by borrowing concepts from dynamics: the concept of *boundary conditions* and the concept of an *attractor*. The higher level of organization of a process can be described as a self-imposed *boundary condition*. Among all the possible solutions of the differential equations describing the dynamics of the parts, such a boundary condition selects a range of *accessible* solutions. This substantiates the idea that the higher level *constrains* the evolution of the constituent elements, at the same time as it is *made up* by them. Alternatively, the higher level of organization can be characterized as a global *attractor* for the trajectories of the constituent elements in phase space. This attractor *shapes out* the evolution of the elements at the same time as it is *determined by* their states and interactions (Emmeche et al. 2000).

The fourth and last species of causality shows up at this point. A two-way process of making up and constraining, within an integrated unit called an "organism," gives rise to the *appearance* of what Kant called "natural purpose" in his *Critique of Judgment* (Sect. 64–66). According to Kant, in order for something to be construed as a natural purpose, a crucial condition is that "(. . .) the parts of the thing combine into the unity of the whole because they are *reciprocally cause and effect of their form*" (Kant 1987, p. 252). Such an integrated process can indeed conveniently be thought of as produced by a *final* cause, insofar as everything in it occurs *as if* the very *concept of the organism* as a whole were the cause of the distribution of its parts.

In another paragraph, Kant characterizes this reciprocity of action between the parts and whole by using a vocabulary that is strikingly close to the current one, yet virtually reverted (Kant 1987, p. 251). In Kant's terms, a "descending series" is a chain of determination of effects by efficient causes. And an "ascending series" is an influence going from the "purpose" of the whole process (as conceived by *reason*), to its constituent "things" (as objectified by *understanding*). The most immediate illustration of this bi-directional circulation of causes is human goal-oriented action. A tool is an efficient ("descending") cause of transformations of our environment, whereas our representation of the form of the desired effect is the final ("ascending") cause of our making and using the tool. By a protraction of reason, a natural system that does *not* result from the intention of a concrete agent can nevertheless be thought of under the *idea* of purpose. This is especially the case of living beings (or "organized beings"), since "(…) an organized being has within it *formative* force, and a formative force that this being imparts to the kinds of matter that lack it (thereby organizing them)" (Kant 1987, p. 253). Organized beings can be considered as imparting an "ascending" ("downward" according to the current lexicon) influence on their material constituents. Their organization can be taken as imposing a feedback on what is organized. It can therefore be thought of *as if* it were the final cause of the whole process. A. Weber and F. Varela summarized Kant's lesson by calling *autopoiesis* (which involves the complete circuit of reciprocal causalities, from parts to whole and from whole to parts) a case of "embodied finality" (Weber and Varela 2002).

These conceptions combining material, formal, and final causes provide us with an elegant ordering scheme of our conception of parts and wholes. But they should not be taken as anything more than that: a satisfaction of our urge for conceptual harmony. Kant himself forcefully pointed this out: Final causality, he wrote, only provides our reflective power of judgment with a *regulative principle* (a compelling guiding thread). By contrast, efficient causality operates as a *constitutive principle* (a condition of possibility of objective knowledge). Then, unlike the category of efficient causality, which is imposed in advance by our *understanding* in order to constitute objective knowledge, final causes are thought of by our *reason* in order to fulfill a subjective need of systematic unity of knowledge (and are also required from our reason by the empirical fact of organismic life).

At any rate, a combination of material, formal and final causality does not solve all the difficulties listed in Sect. 2. It is still dubious whether stating *separately* the self-imposed boundary conditions and their self-imposing elements, or the global attractor and the self-organizing constituents, is more than a verbal (or heuristic) device. And we still have to give the importance it deserves to the fact that changes at upper levels of organization can *trigger* changes at lower levels.

The persistence of this list of problems is a challenge that can hardly be addressed unless one accepts a minimalist and deflationary conception such as "weak emergence." Remember that defending weak emergence means claiming that "the system's global behavior derives just from the operation of micro-level processes, but that the micro-level interactions are interwoven in such a complicated network that the global behavior has no simple explanation" (Bedau 2002, p. 10). The initial motivation of the conception of weak emergence was to indicate a middle way between strong emergence with its dualist flavor, and purely nominal emergence in which the higher level is nothing more than an artifact of language. But at the end of the day, according to its best supporter, it appears that "weak emergence is (nothing but) a proper subset of nominal emergence" (Bedau 2002, p. 11; see Bedau 2008a for a qualification). Weak emergence is the special subset of nominal emergence that corresponds to the case where excessive complexity in practice prevents any detailed micro-explanation of macro-behavior. If one retains the latter criterion, it may appear that in most cases "a weakly emergent property is a *limitation of the observer*, not a property of the system" (Batten 2008). A persistent distinction between weak emergence and flat nominal emergence therefore depends on showing how this limitation of the observer can sometimes be so extreme that it displays the characteristics of an impassable boundary. This may occur whenever the practical limitation in our aptitude to explain macro-behavior is so challenging as to acquire a sort of "in principle" status (Bedau 2008b). This may also occur if many details of the micro-explanations are irrelevant for accounting for the macro-behavior, and if, accordingly, a given macro-explanation can be underpinned by several different micro-stories (Bedau 2002).

No genuine (upward or downward) causation, and ontologically little more than nominal emergence: This is the disappointing outcome of research developed under a substantialist construal of the basic elements. Is there any alternative left? I think there is, provided the substantialist presupposition is dropped at *every single level* of description.

In order to identify this way out I wish to proceed with prudence. Let me first step down, in the next section, from big metaphysical issues about *causality* to practical issues about what counts as a *cause* in a laboratory and in everyday life (Bitbol 2007a).

4 An interventionist conception of downward causation

In ordinary circumstances, we are overwhelmed by the number of antecedents that are likely to have caused a certain event. Among all these antecedents, the one that we usually consider as the cause of this event is either the antecedent that appears to have *changed* recently, or, more convincingly, the one that we can get *control* of Mill (1851) and Pattee (2000). Control, intervention, and action are central touchstones when causes are at stake. This elementary criterion for isolating causes then gives rise to a full-blown conception of causality, called the "interventionist" theory. According to this theory (Von Wright 1974; Price 1992; Woodward 2003; Gillies 2005), configuration A is a cause of the distinct configuration B if: (i) whenever A has been actively set up by any means, B occurs (with probability p); (ii) whenever A has been actively removed, B does not occur (or occurs with probability $p' < p$). Interestingly (Ducheyne 2006), this definition of causality was retained by one of the founders of the modern science of nature, Galilei (1612, p. 425): "The cause is that which, when posited the effect follows, and when removed the effect is removed."

We can then speak of "downward causation" in the following circumstances: (i) whenever an upper-level antecedent has been actively set up by means of a "coarse" instrument able to alter macroscopic patterns, certain elementary phenomena are observed by means of experimental devices designed for microscopic analysis; (ii) whenever this upper level antecedent has been actively removed by "coarse" instruments, those elementary phenomena are not observed. One can define "upward causation" likewise by just inverting the previous order of intervention and observation.

Several difficulties of the concepts of upward and downward causation are formally solved by this theory of causality.

One solution concerns the non-causal (or at least non-*efficiently* causal) status of constraining boundary conditions. Even though the system of mutual constraints or the boundary conditions cannot be considered as causes by themselves, an *alteration* of constraints at one level may cause events at another level. For instance, the constraints set by the mental state of an individual cannot be said to *cause* neural and molecular events in the body of this individual; it is more reasonable, according to the dynamical conception, to consider them as co-extensional. But *actively triggering a change* in this mental state by psychological means may result in changes in her sympathic–para-sympathic balance, thus giving rise to changes in cell metabolism, increased liability to genetic mutations, or activated blood coagulation (Geiser and Halbrecht 2008). If one accepts that causing means *altering* a certain configuration rather than *being* this configuration, one easily makes sense of such psychosomatic phenomena.

This continues to be true when the putative levels are so intermingled that it looks artificial to hold a dualist picture of them. Indeed, each *type of intervention* can be set up in such a way that it (i) concerns specifically and (ii) contributes to *defining* one level of organization, by being more or less coarse-grained and more or less tuned to certain configurations. One can then institute a duality of targets for action, in spite of there being no reason to assume an ontological duality between the levels, and even no reason to think that questions about what levels are "in themselves," independently of any method of assessment, make sense. The difficulty that arose from the tension between a monist conception of natural processes and the dualist or pluralist vocabulary of "levels" is thus overcome.

The interventionist conception of causality also helps to solve another difficulty, about the temporal succession of causes and effects. Even though the evolution of the higher-level patterns can be taken as simultaneous with the corresponding lower-level processes, the *intervention* itself is prior to any alteration of the targeted pattern. True, an intervention on the macroscopic level does not alter *retroactively* the micro-conditions from which the macro-level allegedly arose (which would fall prey to Kim's accusation of vicious circularity); but it can alter the micro-conditions for its *future* evolution. The difficulty of the apparent simultaneity of downward causes and their effects thus disappears. Notice that this solution to the time-lag problem formally resembles a proposition of the paradigm of "weak emergence:" "(. . .) A weak macro-cause cannot alter the conditions from which it arose. At most it can alter the *conditions* for its subsequent survival (. . .)" (Bedau 2002, p. 22). However, there is a major difference. In the interventionist conception of causation, the verb 'to alter' has the concrete meaning of a disturbance introduced from outside, a disturbance which contributes to defining its own specific target through the scale and "coarseness" of the instrument one uses for triggering it. Instead, in the framework of "weak emergence," the verb 'to alter' is used only in order to connect verbally two contingent levels of conceptual or computational analysis of an otherwise unique dynamical process, which concerns a set of interacting "basic elements." No *real* influence is supposed to take place between levels of conceptual or computational analysis. One can *say* that a present traffic jam alters the conditions of the subsequent motion of the cars that compose it. But, admittedly, in the "weak emergence" framework, this is only a way of speaking used in order to compensate for a lack of detailed knowledge, and unable to hide the fact that *ex hypothesi, all there is* is a set of cars that at each instant move according to the distance of their immediate neighbors.

5 Interventionist view of downward causation: first objections and replies

As long as one sticks to a substantialist construal of the basic elements, however, the tentative solution offered in the previous section is bound to look parochial or anthropocentric. Indeed, in this framework of thought, action, intervention, and higher levels of organization as well, are all supposed to be derivative when compared to the basic constituents. Accordingly, Von Wright's theory of causality is usually taken as only epistemological, not ontological; it is supposed to clarify the way we happen to *know* causes, not what causes *are*.

As for the idea that intervening at a high level of organization may be the genuine *cause* of changes at a lower level, it raises at least two objections. Firstly, it may be replied that saying that one has imposed alterations to a given high level of organization is only a way of expressing one's ignorance of the detailed microscopic changes that have been triggered *in reality*. Secondly, it may be argued that the so-called high-level cause of a subsequent change at a lower level is at most a *releasing* cause of this alteration, not its *integral* physical cause, which is likely to involve the entire network of interactions between the basic constituents.

These objections are easily met within another framework of thought, which I would like to characterize as *conceptually symmetric*.

The standard conception of emergence is characterized by a twofold asymmetry:

(i) The usual picture of emergence is asymmetric because it assumes that the basic level is made of individual entities, whereas the higher levels are made of structures, or patterns.
(ii) The usual picture of emergence is also asymmetric because it assumes that the determinations of the basic level are "properties," in so far as they are "proper" to, or inherent to, the elementary entities, whereas the determinations of the higher levels are *relative to* certain (contingent or necessary) *limitations* of the cognitive instruments used to assess them.

But, in view of recent developments of modern physics, both asymmetries appear unwarranted in principle.

On the one hand, in quantum physics, it can hardly be assumed that there is an ultimate basic level of processes made of individual little things able to interact with one another (Bitbol 1996, 2007b). Setting aside any conceptual trick such as hidden variables, the so-called elementary particles have to be treated as non-individuals, as mere units of a limited set of "sorts," and thus as formal rather than substantial entities. This becomes even clearer in Quantum Field Theory, where cardinals of subsets of particles are in one-to-one correspondence with quantized modes of excitation of fields (Teller 1995). Accordingly, particles are *de facto* treated as patterns or configurations, rather than as substantial entities (Bickhard and Campbell 2000). The analysis of a level of organization in terms of structures, patterns, and topological configurations imposes itself throughout the scale of levels, *even at the lower accessible level* (Campbell and Bickhard 2009). The first asymmetry of the standard picture disappears thus.

On the other hand, in quantum physics, one can no longer overlook the fact that determinations are exclusively relational or contextual. Here, the cognitive relation cannot be taken as merely revealing, but as *constitutive*. This circumstance pushes the so-called "basis" of emergence surprisingly close to the higher levels of organization. The determinations of the lower-level entities are just as much *relative* to certain *limitations* of the appropriate cognitive instruments as those of the higher-level entities. The only difference is that the limitations of the cognitive instruments are impassable in micro-physics, whereas they can in principle be overcome (notwithstanding complexity) at higher levels of organization. This latter difference may explain why it is tempting to privilege microscopic physics, but it does not justify forgetting the universality of the mode of definition of attributes relative to a mode of experimental access. Once this universality is recognized, the second asymmetry of the usual picture disappears too.

Without any ontological asymmetry, without any stubbornly substantialist construal of the so-called "basic" level, the dignity of the interventionist conception of causality is dramatically improved. On our construal, it becomes pointless to maintain a distinction between epistemological and ontological conceptions of causation, since nowhere can one point towards something like an ultimate inherent property that sets the standard of ontology. At every level, the cognitive relations and the experimental interventions are constitutive of their object and of their target respectively. Therefore, it is not correct to claim that what is described as an intervention on a higher level of organization is *in reality* an intervention on its elementary parts. For these putative parts are no more inherently real than the upper levels of organization; they are just as much relative to a certain mode of investigation or action as these upper levels. At most, one can say that intervening on a high level of organization means acting on a domain of cognition that *would* be analyzed into parts *if* a microscopic mode of intervention were used instead. But this counterfactual statement should not be mixed up with an ontological statement about the elementary parts. For the analytic activity relative to which the parts are defined is likely to be *exclusive* of another activity performed at an organismic or global level (see the next section for an example of mutual exclusivity).

Similarly, it becomes pointless to make a distinction between the releasing cause of a certain alteration and the "physical" cause. For, in our symmetric framework of thought, the releasing cause is not only triggered but also *constituted* by an appropriate intervention. And the physical (microscopic) cause is evoked only by counterfactual reference to a network of factors that *could* have been constituted by another (possibly incompatible) type of intervention.

6 A quantum model of downward causation

An illustration of this symmetric and interventionist conception of downward causation will now be given by examining a familiar quantum mechanical model. The motivation for displaying this kind of model is that in standard quantum mechanics, the ideas that observables are operationally defined and that experimental interventions are *co-constitutive*, rather than merely revealing of properties, are built into the theory.

Let us concentrate on the state vector of a two-particle system in which each particle (each sub-system) can be submitted to an elementary two-values experiment (such as measurement of a component s_z of spin 1/2). The possible outcomes of the measurement are labeled $+$ and $-$, and the corresponding eigenstates of observable s_z are: $|+\rangle$ and $|-\rangle$.

The generic state vector of a two-particle system then takes the form of a linear superposition of tensor products of these eigenstates:

$$|\Psi^T\rangle = \sum_{ij} c_{ij}|ij\rangle = c_{++}|++\rangle + c_{+-}|+-\rangle + c_{-+}|-+\rangle + c_{--}|--\rangle$$

This state vector $|\Psi^{\mathrm{T}}\rangle$ is usually not factorizable. The contributions of the two sub-systems cannot be set apart; they are said to be "entangled." According to Schrödinger (1935/1983), who first formulated the concept of entanglement, this means that "[b]est possible knowledge of a whole does not necessarily include (best possible knowledge of) its parts. (...) The whole is in a definite state, the parts taken individually are not" (p. 161).

A precise way of displaying this type of relation of parts and wholes in quantum mechanics consists in comparing different states of the two-particle system taken as a whole, and demonstrating that these differences have no counterpart whatsoever in the individual states of the sub-systems.

To show this, let us first consider four mutually orthonormal (entangled) states.

$$\begin{aligned} |\Psi^{+}\rangle &= 2^{-1/2}\,(|+-\rangle + |-+\rangle) \\ |\Psi^{-}\rangle &= 2^{-1/2}(|+-\rangle - |-+\rangle) \\ |\Phi^{+}\rangle &= 2^{-1/2}(|++\rangle + |--\rangle) \\ |\Phi^{-}\rangle &= 2^{-1/2}(|++\rangle - |--\rangle) \end{aligned} \tag{1}$$

These are global states for the two-particle system. But what about the individual states of the sub-systems? These individual states cannot be denoted here by a state vector. At most, one may associate each sub-system with a *density operator* ρ, obtained by taking the *partial trace* of the matrix representing the density operator of the two-particle system. The problem is that these individual density operators are *exactly identical* to each other, irrespective of the whole system's being in any one of the former four states. They *always* write:

$$\rho = 1/2(|+\rangle\langle+|+|-\rangle\langle-|) \tag{2}$$

Therefore, there exists no difference in the states of the parts that may account for the difference in the state of the whole (Maudlin 1998). *The differences in the state of the whole are not supervenient on corresponding differences of the parts.* This is one of the most remarkable holistic-like features of quantum mechanics among those used by Humphreys (1997) to develop his conception of strong emergence based on what he called "fusion" of micro-properties (Seevinck 2003).

Another, less documented consequence of entanglement is that any attempt at *separating completely* the states of the parts from the state of the whole is bound to fail. This sounds strange, since we have accepted (at least tentatively) that the state of each component particle (or sub-system) of a two-particle system can be described separately as a density operator. But the latter kind of density operator is quite peculiar: Using the terminology of D'Espagnat (1989), it corresponds to an *improper* statistical mixture, not a *proper* mixture. Not all the relevant information about the sub-system is contained in its density operator (obtained by *partial tracing*). A fraction of this information is missing, and it is contained only in the state vector of the overall system. Any separate description of parts, and then any *dual* description of parts and whole, then looks artificial.

As a consequence, the concept of inter-level causation looks nonsensical in the highly holistic domain described by quantum mechanics. Since there is no way to separate the states of the parts from the state of the whole, it sounds absurd to call one the cause and the other the effect, as if they were two different things.

Now, even though one cannot separate the contribution of the parts and of the whole within an entangled state, there is a sense in which one can *intervene* selectively either on the parts or on the whole. For one can *intervene* either analytically or globally. Measuring individual observables (or preparing sub-systems in eigenstates of such observables) is tantamount to intervening on the parts; whereas measuring global observables (or preparing a system in an eigenstate of such an observable) is tantamount to intervening on the whole. Thus, measuring the component s_z of the spin of each particle belonging to a two-particle system means intervening on its parts. But preparing this two-particle system in one eigenstate of a global observable (in a 4-dimensional Hilbert space) means intervening on its whole. In order to display the consequences of this distinction, let me give a new list of eigenstates corresponding to an interesting global observable. This list is made of one "singlet" and three "triplet" states (Cohen-Tannoudji et al. 2006) :

$$\begin{aligned} &\text{(Singlet)} \quad |\Psi^-\rangle = 2^{-1/2}(|+-\rangle - |-+\rangle) \\ &\text{(Triplet)} \begin{cases} |\Phi^0\rangle = |++\rangle \\ |\Psi^+\rangle = 2^{-1/2}(|+-\rangle + |-+\rangle) \\ |\Phi^1\rangle = |--\rangle \end{cases} \end{aligned} \tag{3}$$

If a two-particle system initially in state $|\Psi^{\mathrm{T}}\rangle$ is prepared in state $|\Psi^-\rangle$ or $|\Psi^+\rangle$, we can be sure that local measurements of the individual observable s_z performed on the two sub-systems will yield one $+$ and one $-$. And if a two-particle system initially in state $|\Psi^{\mathrm{T}}\rangle$ is prepared in state $|\Phi^0\rangle$ (respectively $|\Phi^1\rangle$) we can be sure that local measurements of the individual observable s_z performed on the two sub-systems will yield two + (respectively two $-$).

This means that *intervening on the whole has effects on values of observables bearing on the parts*.

In other terms, preparing a two-particle system in a certain state of a global observable amounts to a variety of *downward causation*, since this determines the distribution of values one may subsequently obtain if local observables (such as s_z) are measured separately on the sub-systems. Global preparation *downwardly causes* a distribution of values of local observables. The converse also holds, of course. If we measure the values of the local observables s_z on the sub-systems first, and get the outcomes i and j, then by measuring the global observable "total spin" one can find the system only in the factorizable state $|ij\rangle$. Local measurements here *upwardly cause* a global state.

Notice that an intervention on the whole can by no means be identified with a coarse-grained action that "in reality" bears on the parts. Indeed:

(i) In standard quantum mechanics, relations with experimental devices are *constitutive* in the strongest sense: They determine generically which observable is concerned (global or local), and individually which value of this observable is realized.

(ii) Global observables and local observables generally do *not* commute with each other. The global observable associated to the list of eigenstates (3) thus does not commute with the local observables s_z. Therefore, the constitutive act corresponding to measurement of a global observable is generally *incompatible* with the constitutive act(s) corresponding to measurement(s) of local observables.
(iii) As a consequence of points (i) and (ii), the values of local observables generally cannot even be ascribed any *counterfactual definiteness* when the value of a certain global observable is defined and obtained by appropriate measurements.

According to (iii), the conditions that allow ascription of a global property to a quantum system generally *exclude* the conditions that allow ascription of local properties. If a global property is defined and ascribed, *it is then wrong to say that there exist underlying local properties on which one is acting "in reality."*

The non-dualist and anti-foundationalist features of this quantum model are obvious at this point. Let us recapitulate them:

(a) The holistic-like character of the state representation of quantum systems makes any two-level description of these systems look artificial.
(b) In this representation, one cannot even use the usual concept of supervenience to reconcile the ideas of dependence (of the global level on the local level) and autonomy (of the global level with respect to the local level). Indeed, there may exist differences between global states without any corresponding differences between the local states that are supposed to underpin them.
(c) Properties and states cannot be treated as preexistent intrinsic features. They must be construed as relational. Properties are thus replaced by "values of observables," in so far as they are relative to particular instantiations of measurement acts.
(d) Experimental interventions simultaneously single out and *define* a type of property: global properties are defined relative to the measurement of large-scale observables, and local properties are defined relative to the measurement of low-scale observables. The conditions relative to which global observables are defined generally *exclude* the conditions relative to which local observables are defined.
(e) Upward and downward causation make good sense as soon as one does not pay too much attention to state representations, and focuses on the effect of (local or global) experimental *interventions*. Measuring or preparing global observables may influence the results of later measurements of local observables, and conversely measuring or preparing local observables may influence the results of later measurements of global observables.

Points (a) and (b) dispel dualism of levels, and points (c) and (d) challenge foundationalism. As for point (e), it elaborates on the status of downward and upward causation in a non-dualist and anti-foundationalist framework.

7 About some possible misunderstandings

The conception of downward causation developed in the previous sections, which we may now call the "interventionist-constitutive view," is unfamiliar enough to trigger misunderstandings. The purpose of this section is to dispel some of them.

The first misinterpretation concerns the role of quantum-mechanical concepts in my defense of the interventionist-constitutive view of downward causation. The central argument adduced to defend the interventionist theory of causality against the accusation of anthropocentrism in Sect. 5 was borrowed from the philosophy of quantum mechanics, and even quantum field theory; but is it really indispensable to rely on such far-fetched and allegedly exotic domains of physics? The quantum model presented in Sect. 6 has features that make inter-level causation plausible; but is it representative of more standard cases of downward and upward causation? Here are some tentative answers.

In the argument of Sect. 5, relying on the philosophy of quantum mechanics seems to be a contrived strategy only because one forgets that one is dealing with matters of *principle*. The problem here is whether or not there is *in principle* a difference between "ontological" causes and a "purely epistemological or anthropological" definition of causality. Claiming that there is indeed such a difference presupposes a distinction between: (i) lower levels of *basic individual constituents* where the experimental intervention procedure helps us to *reveal* the true causes identified with micro-properties of individual constituents, and (ii) upper levels of *organization* where the experimental intervention can trigger a chain reaction of real micro-causes only by coarse alteration of their global pattern. So, what about the validity of this distinction postulated by the view I am criticizing?

Consideration of intermediate levels is of little help for addressing such a question. Indeed, any intermediate entity (such as a molecule, a cell, an organism, etc.) can be seen indifferently either as an elementary individual constituent of the upper level of organization, or as a pattern emerging from the lower level of organization. Accordingly, an intermediate entity can be construed either as an unquestioned (and therefore de facto absolute) elementary building block for the upper levels of organization, or as relative to the coarse mode of access that singles out its own level of organization. An intermediate entity can also be taken either as the real elementary cause of a certain process taking place at some upper level of organization, or as a coarse target for interventions meant to trigger events at its own level of organization (or at lower levels). In other terms, consideration of intermediate levels can display only *functional* or *instrumental* asymmetry between basic elements and emergent patterns. Intermediate levels present us with no motive to declare that there is a *fundamental* and *irreducible* difference between: (i) absolute, ontological, efficient causes, and (ii) relative, epistemological, releasing causes at higher levels defined (nay *constituted*) by a scale of experimental intervention. Actually, every such intermediate level falls under the second (relative, epistemological, releasing) category of causes, and it is only by reference to upper levels that one can conventionally (and provisionally) grant it the *role* of a repository of basic entities and real causes.

The only unshakable reason for taking seriously the distinction between the two types of causes, and for accepting that there is a difference of *principle* between an ontological and a purely epistemological-anthropological conception of causes, would then be to have the assurance that there exists an *ultimate* basic level of individual entities endowed with intrinsic properties. For, then, these micro-properties would inherently count as real causes, and any other higher-level process circumscribed as a target for macroscopic action could give rise *only* to "epistemological" or "relative" releasing causes. Unfortunately (for this way of thinking) micro-physics has strong arguments in store against such a concept of an ultimate basic level, because, (a) as I mentioned in Sect. 5, the so-called "elementary" particles are replaced by quantum field patterns, and also because (b) *not even* quantum fields can be considered as basic entities, since they all appear (in view of *reorganization*) to be nothing more than "effective fields" for ever deeper levels of field-theoretic entities (Hartmann 2001; Castellani 2002). Therefore, at *no* level, even the deepest and (presumably) most elementary level of micro-physics, are there anything like absolute properties and absolute causes to be opposed to the (so-called) anthropological, epistemological, relative causes that can be defined by varying the scale of intervention on putative antecedents. There is *no* cause that one could consider as a "primary quality" of matter; every cause arises as a "secondary quality" in so far as it is relative to a certain procedure of intervention (Menzies and Price 1993). In the absence of any contrast of this kind *even in the putatively ultimate domain of micro-physics*, the interventionist definition of causes is bound to become paradigmatic *in principle*, not only *in practice*.

A second possible misunderstanding is to think that the example developed in Sect. 6 is an exceptional case of downward causation, which works only in the unique epistemological configuration of quantum mechanics. This example is in fact less isolated and less exceptional than it may appear. Indeed: (i) the epistemological features of quantum mechanics represent only a reinforced version of a universal situation, and (ii) there are some non-quantum cases that fully implement these epistemological features in their reinforced version.

About point (i), even classical physics can be seen as instantiating a type of science that only deals with some sort of "secondary qualities" and *never* with "primary qualities." This was Kant's interpretation. In his *Prolegomena*, he insisted that even spatial and kinematic predicates can (and should) be considered "secondary" in Locke's sense (Kant 1955, Sect. 13; Kant 1992). They should be considered secondary because they are relative to the pure form of our *sensibility*. By extension, Kant could have said that ascribing something the status of a "substance," or the status of a "cause," is also to be considered as "secondary," because both ascriptions are relative to the pure form of our *understanding*. The strategy of universalization of relational/contextual predicates in every domain and at every level can thus already be defended in classical science, against the substantialist and foundationalist conceptions which underpin reductionist approaches of interleave relations. This strategy is still used nowadays, in the framework of the uropoietic (Maturana and Varela 1980) and "ecological" (Gibson 1979) theories of cognition, in biostatistics inspired by Von Ukeüll's theory of *Unmelt* (Hoffmeyer 1997), etc.

About (ii), it is now increasingly recognized (Lambert et al. 2006; La Mura 2005; Atmanspacher et al. 2004) that certain domains of human sciences (economy, psychology of perception, rational choice theory, etc.) share *exactly the same* (and not just analogous) characteristics and backbone structure as quantum mechanics. The reason for this strong similarity is simple to understand: In all these cases, just as in micro-physics and unlike macro-physics, the predicates cannot be disentangled from the methods of access or from the agents (Bitbol 2009, 2010). The most basic mathematical structures of quantum theory can then be transposed to these domains, including those involved in non-factorizations of state vectors. Generalized quantum formalisms able to apply beyond the realm of physics have been formulated, and applied with success to several situations in the human sciences (Atmanspacher et al. 2004).

For all these reasons, the quantum case of interventionist-constitutive downward causation discussed in Sect. 6 is likely to have a much higher degree of generality and relevance than expected.

A third misunderstanding may arise about the topic of *intra*-level causal powers. These have just been ignored in this article until now, in favor of a systematic study of inter-level causation. Has the interventionist-constitutive conception of causation anything to say about intra-level influences? Besides, if (as the interventionist-constitutive conception requires) one grants an equal status to high-level *intra-level* causes and to low-level *intra-level* causes, isn't there a redundancy of causes? How can one reconcile this equality of status between high-level and low-level causes with the idea of *causal closure* of the domain of physics?

The first question can be answered unambiguously in the positive. The interventionist-constitutive conception of causation is indeed applicable to the case of intra-level causation. For this, it is enough to consider that both the instrument of intervention and the instrument of observation are designed for targeting (and thereby co-defining) the same level. The two clauses that define an intra-level cause according to the interventionist conception are then the following: (i) Whenever an antecedent has been actively set up by means of a level L-specific instrument I_{L+}, certain phenomena of level L are observed; (ii) whenever this antecedent has been actively removed by another level L-specific instrument I_{L-}, the former phenomena of level L are not observed.

Now, in order to answer the second question, one has to take seriously the word 'constitutive' in the expression 'interventionist-constitutive conception of causation'; just as seriously, in fact, as in Kant's theory of knowledge. There, the principles of understanding are constitutive of objects in so far as they are conditions of possibility (or necessary presuppositions) for organizing "rhapsodies of sensations" into objective knowledge. In our case, using the word 'constitutive' is meant to express: (i) the relativity of any causal scheme to the method of active substitution of antecedent by means of various instruments adapted to various scales or levels, (ii) the thesis that, therefore, any causal scheme has these methods and instruments as a necessary presupposition, and (iii) the conviction that *there is no "fact of the matter" as to which of the many instrument-relative causal schemes is more "real" than another*. No level of organization can claim any privilege for itself, because every such level is defined (or "constituted") by a certain scale of intervention and observation. Moreover, no absolute meta-observer, no "view from nowhere," is available to select one pattern of causes at a certain agent-relative level as the "truly efficient" one. This does not threaten the thesis of causal closure of the domain of physics, but only denies it any ontological significance. Causal closure here means only that it is possible to establish a systematic and self-sufficient network of causal connections relative to a single scale of intervention and experimental access, without having recourse to any other scale of intervention and access. This being granted, causal closure of a low level of organization (say the level of micro-physics) is perfectly compatible with the thesis that there are also efficient causes at an upper level of organization. Causal closure at one level shows only the efficacy and reasonable exhaustiveness of the procedures of intervention and access that define this level, in yielding a coherent picture *relative to these procedures*. Nothing then prevents one from obtaining another coherent picture, possibly with causal closure, at another level and with other procedures of intervention and access. These two internally coherent pictures and causal closures can *both* be valid, each at its own level, and relative to its own set of procedures of intervention and access. Far from being mutually exclusive (as an ontological conception of causation would require), they can be made mutually consistent. Of course, mutual consistency here does not mean that they are somehow identical, but only that they are *inter-translatable*, as if they were two different languages (say a micro- and macro-language).

For instance, according to the old-fashioned dogma of molecular biology, one can say that:

(a) A gene *causes* a phenotypic feature of the cell.
(b) A DNA strand's opening offers regions of selective affinities for nucleotide basis, RNA polymerase moving along the DNA then *causes* successive pairing of the nucleotide basis of mRNA with the nucleotide basis of DNA (this is called "transcription"), and finally mRNA, along with ribosomes and transfer RNA, *causes* systematic pairing with amino-acids thus resulting in a certain protein primary structure (this is called "translation").

In this standard framework, the biological-genetic causal scheme (a) does not exclude the chemical-molecular causal scheme (b), nor conversely; it is compatible with it under certain rules of (complete or incomplete) inter-translation. These inter-translatable causally closed schemes are not even redundant. Indeed, each one of them is indispensable relative to the corresponding procedure of intervention and access. In the framework posited by *this procedure*, it could not be replaced by another causally closed scheme without loss in explanatory power.

8 Downward causation beyond present interventions

Another challenge to the interventionist conception must be addressed at this point. Is there a sense in which an event at a certain level L can be said to cause an event at the lower level l, *even in the absence of a present intervention* at level L? As we will now see, this question can be answered without reference to "causes in themselves," by studying the mode of interconnectedness between *domains of knowledge defined relative to different modes of access and concepts.*

Bich (2009) recently formulated this strategy, with inspiration from Rosen (1991) concepts of scientific models and relational biology. To begin with, one notices that it looks nonsensical to posit real interactions between entities of two models (say a high-level and a low-level model) of a certain process, since each model is relative to a completely distinct class of modes of access and modes of conceptualization. Figuring out interactions between such entities resembles Ryle's "category mistake." However, there are special kinds of operations that may set the conditions for an indirect comparison between the domains ruled by the two types of models. These operations aim at defining entities of the *same category*, yet starting either from the high level or from the low level. These operations are, respectively, *analysis* and *synthesis*. Analysis is a top-down procedure starting from the higher-level system S and trying to figure out its parts (or its "components"). Synthesis is a bottom-up procedure which starts from "material elements" (whose concept is derived from a model of the basic physical level), and tries to elaborate a model of something like system S as the final outcome. In simple linear systems, the "components" obtained as a byproduct of analysis, and the material "elements" taken as starting points of synthesis, coincide. But in many other cases, this is not true. The model of the whole so to speak imposes a mark on its components, making them different from the alleged basic material elements. Downward causation is then defined as the *non*-coincidence of the models that apply to the components of a system with the models of its material elements; or as *the* non-*coincidence of the end product of the operation of analysis with the starting point of the operation of synthesis*.

Let me give an example, borrowed from quantum chemistry (Hladik 1971; Vemulapalli 2003; Bensaude-Vincent 2005). If one starts from a quantum model of an isolated atom of carbon, one can predict neither its valence nor the geometry of its possible chemical bonds. The model of electron distribution of the atom of carbon taken in isolation is indeed: $[1s^2, 2s^2, 2p_x^1, 2p_y^1, 2p_z^0]$ (two electrons on the layer 1 s, two electrons on the layer 2 s, one electron on each of the layers $2p_x$ and $2p_y$, and no electron on the layer $2p_z$). This electron distribution entails the *di*-valence of carbon (because only two electrons are unpaired: those on layers $2p_x$ and $2p_y$), and an angle between the two predicted bonds equal to 90°. But this is almost always wrong. Methane and other molecules including carbon display the *tetra*-valence of carbon atoms, and a tetrahedral geometry with an angle between the bonds close to 109°. In view of this huge discrepancy, quantum chemists then introduced the ad hoc trick of "orbital hybridization." The states of the two electrons on layer 2 s were linearly combined with the states of the electrons on $2p_x$ and $2p_y$, and one then obtained four unpaired electrons in hybridized states called sp^3. This finally accounted for the observed molecular structures. It is clear at this point that the physical model of the isolated carbon atom taken as the starting point of bottom-up synthesis does not coincide with the chemical model of carbon-components of molecules that is the end product of top-down analysis. The bottom-up *elements* ($2s^2$, $2p_x^1$, $2p_y^1$-isolated carbons) do not coincide with the top-down *components* (sp^3-hybridized carbons). This non-coincidence between bottom-up and top-down procedures and associated models accounts for what we may call a case of "downward causation," *beyond the situation of an immediate intervention.*

Here again, it is tempting to forget the whole cognitive procedure and jump to the claim that ontologically construed molecules have an influence on ontologically construed atoms. One could say, for instance, that molecules of methane downwardly influence carbon atoms by altering their electron distribution, changing it from the $2s^2$

$2p_x^1\ 2p_y^1$ configuration to the sp^3 configuration. But this shortcut immediately raises all the objections listed in Sect. 2. By contrast, mere non-coincidence of bottom-up and top-down experimental *procedures*, and of the two corresponding *models* of the lower level, is immune to these objections formulated in an ontological framework.

To sum up, according to the interventionist-constitutive view, there is a sense in which an event at a certain level L can be said to cause an event at the lower level l, even in the absence of a *present* intervention at level L. This requires only understanding how patterns of experimental interventions are used to formulate permanent law-like statements, and connecting various access-relative models.

9 Downward self-causation

A crucial question is still left over at this point. We have seen that downward causation makes perfectly good sense when an *external* intervention at a given organization level which is at the same time targeted and defined by it, induces alterations observed at a lower level. Now what about *self*-intervening and *self*-altering? This notion sounds nonsensical. Self-alteration looks like a verbal trick for expressing mere unobstructed evolution of a process. However, everybody knows cases of (apparent or real) self-interventions in which a voluntary change of state of mind has an influence on one's own physiology (Petitmengin et al. 2006).

The major sticking point concerns the word 'voluntary' in the expression 'voluntary change of one's own state of mind'. Are we *really* free to change our state of mind? More generally, are we *really* free to choose any one of our actions (Libet et al. 1999)? Even before examining this question, its underlying presupposition must be exposed. The very fact that the debate bears on the *reality* or *irreality* of free will is a further symptom of the dominance of a foundationalist framework. Those who argue in favor of the reality of free will tend to ground it on a reified conception of mind. And those who argue *against* the reality of free will tend to reify objective descriptions of physiology and behavior (Bitbol 2008a,b). The latter authors are prone to accept that our experience (especially our experience of free will) may be an illusion, whereas the law-like objective description of the functioning of our body exposes its *real* working. However, if Kant's "Copernican revolution" is pushed to its ultimate consequences, this kind of cut between experiential illusion and *intrinsic reality* as disclosed by inquiry appears to be an even deeper illusion, generated by reason's mistaking its own horizons of research for metaphysically real entities (Kant 1996).

An alternative pluralist and access-sensitive analysis of free will is then needed. It develops in two steps. Firstly, one dispels carefully the classical dualist view which is inadvertently associated with the phrase 'free will'; to wit the idea that free will means unruly action of a substantial mind *on* another substantial something. Secondly, one formulates a non-dualist alternative to this misleading metaphysical picture. An interesting account of this kind, adapted from Kant's *Critique of Practical Reason* with a touch of Spinozism, relies on the dialectic of the actor and the spectator (Beck 1963). The spontaneous belief in free will here arises from the standpoint of an *actor* immersed in nature, whereas denying free will is unavoidable from the standpoint of a *spectator* who undertakes to posit part of this nature as an object of description.

According to Kant (2002), even though there can be no "theoretical" proof of freedom (i.e., no proof from the standpoint of a spectator), things are different from the standpoint of the actor. Indeed, someone "(…) cannot *act* except *under the idea* of freedom" (p. 247); and this entails that *in some sense* one is "*really* free." One must insist at this point on the unusual but sensible reasoning of Kant:

(i) In order for the very process of *deliberation* to develop and make sense, actors must work under the unquestioned *presupposition* of free will. As a counter-example, let us imagine an actor who decides to do something, but suddenly realizes that, at any rate, what she will do next is already pre-determined. This implies a lived *performative contradiction* (Apel 1984) that, if (too) consistently taken into account, may result in inhibition of the decision and mere passivity.
(ii) This necessary *presupposition* of free will is enough to declare that actors are "really" free from a *practical* standpoint.

The fact that, from the standpoint of a scientific external *spectator* (or even from the standpoint of the same human being who has adopted retrospectively a spectator-like stance towards her own past actions and experiences), the behavior of the agent may appear to be governed by objective deterministic laws does not weaken in the least the practical *reality* of the presupposed freedom. For, here, using the adverb 'really' does not require reference to anything else than the spontaneously felt, and in practice indispensable, *presupposition* of deliberative action. This sense of the word 'reality' can be considered "weak" only if one forgets that, in a deliberately non-metaphysical view, no "stronger" *an sich* sense is available. In particular, the theoretical sense of the word 'reality' can hardly be considered stronger, since it refers only to phenomenal invariants used as practical guides for an intersubjectively coordinated research.

This can be called a "two-stance," as opposed to a "two-substance," account of free will: actor stance/spectator stance, instead of spiritual substance/material substance. On the basis of this account, one could say (in agreement with my initial formulation of the status of downward causation) that free will is neither illusory nor inherently existent, but merely indexical or situational. It holds in the situation of an actor, qua indexed by this situation.

In particular, there is indeed a sense in which one: (i) can voluntarily alter one's own state of mind, and (ii) observe the objective physiological "effects" of this voluntary "downward" alteration. No "disruption of the laws of nature by a spiritual substance" is needed for that. It has only to be accepted that step (i) expresses what is presupposed by oneself as an actor (and is then "real in practice"), whereas step (ii) expresses the outcome of an objective science developed by adopting the stance of spectators (and is then "real in theory").

Finally, one can think of yet another case which is rarely discussed: "indirect self-influence," beyond "direct self-causation." As pointed out by Elster (1983), when we wish to alter our own (psycho-physiological) state, it is often better to act indirectly by choosing to influence ourselves laterally, rather than to act directly by trying to cause the sought state. For instance, when we wish to sleep, it is better not to strive towards this aim, but rather influence ourselves by promoting other (apparently unrelated) processes, such as concentration on a repetitive mental task or renunciation of any aim whatsoever. The altered state thus arises as something that is "essentially" an indirect by-product of the promoted process, rather than being a direct effect of our desire to reach it.

It goes without saying that, just as causing, influencing by self-intervening at a certain level contributes to *defining* this level of organization. Accordingly, downward influence holds exactly in the same sense as downward causation by external intervention or by voluntary self-alteration: it holds qua indexed by a mode of access and a mode of (lateral) action.

10 Conclusion and summary

(1) Making sense of upward and downward causation does not require acceptance of some sort of metaphysical distinction between the higher and basic levels of organization: neither a substantial distinction as in genuine dualism, nor a distinction between properties and structures as in the current popular picture. It is enough to assume a duality of modes of access, or modes of intervention.
(2) If one intervenes at a higher level of organization, some effects of this action can then be detected by a mode of access specifically aimed at a lower level. This is downward causation. Conversely, if one intervenes at a microscopic level, some effects of this action can then be detected by a mode of access specifically aimed at a higher level of organization. This is upward causation.
(3) The modes of access and the modes of intervention are not just revealing; they are *constitutive* throughout. Accordingly, the levels of organization are not only disclosed but also *defined* by corresponding modes of intervention.
(4) Therefore, saying that some intervention at a higher level *downwardly causes* alterations detected at a lower level (or, conversely that some intervention at a lower level *upwardly causes* alterations detected at a higher level) is an accurate expression of a dual mode of *operational definition* of the levels.
(5) The dual mode of operational definition of the levels can be extended beyond the very moment of intervention, and beyond the simple case of external interventions, thus also making sense of cases of permanent inter-level influence or self-transformation.
(6) The emergent levels of organization, and the inter-level causations as well, are neither illusory nor ontologically real. They are *objective* in a transcendental, constitutive, sense.
(7) Alternatively, one can say that the usual dichotomy between epistemology and ontology collapses, because we can ascribe no other meaning to the "being" of levels of organization and causes than that which has been constituted and objectified by an epistemic method of access.

References

Apel, K. O. (1984). *Understanding and explanation: A transcendental-pragmatic perspective*. Cambridge, MA: MIT Press.

Atmanspacher, H., Filk, T., & Römer, H. (2004). Quantum Zeno features of bistable perception. *Biological Cybernetics, 90*, 33–40.

Batten, D. (2008, April 12–16). *Demystifying emergence*. Presentation at the conference emergence in the physical and biological world: A notion in search of clarification, Erice, Italy.

Beck, L. W. (1963). *A commentary on Kant's critique of practical reason*. Chicago: The University of Chicago Press.

Bedau, M. (2002). Downward causation and the autonomy of weak emergence. *Principia, 6*, 5–50.

Bedau, M. (2008a). Is weak emergence just in the mind. *Minds and Machines, 18*(4), 443–459.

Bedau, M. (2008b, April 12–16). *Pluralism about emergence*. Presentation at the conference emergence in the physical and biological world: A notion in search of clarification, Erice, Italy.

Bensaude-Vincent, B. (2005). *Faut-il avoir peur de la chimie*. Paris: Les Empêcheurs de Penser en Rond.

Bich, L. (2009) Downward causation and relatedness in emergent systems: Epistemological remarks. In G. Minati, E. Pessa, & M. Abram (Eds.), *Processes of emergence of systems and systemic properties: Towards a general theory of emergence* (pp. 591–602). Singapore: World Scientific.

Bickhard, M., & Campbell, D. T. (2000). Emergence. In P. B. Andersen, C. Emmeche, N. O. Finnemann, & P. V. Christiansen (Eds.), *Downward causation* (pp. 322–348). Aarhus: Aarhus University Press.

Bitbol, M. (1996). *Schrödinger's philosophy of quantum mechanics*. Dordrecht: Kluwer.

Bitbol, M. (1997). *Mécanique quantique, une introduction philosophique*. Paris: Flammarion.

Bitbol, M. (1998). Some steps towards a transcendental deduction of quantum mechanics. *Philosophia Naturalis, 35*, 253–280.

Bitbol, M. (2000). Physique quantique et cognition. *Revue Internationale de Philosophie, 54*, 299–328.

Bitbol, M. (2001). Non-representationnalist theories of cognition and quantum mechanics. *SATS (Nordic Journal of Philosophy), 2*, 37–61.

Bitbol, M. (2007a). Ontology, matter and emergence. *Phenomenology and the Cognitive Science, 6*, 293–307.

Bitbol, M. (2007b). Materialism, stances, and open-mindedness. In B. Monton (Ed.), *Images of empiricism: Essays on science and stances, with a reply from Bas C. van Fraassen* (pp. 229–271). Oxford: Oxford University Press.

Bitbol, M. (2008a). Is consciousness primary. *NeuroQuantology, 6*, 53–71.

Bitbol, M. (2008b). Consciousness, situations, and the measurement problem of quantum mechanics. *NeuroQuantology, 6*, 203–213.

Bitbol, M. (Ed.). (2009). *Théorie quantique et sciences humaines* (CNRS éditions).

Bitbol, M. (2010). *De l'intérieur du monde, Pour une philosophie des relations*. Paris: Flammarion.

Campbell, R. J., & Bickhard, M. H. (2009). Physicalism, emergence, and downward causation. *Synthese* (in press).

Castellani, E. (2002). Reductionism, emergence, and effective field theories. *Studies in History and Philosophy of Modern Physics, 33*, 251–267.

Cohen-Tannoudji, C., Diu, B., & Laloe, F. (2006). *Quantum mechanics*. New York: Wiley Interscience.

Deacon, T. W. (2003). The hierarchic logic of emergence: Untangling the interdependence of evolution and self-organization. In B. Weber & D. Depew (Eds.), *Evolution and learning: The Baldwin effect reconsidered* (pp. 273–308). Cambridge: MIT Press.

D'Espagnat, B. (1989). *Conceptual foundations of quantum mechanics*. New York: Addison-Wesley.

Ducheyne, S. (2006). Galileo's interventionist notion of 'cause'. *The Journal of the History of Ideas, 67*, 443–464.

Elster, J. (1983). *Sour grapes*. Cambridge: Cambridge University Press.

Emmeche, C., Köppe, S., & Stjernfelt, F. (2000). Levels, emergence, and three versions of downward causation. In P. B. Andersen, C. Emmeche, N. O. Finneman, & P. V. Christiansen (Eds.), *Downward causation, minds, bodies, and matter* (pp. 13–34). Aarhus: Aarhus University Press.

Galilei, G. (1612). *Discorso intorno alle cose che stanno in su l'aqua o che in quella si muovono In Opere I*. (Torino: U.T.E.T).

Geiser, F., & Halbrecht, U. (2008). *Fear that freezes blood in your veins*. Retrieved 20 August 2009, from http://www.uni-bonn.de/Press-releases/110_2008/.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Gillies, D. (2005). An action-related theory of causality. *British Journal for the Philosophy of Science, 56*, 823–842.

Hartmann, S. (2001). Effective field theories: Reductionism and scientific explanation. *Studies in History and Philosophy of Modern Physics, 32*, 267–304.

Hladik, J. (1971). *Elements de chimie quantique*. Paris: Dunod.

Hoffmeyer, J. (1997). Biosemiotics: Towards a new synthesis in biology. *European Journal for Semiotic Studies, 9*, 355–376.
Humphreys, P. (1997). How properties emerge. *Philosophy of Science, 64*, 1–17.
Kant, I. (1955). *Prolegomena to any future metaphysics*. Chicago: Open Court.
Kant, I. (1987). *Critique of judgment*. Indianapolis: Hackett Publishing Company.
Kant, I. (1992). Concerning the ultimate ground of the differentiation of directions in space. In D. Walford & R. Meerbote (Eds.), *The Cambridge edition of the works of Immanuel Kant. Theoretical philosophy 1755–1770* (pp. 361–372). Cambridge: Cambridge University Press.
Kant, I. (1996). *Critique of pure reason*. Indianapolis: Hackett Publishing Company.
Kant, I. (2002). *Groundwork for the metaphysics of morals*. Oxford: Oxford University Press.
Kim, J. (1999). Making sense of emergence. *Philosophical Studies, 95*, 3–36.
Kistler, M. (1999). *Causalité et lois de la nature*. Paris: Vrin.
La Mura, P. (2005). Correlated equilibrium of classical strategic game with quantum signals. *International Journal of Quantum Information, 3*, 183–188.
Lambert, A., Zamir, S., & Zwirn, H. (2006). *Type indeterminacy: A model of the KT (Kahneman–Tversky)-man*. New York: Cornell University Library.
Levinas, E. (2001). *En découvrant l'existence avec Husserl et Heidegger*. Paris: Vrin.
Libet, B., Freeman, A., & Sutherland, J. K. B. (Eds.). (1999). *The volitional brain: Towards a neuroscience of free will*. Exeter: Imprint Academic.
Maturana, H. R., & Varela, F. J. (1980). *Autopoiesis and cognition*. Dordrecht: Reidel.
Maudlin, T. (1998). Part and whole in quantum mechanics. In E. Castellani (Ed.), *Interpreting bodies* (pp. 46–60). Princeton: Princeton University Press.
Menzies, P., & Price, H. (1993). Causation as a secondary quality. *British Journal for the Philosophy of Science, 44*, 187–204.
Mill, J. S. (1851). *System of logic* (Vol. 1). London: John W. Parker.
Pattee, H. H. (2000). Causation, control, and the evolution of complexity. In P. B. Andersen, C. Emmeche, N. O. Finneman, & P. V. Christiansen (Eds.), *Downward causation, minds, bodies, and matter* (pp. 63–77). Aarhus: Aarhus University Press.
Petitmengin, C., Navarro, V., & Baulac, M. (2006). Seizure anticipation: Are neuro-phenomenological approaches able to detect preictal symptoms. *Epilepsy and Behavior, 9*, 298–306.
Price, H. (1992). Agency and causal asymmetry. *Mind, 101*, 501–520.
Rosen, R. (1991). *Life itself: A comprehensive inquiry into the nature, origin, and fabrication of life*. New York: Columbia University Press.
Salmon, W. (1984). *Scientific explanation and the causal structure of the world*. Princeton: Princeton University Press.
Schrödinger, E. (1935/1983). The present situation of quantum mechanics. In J. A. Wheeler & W. H. Zurek (Eds.), *Quantum theory and measurement* (pp. 152–167). Princeton: Princeton University Press.
Seevinck, M. P. (2003). Holism, physical theories and quantum mechanics. In *International workshop on Holism in the philosophy of physics*, Bonn, 4–5 July 2003. Retrieved 12 August 2009, from http://philsci-archive.pitt.edu/archive/00002191/.
Stephan, A. (1998). Varieties of emergence in artificial and natural systems. *Zeitschrift für Naturforschung, 53*, 639–656.
Teller, P. (1995). *An interpretive introduction to quantum field theory*. Princeton: Princeton University Press.
Varela, F. (1976). Not one, not two. *The Coevolution Quarterly, 11*(Fall), 62–67.
Vemulapalli, G. K. (2003). Property reduction in chemistry. *Annals of the New York Academy of Sciences, 988*, 90–98.
Von Wright, G. H. (1974). *Causality and determinism*. New York: Columbia University Press.
Weber, A., & Varela, F. J. (2002). Life after Kant. *Phenomenology and the Cognitive Sciences, 1*, 97–125.
Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford: Oxford University Press.