

Where to look for emergent properties

Agustin Vicente

1. Ikerbasque: Basque Foundation for Science, Alameda Urquijo, 36-5, 48011,

Bilbao, Spain

2. University of the Basque Country. Linguistics and Basque Studies Department.

Universidad Ibilbidea, 5, 01006 Vitoria, Spain.

e-mail: agustin.vicente@ehu.es

Abstract

Recent years have seen renewed interest in the emergence issue. The contemporary debate, in contrast with that of past times, has to do not so much with the mind–body problem as with the relationship between the physical and other domains; mostly with the biological domain. One of the main sources of this renewed interest is the study of complex and, in general, far-from-equilibrium self-preserving systems, which seem to fulfill one of the necessary conditions for an entity to be emergent; namely, that its causal powers are not predictable from the causal powers of basic physical properties. However, I argue that much of the current emergentism debate has misfired by focusing on the interpretation of self-maintaining systems. In contrast I claim that if we want to find emergent properties, we should not look at complex systems, but at selection (natural selection, in particular). I argue that selection processes make the causal world “exuberant” by making non-physical functional and relational properties enter the causal web of the world.

1. Introduction

Recent years have seen renewed interest in the emergence issue. The contemporary debate, in contrast with that of past times, has to do not so much with the mind–body problem as with the relationship between the physical and other domains; mostly with the biological domain. Philosophers discuss whether there are proper (i.e. non-reducible) non-physical properties or whether, as reductive physicalists claim, everything is reducible to the physical. One of the main sources of the renewed interest in emergentism is the study of complex and, in general, far-from-equilibrium (FFE) self-preserving systems. FFE systems are intriguing from a philosophical perspective because it seems as though knowledge of their constituents and of their causal powers does not suffice to predict the behavior of the whole system. Thus, FFE systems seem to fulfill one of the necessary conditions for an entity to display emergent properties, namely, that its causal powers are neither predictable nor explainable from knowledge of the causal powers of basic physical properties. A hurricane, for instance, is a self-preserving structure composed of myriad particles that behave in a very peculiar way, clearly constrained by the global pattern or structure to which they belong. As it moves, the hurricane leaves behind some of its constituents and replaces them with others taken from the surrounding matter, forcing them to behave in an uncommon way so that the global pattern is preserved. The point here is that it seems that the (peculiar or uncommon) behavior of the particles that form part of a hurricane cannot be explained on the basis of our knowledge of their fundamental physical properties, basically because their behavior is constrained by their being part of a whole. So, it seems that the whole exerts a causal influence on its

own constituents that can only be imputed to the whole. Thus, the whole seems to have non-derived, new, causal powers. Note that this view of emergence and downward causation¹ (from wholes to constituents) paints a picture in which the relationship between emergent properties and basal properties is apparently *vertical*, i.e., emergent properties apparently cause changes in their own basal conditions. The picture of downward causation that Jaegwon Kim (1999, 2006) discusses, in contrast, is a picture in which emergent properties cause changes at lower levels *diagonally*. That is, if property *E* emerges from property *F* (its basal conditions), the instantiation of *E* causes the instantiation of *F'*, with $F' \neq F$. In short, the causal powers of emergent phenomena are not directed at their own basal conditions.²

The more complex an FFE system is, the more we are tempted to see emergent properties in it. A hurricane is a very simple constraining system. Living systems, in contrast, involve various nested and interrelated constraints. A eukaryotic cell is a large complex constraint consisting of several components whose behavior constrains the behavior of the other components and also of their own constituents; the final effect is the self-maintenance, self-containment and self-organization of the cell. However, intriguing and perhaps impressive as self-maintaining FFE systems may seem, reductionists are not convinced. It is one thing to say that we are unable to explain a hurricane, a traffic jam or the appearance of life in a bottom-up way; it is quite another to claim that all these phenomena are unexplainable *in principle*. We simply cannot know whether enough knowledge would enable us to explain the

¹ 'Downward causation', a term introduced by Campbell (1974), refers to a causal relation between a higher-level property or system and a lower-level one.

² This is because Kim is mostly interested in mental causation, where the issue is whether or not mental states can be said to produce either bodily movements or the physical states that subvene other mental states. Defenders of vertical downward causation, in contrast, focus on the relationship between wholes

formation and duration of a hurricane. Is it not more sensible to believe that the bottom-up explanations *are there*, even if we cannot reach them? Moreover, where can the new causal powers stem from? If everything is made of the same fundamental entities, how can it be that the behavior of all things is not due ultimately to the behavior of these fundamental entities? Emergentists do not have a clear answer to these questions. In the wake of some suggestions made by British emergentists (see Gillett, 2006, on Alexander), and especially following some of Roger Sperry's (1969) and Donald Campbell's (1974) insights, some contemporary emergentists maintain that wholes do not bring about new causal powers but are able to constrain the causal powers of fundamental entities. Somehow, wholes determine the behavior of their component parts (Viera and El-Hani, 2008), for instance by reducing their degrees of freedom³. Some call this "formal causation" (El-Hani and Emmeche, 2000). The idea, whatever we call it, is that we cannot fully explain the behavior of the constituents of a self-maintaining system unless we take into account their belonging to the system. Their behavior is certainly different when they are inside the system from when they are outside it.

However, there is no end in sight for this dispute between emergentists and reductionists. Reductionists tend to claim that everything that is accounted for in terms of wholes can, as a matter of principle, be accounted for in terms of the many interrelations that exist among their constituents and between these and the surrounding matter (see Davies, 2006, Mossio *et al.*, forth.). Emergentists, in turn, tend to insist on the peculiar behavior of certain components within certain systems,

and their constituent parts.

³ Wilson (2010) makes a distinction between reducing, restricting and eliminating degrees of freedom.

thus trying to elicit emergentist intuitions. Those intuitions, however, are far from clear. What is needed is an argument; which is lacking in the current debate. A radical empiricist approach to the philosophy of science could provide such an argument: if we lack the relevant positive evidence that complex systems can be accounted for in terms of fundamental physics, then we are not justified in believing that complex systems can be explained by fundamental physics.⁴ However, emergentists cannot hope that such an attitude will be shared by their opponents. Rather, opponents will insist that it is possible to distinguish between what *we* can explain and what can be explained *in principle*; their position being that complex systems have a physicalist explanation *in principle*.

Now, I think it is possible to show that there are certain causal explanations that cannot be substituted by physicalist explanations, even as a matter of principle. However, I also think that much of current emergentism has misfired by being focused on the interpretation of FFE self-maintaining systems. If we want to find emergent properties, we should look not at the processes that lie at the heart of complexity, but at selection processes. It is possible to find some indications pointing towards such a position in Campbell's (1974) seminal work on downward causation. While Campbell is usually presented as one of the precursors of the idea that higher-level phenomena have a "vertical" influence on the behavior of their own constituents (such that higher-level phenomena constrain the behavior of their constituents), his

According to this distinction, what goes on in complex systems is not so much a reduction of degrees of freedom, but an elimination of them.

⁴ This radical empiricist attitude is best exemplified by Cartwright (1999). Cartwright argues that, given that we do not have any evidence that Newtonian physics can explain the fall of a dollar bill in the open air, we can doubt that Newtonian physics applies except in very controlled settings. By parity of reasoning, we could say that if there is no evidence for a physicalist explanation of complex systems, we had better believe that physicalist explanation does not account for complex systems. Dupré (2001)

discussion of downward causation focuses on the role of natural selection. Thus, I would agree with what he says in the following quote, which serves to introduce the notion of downward causation for the first time (Campbell, 1974; 180): “Where natural selection operates through life and death at a higher level of organization, the laws of the higher-level selective system determine in part the distribution of lower-level events and substances”.⁵

Let me now begin to explain why I hold that we should focus on processes of selection. Typically, a process of selection involves a mechanism of selection’s “choosing” an entity due to its instantiating at least one interesting property. Mechanisms of selection, however, are not constrained to choosing entities only on the basis of their physical or fundamental properties. Their target can be, and usually is, properties that, *prima facie*, would not count as physical properties, such as some functional or relational properties.⁶ My claim, thus, is that in selection processes these functional or relational properties can be causally efficacious, which means that properties other than physical properties can have causal powers. To put it in the form

is another example of this “empiricist” attitude.

⁵ I may be wrong in pointing out that Campbell’s reflections on the role of natural selection somehow bring him close to my cause, and I do not want to suggest at all that the “orthodox” interpretation of Campbell is defective. I am content with claiming that some of the points he makes are points I would endorse. Here is another interesting quote (Campbell, 1990, 5): “From the point of view advocated here, these macrostructures [perhaps all of them understandable in terms of the science of microparticles] provide the raw material of variations on which biological natural selection can operate. They are, however, inadequate to predict or explain which variant organic forms will be most abundant”.

⁶ As we lack a proper account of what it is for an entity to be physical, this is deemed to be controversial. From here on, I will assume that to be physical is to be a posit of physics, and that the functional and relational properties that I consider are not (and will not be) posits of physics. I think that it is possible to argue for this construal of the physical (see Vicente, 2011b). I also think that this is the notion of the physical involved in the physicalist/emergentist debate (or, at any rate, in the physicalist/emergentist debate as it concerns the present paper). Other approaches, such as the “via negativa” (see Montero and Papineau, 2005) or Crook and Gillett’s (2001) account, tend to blur the terms of the debate, as biological properties (and the properties considered here) turn out to be physical (since “the physical” is very broadly construed).

of a slogan: selection processes make the causal world exuberant.⁷

Emergentism claims that there are more causally powerful properties in the world than physical properties. Moreover, emergentism typically claims that such properties can produce changes in the physical world, that is, that they have powers for downward causation, and that they "arise" when the physical world is sufficiently complex. I will argue that if we want to find properties that fulfill these requirements, we should take a close look at what goes on in selection processes; when the world is sufficiently complex for it to contain selection mechanisms, there emerge new causal properties.

The basic argument will be as follows. When trying to account for what goes on in selection processes, we have to resort to non-substitutable explanations that mention properties other than physical properties, such as some functional or relational properties. There are also physical explanations that account for the end-products of selection processes, but such explanations can never substitute the non-physical explanations. Next, in such cases, both explanations (physical and non-physical) are well-supported and satisfactory. Although *prima facie* this means that we should choose one of them and reject the other, the tension required to apply this so-called 'exclusion principle' does not arise. Finally, that tension is lacking at the

⁷ By this I only mean that selection processes have an effect on which properties have causal powers. A world without selection processes has fewer causally efficacious properties than one with selection processes. If we maintain, as many metaphysicians do, that "to be is to have causal powers", then a world with selection processes is exuberant with respect to a world without them—a world with selection processes, for instance, has functional properties, while in a world without selection all that we find are functional concepts (on this distinction, see, e.g. Kim, 1998, Gillett, 2003). I am aware that the slogan may be misleading in that it suggests that it is selection (not the processes that generate the variability on which selection operates) that is responsible for the manifest exuberance of the world. I am not claiming that at all. I hope this footnote serves to avoid such a misreading.

level of explanations and neither does it arise at the metaphysical level, i.e., at the level of causes. The conclusion is that the predicates/concepts used in these non-physical explanations denote causally efficacious properties. As the *explananda* of these explanations can be considered physical effects, or at any rate to involve physical effects, the upshot is that some non-physical properties have powers for “downward” causation.

The paper develops along the following lines. In the following section, I argue for the idea that explaining what goes on in selection processes involves resorting to non-physical emergent properties. Then in section 3, I contend that the information provided by such explanations cannot be captured in any other way. Furthermore, I maintain that we lack the feeling of tension that would be required to force us to choose between these emergence explanations and the alternative physical explanations; or between the causes emergentism postulates and the causes postulated by physical explanations. Finally, in section 4, I face and discuss the perennial problem with emergentist proposals, namely, how to reconcile their claims with the so-called principle of the causal closure of the physical world. I argue that my proposal is compatible with the truth of the principle, though it certainly clashes with the idea that if you have a physical explanation of a physical effect, then there is nothing left to be known about the etiology of that effect. However, it is possible to refute this idea without denying either the veracity of the principle itself or any other claim that we take to be true.

2. Selection processes and their role in emergence

Selection processes require three components: a selection mechanism, a selected particular, and (in the simplest case) a property instantiated by the selected particular, by virtue of which it is selected. Thus, a clockmaker is a selection mechanism, a cogwheel is a selected particular, and the property by virtue of which the cogwheel is selected is, in this case, a dispositional property, namely, the property of being able to transmit a particular movement. As it is well known, the natural world contains selection mechanisms, which are usually grouped together under the rubric ‘natural selection’ (NS). It may be, though, that NS is not the only selection mechanism in the world: it is plausible that there are pre-NS selection mechanisms that account for the existence of relatively simple self-sustaining wholes (see Moreno and Ruiz-Mirazo, 2009).

Be that as it may, it seems plausible to claim (again, considering simple cases), that when an entity is selected, the cause of its being selected is its instantiating a particular property. In contrast, it is just as plausible (*prima facie* at least) to hold that, generally, the properties that are involved in selection processes are not fundamental physical properties. As in the case of the clockmaker and the cogwheel, many of the properties are dispositional or functional properties. Thus, hearts were not selected because they had certain physical or categorical properties, but because they were blood pumps. In fact, it seems that an entity’s acquiring a (teleo)function (and thus normativity), consists of some selection mechanism selecting it by virtue of its instantiating a particular functional property.⁸ This is well exemplified by Wright’s (1973) definition of the notion of teleofunction: x has the function to F if and only if

⁸ Functional properties are properties defined by their causal roles. Teleofunctional properties are properties defined by what they *ought* to do.

(a) x F s, and (b) x is there because it F s. Wright's definition has been the target of many counterexamples, but all the counterexamples affect the sufficiency part of the claim: no counterexample has shown that definition does not provide a necessary condition for an entity to have a function. Thus, Wright seems to be right in claiming that functional properties are causally relevant in the process of an entity's acquiring a teleofunction (which means that the entity has been the object of a selection process).

So, the recruitment or selection process by which a certain entity acquires a certain teleofunction is a process where the cause is the entity's instantiating a functional property and the effect is the entity's being there. However, this is by no means the only case where we can say that we find a non-physical event acting as a putative cause of an ultimately physical effect. Evolutionary accounts of teleofunctionality claim that NS is the only source of teleology (see Millikan, 1989, Neander, 1991). This is perhaps disputable (see, e.g., Mossio *et al.*, 2009). But what does not seem to be open to dispute is that the role of NS in bringing about new causal properties is not restricted to functional properties. All sorts of relational properties acquire causal efficacy thanks to NS. For the main claim of Darwinian theory is precisely that the organisms that do better than their "rivals" in some respect have a greater chance of surviving and reproducing successfully. An organism that has a better sense of orientation than the rest of its co-specifics will be able to get more food for itself and its progeny, thus ensuring that both have more chances of enduring. A member of a species that is able to eat more kinds of food than the other members of that species will be better off in the event of persistent draught. And so on and so forth. In all these cases, we find relational properties acting as putative causes. It is the relation between the organism and its rivals (or alternatively, the relation of the

organism to the total environment) what explains its greater rates of survival and reproduction.⁹

I can illustrate this point by using a toy example which I assume will be familiar to all and I hope will make matters simple. Think of a lion—a predator—as a selection mechanism. The lion silently approaches a herd of antelopes, lies in wait and, when things settle in a particular way, starts its hunt. It makes the herd move, identifies some easy prey and tries to isolate it. After a while—if things go well—it is able to capture and kill it. Now we may ask: why did the antelope die? A reasonable answer is: because it was the weakest member of the herd. We have thereby invoked a relational property in a causal explanation. Being the weakest member of the herd seems to have causal powers; powers that are not reducible to the powers that any physical property that the antelope (or of any other entity) may have. That, at any rate, is what I want to argue.

To clarify: I am not trying to contrast the relational with the intrinsic; neither am I assuming that all physical properties—or at any rate, all the physical properties of the antelope—are intrinsic. My claim here is that the relational properties I am talking about are not properties that form part of the repertoire of current physics and will presumably not form part of any future physics.¹⁰ By being part of physics, in turn, I

⁹ I have preferred not to talk about fitness because of the many problems associated with the notion (see, e.g., Rosenberg and Bouchard, 2010). Depending on how fitness is understood, it may make no sense to hold that *a's being fitter than b* has causal powers. Whereas it does make sense, I think, to claim that *a's having a better orientation sense than b*, or *a's being weaker than b*—which will be my toy example—have causal powers.

¹⁰ I am aware that speaking about future physics is problematic. Perhaps I could restrict the meaning of “the physical” to that which is postulated by current physics and construe my claim as a claim indexed to the present time and physics. That is, my claim could be that, given current physics, properties such as being the weakest member of a herd are emergent properties. However, I suspect that this kind of predicate will not appear in the causal antecedent of any law of physics, current or future. One reason to think this is that physics already has explanations for the events that these properties explain—see below.

mean being a posit invoked by some of the laws physics holds to be true. There is no law of physics, and in particular, no causal law of physics, which speaks about weakest members of herds.

3. The case for the causal efficacy of non-physical properties

3.1. Physical causal explanations cannot explain everything

It does seem as though non-physical properties play a causal role in selection processes. Moreover, it seems that they can have “downward” causal powers, as the effects they bring about (such as an entity’s being there, the death of an individual, or the persistence of a certain kind of individuals) are, ultimately, physical effects or involve physical effects. So, *prima facie*, these properties are what could be called “classical” emergent properties. Kim (1999, 2006) famously argues that for a property to be emergent it has to fulfill three conditions: (a) being a "new" property which arises from certain configurations of fundamental physical properties and supervenes on them; (b) not being identical to the physical properties that realize it; (c) nonetheless producing effects in the physical world. The non-physical properties I have considered seem to fulfill all these conditions, except that they typically do not supervene on intrinsic physical properties of individuals, but on a wider base; but this is no objection to their being emergent properties.

However, the perennial and key question for the emergence issue now arises: how could these properties have causal efficacy, if all the effects they explain have a physical cause? The heart that is allegedly there due to its being a blood pump is there as a result of a physical process, a long chain of interactions between purely physical

events. The antelope that dies, apparently as a result of being the weakest member of the herd, dies because it loses a lot of blood due to the lion's having perforated its jugular vein, a process which, in principle, could have a physical description. So, how can it be that these events, which have a physical explanation, have a non-physical cause as well?

The response that I argue for is the following: exclusion problems, which is what we are dealing with here, do not affect all kinds of pairs of causal explanations. Kim (1993) famously claims that there is a "principle of causal/explanatory exclusion" which tells us that we cannot have two complete and independent causes/causal explanations for one event (except in those rare cases where there is causal overdetermination). The justification for adopting this principle is that we feel a tension whenever we have two explanations that explain the same thing. Now, I argue that we do not feel that kind of tension in all circumstances and that in particular, we do not feel that kind of tension when we consider a non-physicalist and a physicalist explanation of, e.g., why the antelope died. We do not see the two explanations as competing, one against the other. Each explanation sheds different light on the occurrence to be explained and each of them explains, apart from that occurrence, things that the other explanation does not account for. So, this case is not like the case, for instance, of psychological versus physical explanations, where there seems to be general consensus that the two explanations are in competition. Now, I think that the same line of argument can be applied if instead of talking about causal explanations we talk about causes. In fact, it seems that if we do not see a problem in accepting two causal explanations, each of which we regard as complete, for an event,

then we should not see a problem in accepting that the event has two causes. This is not extracting metaphysical consequences from claims about explanation. Rather, it is just a reflection on what endorsing a causal explanation implies; namely, that you are committed to considering that the property you mention in the *explanans* is a cause of the *explanandum*. So, if, for whatever reason, you endorse two complete causal explanations for one event, then you are thereby committed to considering that the event has two causes. I return to this later on.

The task now is to show that some non-physical explanations do provide causal information that no physical explanation could provide. This is the task that contemporary emergentists who focus on complex systems and dissipative structures fail to meet convincingly enough. They tell us that the behavior of the constituents that belong to such a system can only be explained on account of their being parts of the system. But physicalists reply that there is an adequate thoroughly physical explanation for the phenomena, only that it is not available to us: everything that goes on in these systems and their surroundings is due to the action of basic forces on basic particles. That is, all the causal information that a higher-level explanation gives us could be captured by a fundamental-level explanation, and the fundamental-level explanation would explain the higher-level explanation. As mentioned above, emergentists lack a clear-cut answer to this kind of reply.

When it comes to selection processes, things are quite different. Let us take the case of the death of the antelope. An apparently satisfactory causal explanation of this event is that there is a selection mechanism, the lion, which, we could say, makes its choices as the result of a trade-off between amount of food, and risk and effort.

Weighing up options may lead it to focus its hunt on the weakest antelope in the herd; that is, on an antelope that bears a certain relation with the rest of the members of the herd: the relation of being weaker than them. Perhaps the process is somewhat different, more interactive, so that the lion does not weight many options, but starts running and leaves the rest to the interaction between its own movements and the movements of the herd. However, whatever the right detailed explanation is, it turns out that the property of being weaker than the rest plays an essential role in the explanation of the antelope's death: it is, in the end, the reason why the antelope performs worse than the others, and so the reason why its chances of surviving the lion's attack are fewer. Now, the key question to be asked is: can we come up with an explanation that captures all the information contained in this kind of explanation that does not mention *any* relational property? The answer, I claim, must be negative.

Of course there is another kind of causal explanation of the antelope's death, an explanation which is more physicalist friendly. The death is due to an enormous loss of blood caused by a hole opened in the jugular vein; in turn, the hole is opened by a tooth with a certain shape which pierces the antelope's body with certain force, due to the movements of the lion. The reason why the initial impact is produced lies in the speed and the trajectory of the lion and the speed and trajectory of the antelope. In turn, the movements of the lion and of the antelope are triggered by something that goes on in their brains. The lion's brain sets its body in motion on recognition of a certain scenario, and the antelope's brain triggers its fleeing movements after seeing the lion's approach. This kind of explanation is, very plausibly, reducible to an explanation couched only in physical terms. It can even be considered a "complete" explanation, at least in the sense that it does not require additional information *per se*.

It is a self-sufficient explanation. Yet the explanation does not capture all that has to be said about the *explanandum*. The issue can be put either in terms of modal information or in terms of what God would know or fail to know.

The antelope died in a certain way. However, it could have died in very many other ways: the impact of the lion's tooth could have been in a different place; the antelope could have covered a longer or a shorter distance; the loss of blood could have been the result of not just one bite but of several injuries caused by teeth and also by claws; the antelope could have died from a heart attack; the lion's brain could have set the lion in motion at the sight of a very different scenario, etc. But the important thing, it seems, is that the antelope would have died all the same *insofar as it was the weakest member of its herd*. That is, in all the possible worlds where that particular antelope is the weakest member of its herd, it is highly plausible that, keeping the circumstances alike, it ends up dying. In contrast, in the worlds where the antelope is not the weakest in the herd, it has a far greater chance of surviving. In a nutshell, the relational property *being the weakest in the herd* seems to effect a sharp division in the sphere of possible worlds: on the one hand, we have the worlds where the antelope dies (those worlds where the antelope is the weakest in its herd); and, on the other, the worlds where it survives (those worlds where it is not the weakest in the herd). This kind of modal information will never be available to the physicalist-friendly explanation I sketched before.¹¹ So, the explanation cannot be subsumed, and

¹¹ One can respond that the ultimate cause of the antelope's death is that the lion *represented* it as being the weakest of the herd, and that an explanation that appeals to this fact would not miss the information I have been talking about, while still being reducible to a physical explanation (the lion's representing the antelope as being so is, after all, a brain event). However, there is no need for the lion to represent the antelope as being the weakest member of the herd or to represent it as bearing any other kind of relation to the rest of the group. It can represent it just as being easy prey, or possible prey, or even not represent it at all—it may be that some anti-representational dynamical approach to lions' hunting is

substituted, by the physical explanation.

We can use an analogy to stress the point: there is some debate as to whether determinables can be causally efficacious properties. For instance, Steven Yablo (1992) thinks they can, while Jackson and Pettit (1990) hold that only ultimate determinates have causal efficacy. However, both sides to the debate concede that a true explanation which resorts to determinables provides modal information that a true explanation couched in determinate terms cannot capture.¹² Thus, suppose that a pigeon pecks at red things, and we find it pecking at a scarlet circle: we know that it would have pecked at that circle even if it had been of a different shade of red. That is, we know that it would have pecked at the circle in all the worlds where the circle had been colored any tonality of red. So, the determinable delineates the sphere of worlds where we find the effect occurring. The alternative explanation, i.e., the explanation that cites the “scarletness” of the circle, inevitably misses this modal information. *Mutatis mutandis*, the explanation that tells us the details of the antelope’s death, or even an alternative explanation that abstracts somewhat from such details, inevitably misses the modal information provided by the explanation which attributes causal power to the property of being the weakest in the herd.

The issue can be approached in a different way. In order to know whether a

the correct one. That is, the correlation between instantiating the property of being the weakest of a herd and falling prey to a predator can be mediated by various mechanisms; and it is not necessary that part of the mediating mechanism is a detector of the weakest members of herds. So, my point stands: a physicalist-friendly explanation will plausibly miss some causal information, because (a) it will not invoke relational properties at any stage and (b) relational properties provide causal information. However, let me note one thing: I have chosen the example of the lion and the antelope for ease of exposition. If the fact that the lion is a representational device is judged to muddy the waters, the example could be substituted for one in which predators are not involved (see section 2 for some such examples).

¹² To me, this means that determinables are indeed causally efficacious.

particular explanation e captures all the causal information to be captured about an effect f , it is useful to ask whether God would know everything there is to be known about f when entertaining e . So, suppose God entertains the complete physical explanation of the antelope's death, that is, entertains the minute physical description of the whole chain of causes and effects that ended with the antelope's death. It seems that, no matter how detailed this knowledge is, God would be missing an important, crucial, piece of information. In particular, it seems, God would not know why all that process took place: ultimately, why the antelope was chased and killed. Lacking that explanation, it would be impossible to make adequate predictions of future deaths. For instance, God could mistakenly predict that, when hungry, the lion will chase and kill all the antelopes that have the same intrinsic properties as the one that was actually killed. But the lion will not do that: *ceteris paribus*, it will chase and kill such antelopes only if they happen to be the weakest of their respective herds.

I want to stress that the relational property is indispensable. One could think that there is no need to mention the relation in the causal explanation—i.e., the relation could be seen as superfluous—mentioning the *relata* can be enough. It is such-and-such intrinsic properties in such-and-such a context that cause the effect, it could be said. In our case: the relational property relates the weakness of an individual to the weakness of all the individuals in the herd.¹³ We could try to explain the death of the antelope by using the weakness of the rest of the individuals as the background context and identifying the weakness of the antelope as the crucial causal factor, i.e.

¹³ 'Weak' is a gradable adjective which requires a comparison class and a standard. Suppose, for the time being, that one could talk about something such as "intrinsic weakness".

the one that makes a difference—given the context. That is, we could claim that the reason why the antelope was killed is that it was *this* weak. Being this weak in this context means being killed.

However, this will not do. The information we are after is: which worlds does the antelope die in (modal variant), and: what knowledge does God require in order to know everything that there is to be known about the causes of death of the antelope (God variant). There is a range of contexts where the antelope dies, but there is a limit: if the context includes an animal, or several animals, that are weaker than the antelope, the antelope will not die. Similarly, there is a range of weaknesses that the antelope may instantiate and still die, but again there is a limit: if the antelope is stronger than others, it will not die. It seems essential to know the relation between the weakness of the antelope and that of the rest of the herd in order to know in which conditions the antelope is going to die (keeping fixed the lion, the hole in the jugular, and so on). This kind of information is lost if the only thing we have available is context, on the one hand, and intrinsic properties, on the other. In such a case, we could have a list of pairs (context, intrinsic weakness): context c , intrinsic weakness w : death; context c' , intrinsic weakness w : survival, etc. But we want to know, if possible, what grounds, or explains, that very list. And it seems that the list does have a rationale: in a given context the weakness of the antelope results in its death as long as the antelope is weaker than the individuals that form part of the context. That is, we can explain why in certain contexts certain properties are related to certain effects. So it seems that the explanation that mentions the relational property cannot be substituted by another which mentions only intrinsic properties—and leaves the context as a background factor—without loss of information.

So, however we approach the issue, it seems that there is, in effect, some relevant causal information that an explanation which does not invoke the relational property cannot capture. The lesson to be extracted is that we have to know how certain entities are related to others in order to have some important information about why some things happen to them, for instance, about why they do or do not endure. Something along the same lines can be claimed about the causal information that functional properties such as *being a blood pump* provides (see Bird, 2007, Vicente, 2011a). Hearts are there because they are blood pumps. They are there also because there is a long chain of local relations of causation whose final output is the world as it is. But the second explanation cannot substitute the first. We might look at the world and be able to track its causal history. We would thus see that some types of things have gone out of existence while others, like hearts, have replicated themselves and endured. We would also see that some structures start doing things that they did not do before. We might as well think that there is nothing else to be explained: physical laws just drive the world and its entities in one definite direction. Yet, the fact is that there is something else to be known. We can find a more illuminative explanation of why some things remain and other things disappear. And such an explanation is couched in functional terms: entities which can perform certain functions tend to stay. Entities whose causal powers are not of interest to selection or recruitment processes tend to disappear. And many other entities do what they do because they were able to do it.

3.2. *Explanation, exclusion, and causation*

Let us take stock. I have argued that functional properties such as *being a blood pump* and relational properties such as *being the weakest member of a herd* are properties invoked by non-replaceable causal explanations, that is, explanations whose causal information cannot be captured by any other explanation. Now I want to argue that this implies that such properties are causally efficacious. In general terms, I will defend the notion that, though an explanation that resorts to these properties and an explanation couched in physical terms may explain the same phenomenon, there is no exclusion problem that forces us to choose one or the other. And if we are content with two causal explanations of the same effect, then we are content with two causes of the same effect.

First of all: is it true that the explanation “the antelope died because it was the weakest in the herd”, and the explanation “the antelope died because a hole was opened in its jugular” explain the same *explanandum*? Apparently they do: either if one thinks of effects as Davidsonian unstructured events, or takes them to be Kimian events (i.e., as the instantiation of a property by a particular at a time), it seems that both explanations are aimed at the same effect. But things are a bit more complicated. Jonathan Shaffer (2005) has argued that causation is a contrastive, quaternary, relation: causal claims do not assert that *c* caused *e*, but that *c*, rather than *c'*, caused *e*, rather than *e'*. Are our two explanations explaining the same *explanandum*, if *explananda* are contrastive? One might suspect that they are not, and so one might suspect that, after all, they do not actually explain the same thing. It may be claimed, for instance, that the physical explanation can explain why the heart stopped rather than continuing to pump blood, while the emergent explanation explains why that

antelope, rather than any other, died. Now, while there is clearly something to this (see below), it nonetheless remains true that both explanations do *also* provide information about the very same thing. Both explanations can be made to converge on the same *explanandum*, e.g., why the antelope died rather than survived. That is, even if we take *explananda* to be contrastive—and it does make sense to think about them in those terms in the present case—the two explanations provide different ways to account for the same thing. Each explains things that the other explanation does not account for, but they also both account for the very same thing.

Now, granted this, and granted also that both explanations are complete and independent, should we not eliminate one of them? In the eighties and nineties, Kim was able to convince most philosophers that we accept a principle, called ‘the principle of causal/explanatory exclusion’, according to which there cannot be two independent and complete causes/causal explanations of one event. If we are able to show that one of the two explanations depends on the other, or that one of them completes the other, then we can have both. A case in point is that of functional explanations according to some authors (see Block, 1990, Kim, 1998, Heil, 2003, and Gillett, 2003). The explanation “she fell asleep because she took a sleeping pill” tells us that the person fell asleep because she took some chemical which is covered by the concept of sleeping pill. We can retain that kind of explanation because it is in fact a template which can be completed by a lower-level explanation (one that mentions the chemical directly). However, if we find no way in which one of the explanations can be said to complete or ground the other, we have to eliminate one of them in favor of the other.

However, I want to claim that the exclusion principle Kim put forward is not a principle that should be applied without restriction. He motivated it by appealing to the tension, or instability, that we feel when we face two complete and independent causal explanations (see, e.g., Kim, 1989). Kim was able to convince most philosophers of mind that we feel this kind of tension when confronted with physical and psychological explanations of behavior. But do we feel this tension when we realize that we have two different explanations of the death of the antelope (or of the endurance of hearts)? I do not think so. To be sure, the explanations in question seem to be independent, and they also seem to be complete, or self-sufficient. They are not complete in that each of them leaves out some information that the other one captures. But they are complete in the sense that neither requires the other in order to be satisfactory on its own grounds. If we are told about the physical chain of causes and effects that resulted in the death of the antelope, and that is the kind of explanation we are calling for, we can be perfectly satisfied. But we can also be satisfied if we are interested in another kind of explanation and we are told about the relation that the antelope bore to the other individuals in the herd.

The last point is important: each explanation is complete given certain interests. This suggests that they cannot be considered equivalent explanations, even though both are causal explanations. Kim (1989) motivated his exclusion principle thus: “Another way of putting my point would be this: a certain instability exists in a situation in which two distinct events are claimed to be nomologically *equivalent* causes or explanations of the same phenomenon; stability is restored when *equivalence is replaced by identity or some asymmetric relation of dependence*” (his italics). The concept of a nomologically equivalent explanation is left unexplained.

But while it may be intuitive to think that a psychological and a physical explanation are in some sense nomologically equivalent, it is not, I think, intuitive to regard our two explanations of the death of the antelope as nomologically equivalent. So I do not believe that we feel any tension in committing to both at the same time.

An important respect in which the two explanations are not equivalent is their *explananda*. Here it is interesting to resort to Schaffer's insights about contrastivity. I have said that both explanations explain the same thing, even if *explananda* are accepted to be contrastive. However, we have also seen that the way they arrive at this common *explanandum* is different. They can be said to converge on the same *explanandum*, but as talk of convergence suggests, they do not go down the same path. The physical explanation, on its way to explaining why the antelope died instead of surviving, explains, e.g., why blood escaped from the circulatory system, rather than remaining within the individual, or why the heart stopped rather than going on pumping blood. The relational explanation, in contrast, arrives at the explanation of why the antelope died rather than survived by explaining things like why that antelope, rather than any other, died. It seems that the two explanations result from different interests, and this is revealed by looking at the different things they can explain.

Whether or not the reader finds this last point convincing, I want to insist on what I believe to be essential for my position. When we explain why the antelope got killed, or why hearts are there, we produce an explanation that sheds a particular light on the event to be explained. When we detail the chain of local causes and effects that ended with the death of the antelope, or with hearts' being there, we get a very different kind of causal information. Do we think that one of the explanations has to

go? My intuition is that we do not. They are explanations that can cohabit in harmony, so to speak.

There is an issue regarding whether we can move from these reflections about explanation to claims about causation. *Prima facie*, the move is not problematic if we grant what has been said until now. There are two sides to the exclusion principle: on the one hand, it is an explanatory principle, related to our explanatory practices; on the other, it is a metaphysical principle which concerns causal relations. There is usually no question about which side of the principle is more fundamental: usually, whenever we have two causal explanations for one event we have an experience of instability, a tension, which has to be relieved; and we feel that tension at the metaphysical level too. But it is reasonable to say that the tension at the metaphysical level may arise depending on the way our explanations picture the situation (which is to say: it may arise depending on whether there is a tension at the explanatory level or not). If our explanations tell us that there are two complete and independent nomologically equivalent causes (i.e., if the explanations are nomologically equivalent, whatever that means exactly), then we perceive a tension that has to be relieved. However, common though it is, this kind of situation may not be the only possible situation. It may well happen that our explanations present a different picture to us. For instance, it may be that we do not think that the two explanations are nomologically equivalent, and so we do not feel that there is a tension to be relieved. If, for instance, we think that each explanation sheds a different light on the *explanandum*, we may think that both explanations are acceptable. This has to have an impact at the metaphysical level: our explanations talk of two causes for one event,

but two causes that should not be in competition. That is, the way our explanations picture the situation entails that the causal exclusion principle does not have to be applied.

Things could be more complicated, however, depending on what we say about the nature of the causal relation itself. It could be that we feel no instability at the level of explanation, but that our commitments to a particular view of causality force us to discard one of the putative causes (and so, to consider that one of the explanations is not causal). For instance, if we think that causation consists of, or involves, transmission of marks (Salmon, 1984) or of conserved quantities (Dowe, 2000), we will, as a general practice, eliminate all non-physical causes. In particular, we cannot consider that a relational property such as the property of being the weakest member of the herd is a cause. So, adopting a transmission theory of causality forces us to exclude non-physical explanations regardless of whether we feel some tension at the epistemic level or not. Now, this is not the place to argue for or against transmission theories. I think that the transmission theory is probably a good theory of causation *in physics*, but being a good theory of physical causation does not imply being a good theory of causation (see Vicente, 2006). Thus, there is no reason to believe that causal exclusion problems may behave differently from explanatory exclusion problems. Kim was right in talking about problems of causal/explanatory exclusion.

On the other hand, other theories of causation give results that accord with what I have been defending so far. For instance, James Woodward's interventionist theory seems to fit with my position very well. Woodward (2003) holds the following: for X

to be a cause of Y with respect to a variable set \mathbf{V} , it must be possible to intervene on X and change Y or the probability distribution of Y when one holds fixed at some value all other variables Z_i in \mathbf{V} . That is, a cause is characterized by whether intervening on it would result in a change in the effect. Now, to put this very simply (simplistically, in fact), and applying it to the case of the death of the antelope: if you manipulate the relation between the antelope and its herd, you may save its life. If you manipulate its physical intrinsic properties, you will not, at least not as long as the antelope remains the weakest in its herd. This is in fact a point that has been developed above in the context of discussing whether physical explanations can explain everything: changes in the respective intrinsic properties of the antelope and of the other individuals sometimes do, and sometimes do not, result in changes in the effect. They result in changes in the effect whenever the changes also imply changes in the relation that the antelope bears to the rest of the animals. Thus, if we are looking for changes in the death of the antelope, we have to intervene on the relational property, which means that *being the weakest member of the herd* does seem to have causal efficacy.¹⁴

However, I am not interested in committing myself to any particular theory of causality. My point is intended as a general one which could even serve to constrain theories of causality. The attitude that I would favor is analogous to the attitude Tyler

¹⁴ Things are not nearly as clear in the case of functional properties such as *being a blood pump* and their role in selection processes. On the one hand, you cannot manipulate the functional properties of a particular, x , without manipulating its categorical properties: if you want to test whether, all things being equal, a non-blood-pumping heart does not endure, you have to alter its categorical properties. On the other hand, you can alter the categorical properties of a particular, x , without altering its functional properties, but only as long as the new categorical properties of x have the same causal powers as the old ones. That is, you cannot really test the causal powers of the categorical properties of

Burge (2007, pp. 380-382) recommends in the case of mental and physical causation. Burge holds that as yet we lack a good theory of causation, and that one of the tasks of any theory of causation is to explain how mental and physical causation operate together. I am not that sympathetic to this idea in particular, but I hold a similar view with regard to the cases I am dealing with here. A satisfactory theory of causality has to account for, among other problematic things—such as the causal role of omissions—the fact that relational and functional properties can have effects.

Now let us move to the issue of how this all fits with the widespread assumption that the physical world is causally closed.

4. Last point, and summing up: on the causal closure of the physical world

As is well known, one of the main problems that emergentists have to face lies in how to reconcile their claims with the assumption that the physical world is causally closed. As a matter of fact, this can be regarded as *the* problem for the emergentist. For instance, Kim's (1999, 2006) strategy against emergentism consists of forcing the emergentist to make a commitment to downward causation, so that, once obtained, he can claim that downward causation is incompatible with the completeness of physics. Thus, any work on emergentism has to consider the issue of the causal completeness of physics.

Now, it is possible to wonder why we should believe in the causal completeness of physics. I have argued elsewhere that the so-called causal closure principle (CCP) is not, as some claim (see, e.g. Antony and Levine, 1997), a "truth of contemporary

x vis-à-vis its functional properties.

physics". It is not an *a priori* truth, and, as an empirical truth, the evidence for it, which cannot but be inductive, can be questioned. On the other hand, it is possible to try to make it follow from some real truths of contemporary physics, such as conservation laws, but the steps are controversial (Vicente, 2006, Montero, 2006, Gibb, 2010).

However, for the time being, let us take the CCP for granted, that is, let us assume that it is true that every physical effect has a sufficient physical cause. Let us assume also that we all know what we mean by "physical", so that the CCP has a clear, well-understood, sense. Now the question is: am I indeed forced to deny the CCP in order to maintain my point, that is, that non-physical properties have a role to play in selection processes? My answer is negative: no, there is no need to deny the CCP. Let me explain.

The CCP is a metaphysical principle which states that every physical effect (caused event) has a sufficient (or complete) physical cause. I have conceded that all the physical effects that non-physical properties explain do have physical causes as well. What I have claimed is that the explanations that cite these physical causes do not tell us all we can know about the effects to be explained. This claim can be understood as an outright rejection of the sufficiency claim of the principle. Even more clearly, it seems to be an outright rejection of an epistemic counterpart of the CCP, namely, the principle that physical causal explanations are sufficient causal explanations, or in other words, the idea that physics is explanatorily comprehensive. However, this is not necessarily so: my claim can be made compatible with the CCP and even with this epistemic counterpart. Let me start with the metaphysical principle.

In the context of the CCP, that a cause c is sufficient for an effect e means that c alone is enough to bring about e : if c occurs, then, no matter what, e will occur. This seems to mean that c is a *total cause*, that is, a cause that takes on board all the causal factors required to ensure the occurrence of e . So, what the CCP claims is that physical effects have total causes which are also physical. In the light of the examples discussed above, this seems undeniable: the hole in the jugular opened up by the lion's tooth (plus some further facts about blood circulation in living organisms) is sufficient to cause the death of the antelope. Furthermore, the immediate causal antecedent of this event seems to be sufficient for it as well, and it can be claimed that there is no point in the chain of local physical causes and effects where the cause is not sufficient for the effect.

Thus, nothing that I have argued for enters into contradiction with the CCP. The examples do not show that the CCP is false. Rather, what they show is that, although physical causes are sufficient, they are not the only causes. To say that I am advocating causal overdetermination would not be unfair; except that it is certainly an odd kind of causal overdetermination, very far from the usual case of the two causes that coincide in bringing about an effect. Indeed, if this "usual case" gives the content to the notion of causal overdetermination (i.e. if we understand that overdetermination is what happens when, e.g., two different shots reach the heart of a subject at exactly the same time), then perhaps it would be best not to talk of overdetermination here at all. Be that as it may, what I have wanted to argue for is that in selection processes, sufficient physical causes are not the only causes that matter. That is, there are indeed sufficient physical causes in selection processes, but in spite of being sufficient, they are not the only causes.

Turning now to the epistemic counterpart of the CCP: the preceding paragraphs mean that physics can indeed be explanatorily comprehensive: it can come up with an explanation couched in its own terms that points at a sufficient cause of the *explanandum*. However, it seems clear that my position entails that physics cannot provide all the information that can be provided about certain events within its realm. Above I distinguish two senses of “completeness” and claim that physical explanations are complete in one sense (they are self-sufficient) but not in another (they do not capture all the existent causal information about the effect). So, physical explanations are complete in the sense required for the epistemic counterpart of the CCP to be true, but they are not complete in the sense of exhausting all that there is to know. However, I do not find this result problematic. It may be that the CCP is a truth of physics, and accordingly, a truth that philosophers are not in a position to deny. But the claim that physics can exhaustively explain all that happens in the physical world is not a truth of physics, and it does not follow from the CCP either.

Acknowledgements

This paper has benefitted from discussions held in several seminars on emergence and downward causation organized by the IAS Research Group at the University of the Basque Country. I want to express my gratitude to Leonardo Bich, Arantza Etxeberria, Matteo Mossio, Alvaro Moreno, and particularly to Jon Umerez, who made a thorough reading of the draft and suggested improvements that I found very helpful. Thanks as well to the audience at the 6th SEFA Conference, in particular to Jordi Fernández. I also have to express my gratitude to two anonymous referees, who

did a wonderful job providing comments that enabled me to improve the paper significantly and to correct several mistakes, as well as to the editor of *ISPS*, James McAllister. Finally, I want to thank Toffa Evans for the English revision. Research for this paper was funded by the Spanish Government, research projects FFI2010-15717 and FFI2011-30074-C02-02.

References

- Antony, L. and Levine, J. (1997) "Reduction with Autonomy". *Philosophical Perspectives* 11: 83-107.
- Bird, A. (2007) "Causal Exclusion and Evolved Emergent Properties", in Groff, R., ed., *Revitalizing Causality: Realism about Causality in Philosophy and Social Science*, Abingdon: Routledge.
- Block, N. (1990) "Can the Mind Change the World?" in Boolos, G., ed., *Essays in Honor of Putnam*. Cambridge: Cambridge University Press.
- Burge, T. (2007) "Postscript to Mind-Body Causation", in *Foundations of Mind, Philosophical Essays, vol. 2*. Oxford: Oxford University Press.
- Campbell, D. T. (1974). 'Downward causation' in hierarchically organised biological systems, in: Ayala, F. J. and Dobzhansky, T., *Studies in the Philosophy of Biology: Reduction and Related Problems* (pp. 179-186). Berkeley and Los Angeles, CA: University of California Press.
- Campbell, D. T. (1990) "Levels of Organization, Downward Causation, and the Selection-Theory Approach to Evolutionary Epistemology", in: Greenberg, G. and Tobach, E., *Theories of the Evolution of Knowing* (pp. 1-19). New Jersey: Lawrence Erlbaum.

- Cartwright, N. (1999) *The Dappled World*. Cambridge: Cambridge University Press.
- Crook, S. and Gillett, C. (2001) ‘Why *Physics Alone* Cannot Define the ‘Physical’’, *Canadian Journal of Philosophy* 31: 333-360.
- Davies, P. (2006) “The Physics of Downward Causation”, in *The Re-Emergence of Emergence*, Davies, P. and Clayton, P., eds., Oxford: Oxford University Press.
- Dowe, P. (2000) *Physical Causation*, Cambridge: Cambridge University Press.
- Dupré, J. (2001) *Human nature and the limits of science*. New York: Oxford University Press.
- El-Hani, C. N. and Emmeche, C. (2000) “On some theoretical grounds for an organism centered biology: Property emergence, supervenience and downward causation”. *Theory in Biosciences* 119: 234-75.
- Gibb, S. (2010) “Causal Closure Principles and the Laws of Conservation of Energy and Momentum”. *Dialectica* 64: 363-384.
- Gillett, C. (2003). “The Metaphysics of Realization, Multiple Realizability, and the Special Sciences”, *Journal of Philosophy* 100: 591-603.
- Gillett, C. (2006) “Samuel Alexander's Emergentism”, *Synthese* 153: 261-296.
- Heil, J. (2003) *From an Ontological Point of View*, Oxford: Clarendon Press.
- Jackson, F. and Pettit, P. (1990) “Program Explanation: A General Perspective”, *Analysis* 50: 107-117 (reprinted in Jackson, F., Pettit, P. and Smith M., *Mind, Morality and Explanation*, Oxford University Press, 2004).
- Kim, J. (1989) “Mechanism, Purpose and Explanatory Exclusion”, *Philosophical Perspectives*, Vol. 3: Philosophy of Mind and Action Theory, 77-108. Reprinted in Kim, J. (1993).
- Kim, J. (1993) *Supervenience and Mind*. Cambridge: Cambridge University Press.

- Kim, J. (1998) *Mind in a Physical World*. Cambridge: Cambridge University Press.
- Kim, J. (1999) "Making Sense of Emergence", *Philosophical Studies* 95: 3-36.
- Kim, J. (2006) "Emergence: Core Ideas and Issues", *Synthese* 151: 547-559.
- Millikan, R. G. (1989) "In Defense of Proper Functions", *Philosophy of Science*, 56: 288-302.
- Montero, B. and Papineau, D. (2005) "A Defence of the *Via Negativa* Argument for Physicalism", *Analysis* 65: 233-237.
- Montero, B. (2006) "Physicalism and the Conservation of Energy", *Dialectica* 60: 383-396.
- Moreno, A. and Ruiz-Mirazo, K. (2009) "The Problem of the Emergence of Functional Diversity in Prebiotic Evolution", *Biology and Philosophy* 24: 585-605.
- Mossio, M., Saborido, C. and Moreno, A. (2009) "An Organizational Account of Biological Functions", *British Journal for the Philosophy of Science* 60: 813-841.
- Mossio, M. Bich, L. and Moreno, A. (forth.) "Emergence, closure and inter-level causation in biological systems", *Erkenntniss*.
- Neander, K. (1991) "Function as Selected Effects: The Conceptual Analyst's Defense", *Philosophy of Science*, 58: 168-184.
- Rosenberg, A. and Bouchard, F. (2010) "Fitness", *The Stanford Encyclopedia of Philosophy (Fall 2010 Edition)*, Edward N. Zalta, ed., URL = <<http://plato.stanford.edu/archives/fall2010/entries/fitness/>>.
- Salmon, W. (1984) *Scientific Explanation and the Causal Structure of the World*, Princeton, Princeton University Press.

- Schaffer, J. (2005) "Contrastive causation". *Philosophical Review* 114: 327-358.
- Sperry, R. (1969) "A modified concept of consciousness", *Psychological Review*, 76: 532-536.
- Vieira, F. S. and El-Hani, C. N. (2008) "Emergence and Downward Determination in the Natural Sciences", *Cybernetics & Human Knowing* 15: 101-134.
- Vicente, A. (2006) "On the Causal Completeness of Physics", *International Studies in the Philosophy of Science*, 20: 149-171.
- Vicente, A. (2011a) "Functions and Emergence: When Functional Properties have Something to Say", *Philosophical Studies*, 152: 293-312.
- Vicente, A. (2011b) "Current Physics and "the Physical"", *British Journal for the Philosophy of Science*, 62: 393-416.
- Wilson, J. (2010) "Non-reductive Physicalism and Degrees of Freedom", *British Journal for Philosophy of Science*, 61: 279-311.
- Woodward, J. (2003) *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press.
- Wright, L. (1973) "Functions" *The Philosophical Review*, 82: 139-168.
- Yablo, S. (1992) "Mental Causation", *Philosophical Review* 101: 245-80.