

## THE DUAL *EXPLANANDUM* STRATEGY\*

AGUSTÍN VICENTE

Departamento de Filosofía, Lógica y Filosofía de la Ciencia  
Universidad de Valladolid  
agusvic@fyl.uva.es

**SUMMARY:** In this paper I try to fix the price that a non-epiphenomenal dualism demands. To begin with, the defender of non-epiphenomenal dualism cannot hold that mental events cause physical events, since the physical world is causally closed. Hence, she must say that mental events cause events that are not physical, or at least, events that are not affected by the principle of the causal closure of the physical world (this is the “dual *explanandum* strategy”). However, this is not all: the events mental causes bring about must fulfill certain further conditions, which I spell out. When properly analyzed, it will be seen that these conditions make the dual *explanandum* strategy highly demanding.

**KEY WORDS:** mental causation, dual *explanandum* strategy, conceptual divides, dualism

**RESUMEN:** En este artículo trato de fijar el precio que un dualismo no epifenoménico tiene que pagar. Para empezar, el defensor del dualismo no epifenoménico no puede mantener que los eventos mentales causan cambios en el mundo físico, ya que éste está causalmente cerrado. Por lo tanto, ha de decir que los eventos mentales causan eventos que no son físicos, o, al menos, que no están sometidos al principio del cierre causal del mundo físico (ésta es la “estrategia del doble *explanandum*”). Sin embargo, esto no es todo: los eventos que son efectos de las causas mentales tienen que cumplir ciertas condiciones ulteriores, condiciones que detallo. Cuando éstas se analizan propiamente, se ve que vuelven muy exigente la estrategia del doble *explanandum*.

**PALABRAS CLAVE:** causación mental, estrategia del doble *explanandum*, abismos conceptuales, dualismo

<sup>1</sup> I have to thank Manuel García-Carpintero for his comments and advice, and an anonymous referee for his/her comments. My gratitude also to the people at the CESAP conference of Budapest, where a preliminary version of this paper was presented. Research for this paper has been funded by a postdoctoral fellowship provided by the Basque Government.

## 1. *Introduction*

Physicalism has been dominant in contemporary analytic metaphysics for quite a long time. There is, however, no clear idea about what it amounts to or what the arguments for it are, if any. There are some doctrines that look like twins of physicalism, but can nevertheless be separated from it, such as naturalism and materialism. According to Crane and Mellor (1990, p. 186), materialism is “a seventeenth-century [...] metaphysical doctrine; it attempted to limit physics *a priori* by requiring matter to be solid, inert, impenetrable and conserved, and to interact deterministically only by contact”, while naturalism is the thesis that there is nothing in the world that is not reducible to, or explainable in terms of, entities postulated by the true natural sciences. Physicalism, in a first approach, can be considered as a species of naturalism, and very different from materialism, which is, as depicted, false. The specificity of physicalism, as contrasted with other possible varieties of naturalism, lies in the preeminence given there to physics: physical entities, physicalism says, constitute the basis of the rest of natural entities, hence of all entities whatever.<sup>1</sup>

As I say, there is no definite consensus about the reasons to become a physicalist. One possible argument makes use of epistemic considerations, having to do with our need of unificatory movements. But it is also possible to build a straightforward ontological argument. In this paper I will be concerned with this latter route to physicalism; in particular, with what

<sup>1</sup> “Liberal” physicalists take it that this “constituting the basis of all other entities” means that the rest of entities in the world supervene, depend on, or are determined by, the physical entities. Usually, higher level particulars are said to be linked by the composition relation to basic particulars, while higher level properties, they say, are realized by basic physical properties (see Poland 1994 for a full development of this account). More restrictive physicalists, on the other hand, want physical entities to exhaust (at least) the whole of actuality, so that physical entities constitute the basis of the rest of entities in the sense that these latter are nothing but physical entities. If we concede, as I am about to do, that the argument from exclusion is a good argument for physicalism, then, liberal physicalism is thereby excluded as a viable version of physicalism: supervenience accounts of mental causation have been shown to be mistaken (see specially Kim 1993, 1998).

some philosophers have lately presented as *the* argument for physicalism (see e.g. Peacocke 1979, Papineau 1990, Antony and Levine 1997, Sturgeon 1998, Levine 2001): the argument from exclusion.<sup>2</sup> Very briefly put, the argument says that if non-physical events<sup>3</sup> are to interact causally with the physical world, then they have to be physical, because only physical events bring about changes in the physical world. As it can be seen, this argument is an argument for physicalism only if one is committed to the causal ability of mental events to change the physical world. In this sense, the argument does not concern the growing number of epiphenomenal *qualia* dualists as such. But what about those dualists that have the intuition that mental events, or at least, *some* mental events, e.g. intentional events, are causally efficacious? What I aim to do in this paper is to explore the chances that such a dualism may survive the argument from exclusion.

## 2. *The Argument from Exclusion*

The argument that I am about to present derives from the famous and very much discussed problem of causal exclusion for mental causation. According to the proponents of the argument, physicalism emerges as the only possible solution to such a problem. Given that (apparently) there is no plausible way to deny any one of the propositions that make up the problem, this only possible solution comes as the conclusion of an argument that has the problematic propositions as premises.

The propositions whose denial looks so implausible are:

<sup>2</sup> I do not think it is *the* argument for physicalism: on the one hand, it leaves *qualia* dualists untouched. On the other, the most promising and interesting naturalist/physicalist accounts of the mind, such as Millikan's and Dretske's construe mental properties as teleological. And teleological properties are not, as such, causally efficacious. So the basic reason for becoming a physicalist cannot be related to the causal efficacy of mental properties.

<sup>3</sup> I will be working with Kim's account of events, that is, events are structured wholes made up of individuals, properties and times, such that an event is an individual instantiating a property at a time. Although the argument may well affect all (presumed) non-physical events, I will restrict my discussion to mentality.

- (i) principle of the causal closure of the physical: every physical effect has a sufficient causal antecedent that is also physical;
- (ii) causal efficacy of the mental: mental events produce changes in the physical world;
- (iii) principle of the causal-explanatory exclusion (PEE): there cannot be two causes/causal explanations that are both complete and independent for one event, except in cases of overdetermination.

Proposition (i), it is said, is backed up by contemporary physics. Proposition (ii) stems from our most common experiences of the world and ourselves. Proposition (iii), finally, can be seen either as a summary of some of our explanatory practices and metaphysical beliefs or, more simply, as a somewhat oblique definition of overdetermination. Then, if these three propositions are indeed undeniable, and leaving at one side the possibility that there is mind-body overdetermination, we are forced to conclude:

- (iv) mental events are physical events.<sup>4</sup>

The anti-physicalist may try to resist the argument from exclusion by directly denying one of its premises or by rejecting perverse assumptions they think lurk behind them (cf. Burge 1993). Emergentism denies premise (i), epiphenomenalism dispenses with (ii) and some yet unnamed group of philosophers try to deactivate (iii), the principle of causal exclusion.<sup>5</sup> I think,

<sup>4</sup> This conclusion should rather be: mental events are either identical to, or dependent on, physical events. However, as said above (fn. 1), it has been shown that no dependency relation can solve the exclusion problem for mental causation. At least, I am going to assume that such is the “state of the art” in that discussion.

and will assume, that the denial of (i) and the denial of (iii) have been shown to be wrong (see, specially, Kim 1998) and that epiphenomenalism about intentional properties is not a tenable position either. Premise (ii), however, can be rejected on grounds different to those of the epiphenomenalist. Instead of saying that mental events are inert, it is open to the critic of physicalism to claim that mental events are causally potent, only that they do not cause events that are physical. This is the “dual *explanandum* strategy”.<sup>6</sup>

<sup>5</sup> One can reject the principle of the causal-explanatory exclusion, or restrict its range of application so that it does not pose a problem for mental causation. In a reading of the principle, this would mean that mental and physical properties are overdetermining some effects, but defenders of this approach would for sure deny that this could be a normal type of overdetermination. Rather, what these anti-physicalists want to say is that the application of the principle introduces a debate —being one of its positions the overdetermination case— from which one had better step outside. The truth is that there is no competition, neither causal nor explanatory, between mental and physical properties. Those defending this position have it that epistemic/explanatory facts rule over ontological facts, so that if we have two explanations for one event that seem to “illuminate” it from different perspectives, then our ontology should accommodate this explanatory fact of the matter. In R.A. Wilson’s (1995) words “where explanation goes, ontology must follow”.

Some of these philosophers combine this approach with the dual *explanandum* strategy to be discussed. I take Rudder Baker to be one such author. As a matter of fact, it is not easy to discern when one strategy ends and the other begins in these cases, but I think the following positions are intelligible: (a) ontology-first plus dual *explanandum* strategy; (b) epistemology-first plus dual *explanandum* strategy; (c) epistemology-first without dual *explanandum* strategy. This last position, though conceivable, is more unstable than position (b), since there is a considerable purely epistemic pressure for the principle of explanatory exclusion when two explanations are aimed at the same event. That is, even if we play an epistemic non-ontological game I find it difficult to deny the exclusion principle. In what follows, I will restrict my discussion to what I take to be an interesting response to the argument from exclusion, that is, the realist dual *explanandum* strategy.

<sup>6</sup> It is to be noted that I am only concerned with the dual *explanandum* strategy *vis-à-vis* physicalism, and not as a solution to the exclusion problem *simpliciter*. Fred Dretske has been read as a defender of the dual *explanandum* approach, but inside the physicalist framework. What he can be read as claiming is that mental properties are physical though extrinsic properties that cause the instantiation of physical properties other than bodily

Two critics of physicalism have lately made use of this strategy: Scott Sturgeon 1998, 1999, and Jennifer Hornsby 1997. My aim in what follows is to present and criticize their respective accounts in order to show what (high) price the strategy exacts.

### 3. *The dual strategy*

#### 3.1. Preliminaries

I will start by posing a problem to the defenders of the dual *explanandum* strategy so that it begins to become clear what such a defender must be committed to. The problem is none other than the exclusion argument again, but now a stronger version of the principle of causal-explanatory exclusion is used, namely:

- (iii') there cannot be two explanations/sufficient conditions for one event that are independent, except in cases of overdetermination.

Kim (1998) makes use of this principle in step (vi) in his argument from supervenience, destined to force downward causation. There he is considering mental-to-mental causation, and he asks: how can the instantiation of a mental property cause the instantiation of another taking into account that the instantiation of the second mental property can be explained by the instantiation of a physical property that subvenes it? Answer: only by causing such a physical property. The case I want to consider is not different. For suppose I say that mental events cause behavioral events, which are not identical to any kind of physical events. However, as it happens, every time a behavioral property is instantiated, a given physical property gets instantiated simultaneously, and the instantiation of this physical property is sufficient for the instantiation of the behavioral property. If this is the case, then we are forced to say that mental events cause behavioral events only by causing some physical events, and the strategy is thereby dismantled. Thus, it seems crucial

movements —which are the effects of the instantiation of intrinsic properties. I leave Dretske's account at one side.

for the defender of the dual *explanandum* approach to maintain that behavioral events do not supervene on physical properties.

### 3.2. The Conceptual Divide between the Macro and the Micro

Scott Sturgeon (1998, 1999) has defended a version of the dual *explanandum* strategy that accords to the constraint I have just presented. His view is first, that macro physics is not causally closed, second, that behavioral events belong to the macro image and third, that micro events are, for all we know, not sufficient for macro events. I will explain these in turn.

First of all, Sturgeon distinguishes between quantum-mechanical completeness and broadly physical completeness. This distinction gives us two different arguments, depending, that is, on whether we have

- (i') every quantum effect has a sufficient causal antecedent that is also quantum-mechanical

or

- (i'') every broadly physical effect has a sufficient causal antecedent that is also broadly physical.

Now the problem for the physicalist, according to Sturgeon, is that while (i') is true, nobody holds that mental events have quantum effects. Rather, what both science and common sense claim is that mental events have broadly physical effects. But, for that same reason, (i'') seems to be false. In other words, contemporary physics only backs (i'), which does not affect mental causation, since mental events cause just broadly physical events.

However, things are not so easy for the anti-physicalist. The defender of the argument from exclusion, Sturgeon admits, can invoke the fact that quantum events compose macro events in order to impose a completeness principle. This can be obtained by holding (p. 417)

- (a) *Closure under downward-composition*: If  $C$  causes  $E$  and  $E$  is composed by  $E^*$ , then  $C$  causes  $E^*$ .

- (b) *Closure under upward-composition*: If  $C$  causes  $E$  and  $E$  composes into  $E^*$ , then  $C$  causes  $E^*$ .

Then, as every macro event is composed by quantum events and every quantum event has a complete quantum cause, then every macro event has a complete quantum cause (by (b)). Also, as every macro event is composed by quantum events, then (by (a)) some macro causes cause quantum events (and thus compete with quantum causes).

But this is not the end of the story. Principles (a) and (b) are false as they stand. Sturgeon presents three counterexamples of increasing strength to them, and moves on to offer a rewriting. The basic idea now is that causal relations can move across different levels of reality *only if* there is a relation of essence inclusion between the effects (see Yablo 1992). Then, principles (a) and (b) can be true when, and only when, the relation of composition is a relation of essence inclusion.

In Sturgeon's words (p. 422): "The moral is this: intuition sees causation flowing across composition to just the degree it sees composition respecting an effect's essence. Intuition reflects our commitment to the Cause-and-Essence principle:

- (C&E)  $C$  causes  $E$  iff  $C$  is sufficient to bring about what is essential to  $E$ ."

Stemming from this principle, we have:

- (a') If  $C$  causes  $E$  and  $E$  is essentially composed by  $E^*$ , then  $C$  causes  $E^*$ .  
 (b') If  $C$  causes  $E^*$ , and  $E^*$  essentially composes into  $E$ , then  $C$  causes  $E$ .

So the reasoning must be the following: if every macro event is essentially composed by quantum events and every quantum event has a complete quantum cause, then every macro event has a complete quantum cause. Also, if every macro event is essentially composed by quantum events, then (by (a)) some macro causes cause quantum events (and thus compete with quantum causes).



The role of principle (a') is, as can be seen, similar to that played by the generalized version of the exclusion principle in the last section. There, what the physicalist wanted to obtain was the commitment to downward causation, and this can be done by pressing the defender of the dual *explanandum* strategy either with the exclusion principle or with principle (a'). The difference is that this second principle gives rise to an argument that is considerably softer: for the argument from generalized exclusion to work we only need that parts  $E^*$  be sufficient for  $E$ , whereas here we require  $E$  to be a mereological whole made of essential parts. That's a far stronger condition that generates an argument with a much narrower range of application.

It is also interesting to contrast principle (b') with an hypothetical argument for upward causation starting from the generalized principle of exclusion. Parallelisms here are less, as can be noted. First, the argument using the generalized principle of exclusion cannot provide an equivalent to (b'). For that we would need that the instantiation of  $E$  were sufficient for the instantiation  $E^*$ . Under that assumption, it would be possible to say that since  $E^*$  cannot have two sufficient conditions,  $E$  and  $C$ , then  $C$  must have caused  $E$ . But even if  $E^*$  essentially composes into  $E$ , that does not mean that  $E$  is sufficient for  $E^*$ : to use Sturgeon's example, certain duck movements essentially compose a duck flock movement, but the presence of that duck flock movement is not sufficient for the presence of those duck movements. On the other hand, and much more important, the argument from the generalized exclusion principle commits the physicalist to something he would never accept, namely, that micro events cause other micro events via the instantiation of macro properties.

But, of course, all this is good news for the physicalist: his position is now stronger, for he has a good argument for downward causation (which he previously had) and another for upward causation (which he did not).

Now let's see how Sturgeon tries to resist the argument he has construed. Of course, he admits the Cause-and-Essence principle, and its corollaries (a') and (b'). What he rejects is that the

composition of quantum events are essential for macro events. And this he justifies by pointing to the conceptual gap that there is between the macro and the micro images.

Sturgeon's denial of the supposition that the quantum world compose essentially the macro world applies not only to relations between objects but also to relations between spatial events. The argument is in both cases the same: macro events are conceptually remote from quantum events. In the case of objects, this applies as follows:

The conceptual divide between quantum and macro reality is great. We should be cautious in accepting unrestricted causal flow between the two. For we cannot see how quantum events build into the splendor of macro reality. We cannot see, for example, how quantum tunneling *could* build into jalapeños pepper. (1998, p. 424)

The case of spatial events (position and movement), however, is more difficult, since “if the position and movement of anything micro is fixed at the micro level, while the position and movement of everything macro is fully constituted by micro positions and movements, then, it seems, macro movements are caused by micro causes”. Nonetheless, Sturgeon goes on to say, there is again a conceptual divide between macro and micro reality that makes this suggestion false: quantum spatial events have two intriguing characteristics, *superposition* and *projection*, that macro spatial events do not have, and these make it difficult to guess how quantum spatial events could build into macro spatial events. According to the superposition effect, a particle's position when measurement is not taking place is best conceived as a wave-like phenomenon that has a certain probability of being at various positions: “if a particle can be located at P1 or P2 or P3, then it can also be characterized by a combination such as  $(1/3 P1 + 1/3 P2 + 1/3 P3)$ ” (p. 425). According to the measurement effect, on the other hand, when measured the wave collapses into a particle in one particular position. Needless to say, none of these striking effects govern macro spatial events. What is more, we cannot understand how

things that accord to them can give raise to the determinacy we observe in the macro world.

So this is the argument for the conceptual gap. As said above, Sturgeon goes from here to the claim that there is (perhaps) no essence inclusion relationship, hence no causal relations, between quantum events and macro events. The conceptual gap, says Sturgeon, precludes the occurrence of the following two conditions for (the existence of) causal relations between distinct levels: (a) the regions indexed to macro causal claims should be contained within those indexed to micro causal claims; (b) the patterns indexed to macro causal claims should be identical to those indexed to micro causal claims. (A) does not hold because, for all we know, i.e. for all our concepts tell us, there can be quantum worlds in which macro events are absent, and macro worlds with no quantum events in them. As for (b), macro and micro patterns of counterfactual dependence do not coincide: again, the diversity of our concepts permits macro counterfactuals be independent from quantum counterfactuals.

The idea here seems to be that, even though there is some quasi-causal relation between quantum events and macro events in actuality, this does not have any modal force whatever, and hence it cannot be considered as causal. Sturgeon calls this *inducing*: “actual quantum events induce hand movements via composition and causation; and actual hand movements induce quantum events via composition and causation. What fails, at our world, is the link between so inducing and event and causing it” (p. 429). To repeat, this failure is due to the diversity of our concepts for the macro and the micro: there cannot be modal force—hence no causality—if there is no relation between essences, and this we do not know. There might be, or there might not; our conceptual apparatus does not justify either position. This simple epistemic point seems sufficient to block the application of the Cause-and-Essence principle.

At the same time, it permits resisting the argument from the generalized principle of exclusion for downward causation: we as yet do not know whether a certain sum of quantum events is

sufficient for the instantiation of a given macro property. Macro events, for all we know, may not supervene on micro events.<sup>7</sup>

My criticism of Sturgeon's position is quite straightforward. Before passing to it, however, I want to make some comments about his argument. What Sturgeon has shown is not that quantum events cannot cause macro events across the divide and vice versa. All he has established, if the rest is conceded, is that we do not know whether there are these causal diagonals; for all we know, there could be. His point must be that there is no way to know whether quantum event composition provides the essence of macro events. He cannot deny that they do. From what he says, it only follows that we must suspend our judgment about causal relations between these different realms. A claim for their non-existence is as unmotivated as it is a claim for their existence. I do not know how strong a mysterianist Sturgeon is. Perhaps it is reasonable to be skeptical about a meeting of the concepts of the macro with the concepts of the micro, but even then you are not allowed to deny that quantum events are causing macro events so that there is not a problem about *causal* exclusion.<sup>8</sup> (Whether there is a problem about explanatory exclusion is quite another matter.)

Now, conceded this "mysterianism" about macro-micro relations, my doubts concern its import as to mental to physical relations. To put my point in a nutshell: the argument allows macro causation run free, but maybe what constrains mental causation is not quantum mechanics but macro causation itself. Let me explain.

As said above, Sturgeon starts by distinguishing quantum-mechanical completeness from broadly-physical completeness.

<sup>7</sup> Or perhaps the supervenience is stochastic, that is, it conforms to the following pattern: Necessarily, for any  $x$  and  $A$ -property,  $x$  has  $A$  only if there is a  $B$ -property  $B$  and a number in  $[0,1]$  such that  $x$  has  $B$  and  $\text{prob}(A|B) = n$ . As Sturgeon says, "[this] is yet another odd possibility opened by the conceptual gap" (Sturgeon 1999, p. 379).

<sup>8</sup> It could also happen that our concepts of the macro and our concepts of the micro evolve and finally come more or less together. Now given that physicalism is the thesis that all entities are reducible to the entities postulated by a true future physics, this possibility would undermine Sturgeon's argument. So he needs a further argument motivating constant concept separation.

Quantum mechanical completeness is true, he says next, but broadly-physical completeness is false. Then comes the argument for the skepticism about causal diagonals. However, the following positions are possible: (1) broadly-physical completeness is true, and so there is still an argument for physicalism and against mental autonomy; (2) it is not compulsory that the broadly-physical be causally closed for mentality to be in great trouble: trouble comes as soon as the broadly-physical provides its own candidate for causing behaviors. First: Way before Max Planck started to speak about *quanta*, physicists believed the physical world was causally closed by way of conservation principles such as the principle of the conservation of energy and the principle of the conservation of quantity of movement. This belief about the world of macro physics was perceived to be in conflict with our belief in the causal efficacy of the mental; hence, Cartesian interactionism, Malebranche's occasionalism, etc. Thus, it seemed that if you wanted mental events to cause macro physical events, you had to have them physical, otherwise you would be violating what our best theories about the macroscopic world said. These days we think that the best theory of the macro world is Relativity Theory. However, the relativistic description of the macro world is as causally closed and —apparently— as conceptually remote from quantum mechanics, as was classical mechanics. *Contra* Sturgeon, I would say that there is no room for mental causation in the best description of the macro world available.

But even if you can resist this, you are going to have a hard time trying to escape Malcolm's problem of exclusion, namely that "if we bear in mind the comprehensive aspects of the neurophysiological theory—that is, the fact that it provides sufficient causal conditions for all movements—we shall see that desires and intentions could not be causes of movements" (Malcolm 1968, p. 136). This is easily converted into an argument for physicalism of a sort, or, to be faithful to my own distinctions, for naturalism. The argument, following the lines of the argument from exclusion explained above, establishes that mental properties have to be reduced to neurophysiological

properties. And this it does without invoking any principle of causal closure.

So Sturgeon has shown that the quantum world and the macro world may lead independent lives. However, he has not shown that the psychological world has that independence from the macro physical world. If mental events cause macro physical events, then mental events have to be macro physical events (neurological or whatever). So we are not very far from where we started, and this in its optimistic interpretation means that the dual *explanandum* strategy is still open for the anti-physicalist.

### 3.3. The Conceptual Divide between the Physical and the Mental

Most philosophers that take the dual *explanandum* way out of the exclusion problem have it that mental concepts and macro concepts are part of the same world-view, so to speak. They call it common sense. What these philosophers (see, e.g. Baker 1995, Hornsby 1997)<sup>9</sup> defend, then, is that common sense or folk theories cannot be reduced to scientific theories, both in their *explanatia* and in their *explananda* and so we are not entitled to claims for a unique privileged level of causal reality. As a matter of fact, the tendency is to hold that there are two kingdoms, of equal status, of causal reality, but that is, as far as I can see, only supported by the anti-realist turn they tend to take. As I have said before, the presence of a conceptual divide only allows skepticism, neither the claim for causal diagonals nor its denial.

Now I think these philosophers had better keep macro events and mental events separate. The science of physics, by way of conservation laws, has shown that macro events are subject to the principle of the causal closure of the physical world:<sup>10</sup> men-

<sup>9</sup> See her “Causation in Intuitive Physics and in Commonsense Psychology”. The rest of the papers included in her 1997 point to a divide between the physical in any sense and the mental, but in this paper she seems to commit herself to a causal continuum between mental events and “intuitive physical” events.

<sup>10</sup> In Vicente 2001, I try to establish the connection between conservation laws and the causal closure principle.

tal events would have a difficult life within that framework. It could be claimed that folk physics is quite another matter, but I do not see how it can be defended that there is, as Sturgeon puts it, a conceptual Grand Canyon between folk physics and scientific physics, specially if we are concerned with movements of objects. Where does the abyss lie here? Besides, I do not really see why the question should be whether common sense theories are to be reduced to scientific theories. The problem can also be put in this way: here we have different sciences and an argument for them becoming either just one or at least different descriptions of just one reality. Physicalism, as far as I can see, is not synonymous with scientism.

Therefore, the question I am going to address in this section is whether mental entities, both causes and effects, are essentially related to physical properties (or supervenient on them), or, quite to the contrary, are so remote conceptually that we cannot maintain that both types of events —physical and mental— have common effects.

In my discussion I will make use of Jennifer Hornsby's ideas, for she is someone who has been defending the approach I am now considering very explicitly and for some considerable long time. However, she is of course not the only one: these ideas do not come *ex novo*.

Hornsby claims to be defending a kind of naturalism, which she calls "naive naturalism". This, however, does not conform to the view I have taken about naturalism, nor, I guess, to what other philosophers speak about when they use the term. As a matter of fact, her arguments are anti-naturalist in the usual sense of "naturalism". This will be a part of her position that I will just skip. Another part that I am not going to consider is her anti-realism privileging explanation over ontology. As I have explained before, I am here evaluating only the dual *explanandum* strategy assuming realism. From this perspective, from the existence of two types of causal explanations plus a conceptual divide between them only skepticism about causal interactions across the divide follows. Finally, although the point seems very dear to her, I will obviate the distinction Hornsby draws be-

tween causes and causal-explanations. The idea is that causal relations have events as *relata*, while causal-explanations can be relations between states. This means that actions do not have causes, but causal-explanations, because their *explanantia* are states of the entire person, not mental events. This idea introduces a complication in the argument which I believe is unnecessary.

Hornsby's basic claim is that there are two different chains of events occurring before and when a human or animal body is moved. First (see the graph in Hornsby 1997, p. 111) there is the chain starting with the presence of objects of perception; this causes mental events/states of a person; and this, in turn, brings about the desired effects of an action. On the other hand, we have events at sensory surfaces causing brain events/states of a subject of neurophysiological states, and these causing in turn movements of a body. Now according to such a view, there are two "moments" in which the chains can touch each other, so to speak. First, it could be argued that actions are movements of a body or some other kind of physical entity; second, it is possible to hold that it is the desired effects of the action which are identical to bodily movements. If any of these is true, that is, if the chains happen to have some element in common, just the same entities would be involved in one and the other (by the argument from exclusion), hence physicalism would be true. Hornsby's response to the first possible attack by the physicalist is that actions are neither bodily movements nor any kind of physical on-goings internal to the subject: an action, she says, "*is a cause of a bit of the agent's body's moving, and [ . . . ] is an event of the agent's trying to do something*" (p. 85). Also, "actions are not to be identified with anything that can be latched onto from an impersonal point of view" (p. 87). Action reports, she says, are essentially transitive —A's moving of her arm *vs.* A's arm moving— and so actions must involve more than just bodily movements: they involve psychological states of a person.

So Hornsby's response to this attack consists basically in having actions as problematic for the physicalist as purely mental



events or states themselves. But the response is useless if the second attack is successful: what kind of things are the effects of actions? Aren't they bodily movements? And Hornsby's answer here is: yes, but of a different sort. There is not a unique category of bodily movements: rather, our conception of them is disjunctive. On the one hand, we have the bodily movements which are effects of neural events; on the other, we have bodily movements caused by actions. Here is how she puts her point:

A non-disjunctivist conception of bodily movements, as this is now to be understood, is the basis of much physicalist doctrine. [...] The argument [for physicalism] relies on supposing that there is, as it were, just one bodily movement when there is an action. It holds that we have to identify the causes of actions with the causes of bodily movements discovered by neuroscientists, for we should otherwise have to think that bodily movements are overdetermined. [But] if the disjunctive conception of bodily movements is correct, then a premise of the argument is false. (p. 107)

This is just the first step of what she, as any defender of the dual *explanandum* strategy, needs. But there is more that she needs and does not provide, as far as I can see. First, (i) it is necessary that bodily movements that are effects of neural events (*A*-movements) are not sufficient conditions for bodily movements that are effects of actions (*B*-movements), and that there is not a relation of essence inclusion between them. Or at least, we need reasons to be skeptical about all this. That is, we at least need that a conceptual Grand Canyon separates both kinds of bodily movements. Second, (ii) the difference between one kind of bodily movement and the other cannot consist merely in their being effects of different causes. Finally, (iii) the effects of *B*-movements cannot at the same time be the effects of *A*-movements, nor the effects of the latter be the effects of the former, etc. That is, the argument from exclusion must not be reproducible at any other stage of the chain of causes and effects.

These, I claim, are in the end the demands of the dual *explanandum* way out of the argument from exclusion. Thus, as I have explained, these are the conditions a dualist must go through. Let me now explain, taking them in reverse order.

#### 4. *Necessary Conditions for Dualism*

Robert Matthews puts point (iii) in this way:

Unless Hornsby is prepared to extend this disjunctive conception throughout the natural world, in what would in effect be a sort of conceptual dualism, there will be *explananda* for which the two sorts of explanation are in competition. At that point, the usual worries about causal overdetermination arise. (1998, p. 893)

Here Matthews offers, somewhat rhetorically, an escape for the anti-physicalist that some have indeed taken (Rudder Baker can be one case). However, these authors tend to believe in causal connections between mental events and behaviors and macro events. This means that the escape would work only if the mental-to-macro connection were not problematic and the macro was, as it seems to be, conceptually remote from the micro. This, as I have argued, is not the case: the macro world is causally closed in its physical description, and so the “usual worries about causal overdetermination arise” anyway.

But maybe the defender of the dual *explanandum* does not need that much. It is open for her to claim instead that the divergent causal chain stops somewhere after *B*-movements are instantiated.<sup>11</sup> To get the flavor of what I say, we can suppose that the claim is that the mental causal chain interrupts right after *B*-movements are instantiated. Or, to be more precise, the mental causal chain runs parallel to the physical causal chain only until *B*-movements and *A*-movements are instantiated; at that point, the mental causal chain stops while the physical causal

<sup>11</sup> Nothing that I am saying now and will say in the next paragraphs should be ascribed to Hornsby. I am only supplying possible responses to rebuttals she, or any defender of the dual *explanandum* strategy, may receive.

chain goes on. This sounds like epiphenomenalism about *B*-movements, but it is open for the anti-physicalist to say that *B*-movements can in turn cause mental events by being perceived.<sup>12</sup> Some anti-physicalists would not want to hold exactly this position. Rather, they would claim also that *B*-movements have further non-physical effects outside the body of the subject (see the references to the “extended mind” account below). The response to Matthews’ challenge, however, should be, *mutatis mutandi*, the same I have just offered.

Now the condition (ii). *B*-movements, as will be explained, must be essentially different from *A*-movements, or at least be conceptually remote from them. What this condition establishes is that this difference in concepts cannot be obtained simply by resort to their different causal antecedents. The difference between *B*-movements and *A*-movements cannot consist merely in that *B*-movements are caused by mental events while *A*-movements are caused by neural events. This would beg the question, for what we are trying to establish is precisely whether or not actions or mental events cause something different from *A*-movements. We cannot start by claiming that they do.

What about the condition (i)? *A*-movements cannot be sufficient for *B*-movements and *B*-movements cannot be essential for *A*-movements. That is, neither the argument from generalized exclusion nor the Cause-and-Essence principle should

<sup>12</sup> Precision is needed here. Epiphenomenalism can be taken as claiming that (a) some properties have no causal powers or (b) some properties, or some instantiations of such properties, do not cause anything whatever. What I say about *B*-movements excludes epiphenomenalism in sense (a) but is compatible with their being epiphenomenal in sense (b). Now a defender of this approach ought to be interested primarily in avoiding the (a) sense, for that kind of epiphenomenalism is a sure road to elimination. The (b) sense, though apparently threatening, can be handled more easily. It is not a viable position if beliefs and desires are concerned, for we cannot but hold that beliefs and desires are actual—not merely potential—causes. But we do not have an equivalent strong intuition about the causal efficacy of *B*-movements. We know that our beliefs and desires cause behaviors, but we cannot say that we know that our behaviors cause further things. So the defender of the strategy has no obligation to prove that *B*-movements cause, even though the threat of elimination commits her to a defense of their causal potency. (I thank Manuel García-Carpintero for pressing this point.)

be applicable. Or, at least, we need reasons to be skeptical about their applicability. If the first situation obtains, the success of the dual *explanandum* strategy against the argument from exclusion would be complete. Otherwise, it would be partially successful and only in its negative claims, as the discussion of Sturgeon's paper has made clear.

So something like this is what one may want to hold: Bodily movements are to be approached from two different perspectives. On the one hand, they have to be studied as simple movements of a physical body; on the other, they are movements of a human intentional body. Of course, it must be acknowledged that there is a relation of composition between one kind of movements and the other, but it can be said that physical movement composition does not exhaust intentional movements. As happened before with micro and macro movements, it can be said that the movements of a physical body have some features—provided by physical theory—out of which it is difficult to build the defining characteristics of the movements of a human body, for instance (see below) intentionality, expressiveness, and even consciousness of a kind.

I am sure that many philosophers—even many with dualistic intuitions—will find any account complying with conditions (i–iii) (specially with condition (iii)) highly implausible. Nonetheless, I want to close this paper with a short discussion of some ideas normally introduced as interpretations of the last developments in cognitive science that may respond to the demands here described.

Authors who are close to the field of robotics, or to that which is usually known as “situated cognition”, speak about the necessity to think of the mind as something essentially embodied and embedded.<sup>13</sup> “Classical” cognitive science and philosophy of mind, according to these authors, have been misled by the “computer's analogy” and have construed a mistaken view of

<sup>13</sup> There is a huge literature dealing with the embedded and embodied mind. Clark 1997 is a good summary of recent developments in these fields (robotics and situated cognition). The volume of *Cognitive Science* 1993, vol. 17, no. 1) dedicated to the discussion of situated action *vis-à-vis* symbolic interpretations is also a good introduction to the topic.

the human mind. The minimal claim of the defenders of situated cognition is that classical cognitive science has not taken into account the role of the body and of the environment in shaping the human mind. A lot of in principle very demanding processes can become rather simple thanks to our having a body by which we can act on the environment. Think of playing Tetris solely “in your mind” instead of helping yourself with your hands. Moreover, if you think about it, you will realize not only the importance of our having a body, but also the way we use the environment—in this case, the screen—in order to make our (mental) life simpler.

Now some of these authors—Andy Clark (1997) for instance—draw a parallel between their views and those of the phenomenologist Merleau-Ponty. According to this philosopher, and also most probably to the late Husserl, the human body has some distinguishing and essential features that cannot be explained from a physical point of view. The human body is expressive, and it has what Merleau-Ponty calls operative and pre-reflective intentionality (these are concepts related to the idea that the body is capable of understanding the world without having to make use of representations). The human mind is essentially embodied, and the human body is essentially mental in that it is the origin, condition of possibility and expression of our mental lives. What has to be kept apart are not minds and bodies, but *physical* bodies and *human* bodies.

These ideas about the human body can perhaps be made to fulfill the conditions for *B*-movements explained above. First, (i) it may be argued that there is no reason why *B*-movements so conceived should interact with the physical world while, at the same time, they may well be able to cause other mental events by being perceived. That is, the movements of the human body are not causes of anything that goes on in the physical chain: all those events are due to the movements of the *physical* body. On the other hand, the movements of the human body cause something, since they are perceived (and directly perceived as such) by other subjects. This idea seems to fit Clark and Chalmers’ (1998) theory of the extended mind (see

also Wilson's (1995) "wide computationalism"). According to this idea, the individual mind is not confined to "skin and skull" but rather involves parts of the environment and other minds: the environment contains pieces of information that the mind does not interpret, but directly perceive as such. This information contained in the environment that comes to the mind "as it is" can be said, according to these authors, to be part of the mind.<sup>14</sup>

Also, (ii) these movements would not be different from *A*-movements just in their etiology. Their difference would be due to their different intrinsic properties. Finally, (iii) a conceptual Grand Canyon would keep them safe from being reduced to a composition of physical events. A way to illustrate the latter is by pointing to the distinction between reflexes and intentional movements. Bodily movements which are reflexes may be identical, as far as *A*-movements are concerned, to bodily movements which are not reflexes, that is, that are *B*-movements. Therefore, *B*-movements do not supervene on *A*-movements. So the situation here diverges from that of macro to micro relations: it is not just that we have reasons to be skeptical about the truth of supervenience claims. It is that we know them to be false. The same goes for essence relations: we know intentional movements are not essential to physical movements and vice versa.<sup>15</sup>

<sup>14</sup> Hutchins 1995, for example, proposes to study the case of what goes in a cockpit in a commercial airliner. His thesis is that the cognitive tasks involved in flying an airplane cannot be ascribed only to the pilot. The mind responsible for the flight consists of the pilot's representations plus some other representational devices providing information that the pilot directly perceives as providing such information. My own view is that this is all unnecessarily complicated. Human minds use bodies and environments to make cognitive processes as simple as they can, but it is quite another thing to say that the mind is partly in the environment, or even that the individual has a direct access to information contained in the environment.

<sup>15</sup> There is yet a third point of divergence between the micro-macro case and this: in a reading of what Sturgeon says about it, there is no *inducing* relation going upwards here. As Sturgeon explains it, it seems necessary that there is weak supervenience for events of a family to induce, via causation and composition, events belonging to another family. This, so the examples

I am not really sure whether these ideas exactly fit within the profile conditions (i–iii) have drawn or whether they can be made to come together. It is clear to me, however, that they should, because if any of the conditions fails, physicalism follows, i.e. the denial of any of them entails the denial of all. So my conclusion is that the dual *explanandum* strategy is not incoherent, and perhaps there is a way to make it work. However, I find the strategy is extremely demanding, and (*prima facie*) rather counterintuitive when its implications are spelled out in some detail.

## REFERENCES

- Antony, Louise, and Joseph Levine, 1997, “Reduction with Autonomy” *Philosophical Perspectives*, 11, pp. 83–107.
- Baker, Lynne Rudder, 1995, *Explaining Attitudes*, Cambridge University Press, Cambridge.
- Burge, Tyler, 1993, “Mind-Body Causation and Explanatory Practice”, in John Heil and Alfred Mele (eds.), *Mental Causation*, Clarendon Press, Oxford.
- Clark, Andy, 1997, *Being There: Putting Mind, Brain and Body Together Again*, MIT Press, Cambridge, Mass.
- , and David Chalmers, 1998, “The Extended Mind”, *Analysis*, 58, pp. 7–19.
- Crane, Tim, and Hugh Mellor, 1990, “There is no Question of Physicalism”, *Mind*, 99, pp. 185–206.
- Dretske, Fred, 1988, *Explaining Behavior: Reasons in a World of Causes*, MIT Press, Cambridge, Mass.
- Hornsby, Jennifer, 1997, *Simple Mindedness: In Defense of a Naive Naturalism in the Philosophy of Mind*, Harvard University Press, Cambridge, Mass.
- Hutchins, E., 1995, “How the Cockpit Remembers its Speed”, *Cognitive Science*, 19, pp. 265–288.
- Kim, Jaegwon, 1998, *Mind in a Physical World*, MIT Press, Cambridge, Mass.
- , 1993, *Supervenience and Mind*, Cambridge University Press, Cambridge.

of reflexes show, is not the case for *A*-movements and *B*-movements. So the relation seems even weaker than induction.

- Levine, Joseph, 2001, *Purple Haze: The Puzzle of Consciousness*, Oxford University Press, Oxford.
- Malcolm, Norman, 1968, "The Conceivability of Mechanism", *The Philosophical Review*, 77, pp. 45–72.
- Matthews, Robert J., 1998, Review of *Simple Mindedness: In Defense of a Naive Naturalism in the Philosophy of Mind* by Jennifer Hornsby, *Mind*, 107, pp. 890–894.
- Noordhof, Paul, 1999a, "The Overdetermination Argument versus the Cause-and-Essence Principle —No Contest", *Mind*, 108, pp. 367–375.
- , 1999b, "Causation by Content?", *Mind and Language*, 14, pp. 291–320.
- Papineau, David, 1990, "Why Supervenience?", *Analysis*, 50, pp. 66–71.
- Peacocke, Chris, 1979, *Holistic Explanation*, Clarendon Press, Oxford.
- Poland, Jeffrey, 1994, *Physicalism*, Clarendon Press, Oxford.
- Sturgeon, Scott, 1999, "Conceptual Gaps and Odd Possibilities", *Mind*, 108, pp. 377–380.
- , 1998, "Physicalism and Overdetermination", *Mind*, 107, pp. 411–433.
- Vicente, Agustín, 2001, "El principio del cierre causal del mundo físico", *Crítica*, vol. 33, no. 99, pp. 3–17.
- Wilson, Robert A., 1995, *Cartesian Psychology and Physical Minds*, Cambridge University Press, Cambridge.
- Yablo, Stephen, 1992, "Cause and Essence", *Synthese*, 93, pp. 403–449.

*Received June 6, 2002; revised September 12, 2002; accepted September 23, 2002*