

# Two Myths of Psychophysical Reductionism

Ricardo Restrepo

Escuela de Constitucionalismo y Derecho, Instituto de Altos Estudios Nacionales, Quito, Ecuador  
Email: [ricardo.restrepo@iaen.edu.ec](mailto:ricardo.restrepo@iaen.edu.ec), [ricardo.restrepo@yahoo.co.nz](mailto:ricardo.restrepo@yahoo.co.nz)

Received January 30<sup>th</sup>, 2012; revised February 22<sup>nd</sup>, 2012; accepted March 9<sup>th</sup>, 2012

This paper focuses on two prominent arguments claiming that physicalism entails reductionism. One is Kim's causal exclusion argument (CEA), and the other is Papineau's causal argument. The paper argues that Kim's CEA is not logically valid and that it is driven by two implausible justifications. One is "Edward's dictum", which is alien to non-reductive physicalism and should be rejected. The other is by endorsement of Papineau's conception of the physical, immanent in Papineau's causal argument. This argument only arrives at the physical property-property identities by using a conception of the physical that licenses anything to be reductively physical, including putative core anti-physical entities; thus, leaving Papineau's causal argument and Kim's CEA without a reductive physicalist conclusion of philosophical interest.

**Keywords:** Causal Exclusion; Mental Causation; Physicalism; Supervenience; Reduction

## Introduction

Jaegwon Kim and David Papineau have each developed influential arguments to justify reductive physicalism: the *causal exclusion argument* (CEA) and the *causal argument*, respectively. Both approaches purport to show that physicalism entails reductionism. The thesis of reductive physicalism is that each mental property is itself identical with a physical property. There will be more to say in this paper about the nature of this claim, of course.

Using causal considerations, reductionists argue that physicalists can't be "soft and cuddly non-reductionists" (Melnik, 1995: p. 370, 2008), rejecting with their characteristically cool rationality the "myth of non-reductive materialism" (Kim, 1989), and with stoic discipline "swallow" the "unpalatable implications of our assumptions and presumptions" (Kim, 2010: p. 104). The reductionist pill is not without its rewards, however, for its conclusions are of "great philosophical interest" (Papineau, 2002: p. 41). Or so we are told.

Despite the mentioned arguments to the contrary, physicalism, I think, does not entail reductionism.<sup>1</sup> In this paper, I begin by looking at the premises under which Kim's argument is formed. Kim's intention is to reduce non-reductive physicalist theses to absurdity, and derive reductionism from the apparent contradiction. I present a basic non-reductive physicalist reading of the premises, however, and find no such contradiction. The idea is not to argue for one particular way in which to be a non-reductive physicalist, but to show the existence of the logical space available to this position, compatible with the premises. The existence of this logical space has the consequence that the causal exclusion argument is not logically valid. The premises may all be true but the conclusion false.

Because one way Kim (2005) justifies his reductive physicalism is by endorsing Papineau's (2002) conception of the physical, it will be convenient, from an expository point of view,

to address Papineau's causal argument for reductionism in the process of addressing Kim's. What I think is Papineau's myth of psychophysical reductionism is embedded in Kim's argument. I argue that Papineau advocates a position that does not in substance contradict non-reductive physicalism, because it allows for higher-level causal properties, and these are precisely what the non-reductionist thinks mental properties are. The deeper reason, however, is that Papineau's position licenses calling any metaphysics "reductive physicalist", including positions which endorse the existence of things clearly outside physicalist metaphysics, so it is trivial. Thus, Kim's attempt to use Papineau's conception of the physical in the CEA to yield the reductionist conclusion is unsuccessful, since using that conception still does not generate the required contradiction between the premises.

The other way Kim justifies reductionism is by endorsing Edward's dictum, which says that basic properties exclude non-basic properties from making causal contributions (Kim, 2005: p. 36). This justification fails, I argue, for several reasons. Edward's dictum is not part of the official CEA. Kim does not identify it as a premise, but rather introduces it as an aside, even though he acknowledges that it is the driver of the reductionist conclusion of the CEA. Edward's dictum, however, is not part of non-reductive physicalism and consequently cannot be taken to be an internal generator of the required contradiction between the non-reductive physicalist premises. Edward's dictum, I add, is in serious need of justification.

## Physicalism and the Causal Exclusion Argument

Kim's CEA is well-known. It purports to take the form of a *reductio ad absurdum* of non-reductive physicalism. Thus, typical theses of non-reductive physicalism are identified, a contradiction is said to be found, and a premise is rejected. Kim opts for rejecting the thesis which distinguishes the non-reductionist from the reductionist: the one which claims that mental properties are not reducible to physical properties. The premises of the argument are as follows:

<sup>1</sup>Melnik's reductionism is outside the scope of this paper, though it is certainly a view worth careful analysis.

**Supervenience:** Mental properties supervene on certain physical properties. That is, if any system *s* instantiates a mental property *M* at *t*, there necessarily exists a physical property *P* such that *s* instantiates *P* at *t*, and necessarily anything instantiating *P* at any time instantiates *M* at that time.

**Irreducibility:** Mental properties as such are not reducible to the physical properties on which they supervene.

**Mental Causation:** Mental properties have causal efficacy and relevance—that is, mental properties and their instantiations causally contribute to bring about other events with their properties.<sup>2</sup>

**Exclusion:** No single event can have more than one sufficient cause occurring at any given time—unless it is a genuine case of overdetermination.<sup>3</sup>

**Closure:** Each physical event has a sufficient physical cause (insofar as causes of events are sufficient) (Kim, 2005: pp. 34-43).<sup>4</sup>

The causal argument says that since each physical effect has a sufficient physical cause, each mental cause must itself be physical (Papineau, 2002). The premises of the CEA can be seen to be those of the causal argument for physicalism, plus non-reductionist views of mental causation. If indeed a contradiction were to be found, then physicalism would be incompatible with non-reductionism, and reductionism would be a consequence of physicalism.

Suppose Sally picks up a glass of beer, which by **Mental Causation**, her intention to do so causes. By **Supervenience**, her intention is supervenient on a physical property. By **Irreducibility**, her mental intentional property does not reduce to a physical property upon which it supervenes. By **Closure**, there was a sufficient physical cause for her picking up of the glass of beer. Now **Exclusion** rules out the mental from making a causal contribution on pain of degenerate overdetermination. So **Mental Causation** can be true only if some other premise is false. Kim's final solution is to reject **Irreducibility**, and with it goes non-reductive physicalism. Physicalism must consequently be reductive.

The argument seems compelling, but I believe it is misleading. One should first note that it is to be expected, given physicalism, that mental properties, like any other property of our world, must be physical in some significant sense, since every property inhabiting our world is physical. So, mental properties are identical with certain physical properties. This follows just from the fact that our world is completely physical and that mental properties have identities, and consequently, that every property is completely physical. The bare statement identifying mental properties with physical properties should itself be unimpressive.

<sup>2</sup>In a recent reply to Shoemaker's non-reductive physicalism, Kim (2010, p. 107) says that "For a positive resolution of the mental causation problem (the CEA), I believe we should reach the statement that mental properties are causally efficacious as a conclusion, not start with it as an assumption. What troubles me is that Shoemaker's procedure seems the opposite; for him mental causal efficacy is a 'starting assumption', as he puts it" (parentheses added). Notice that this shifts the goal posts of the CEA. For the problem of mental causation is supposed to be one which arises from *within* non-reductive physicalism, a view which, as Kim himself formulates it, includes the claim that mental properties have causal efficacy.

<sup>3</sup>It seems correct that systematic overdetermination could not be genuine, as Kim supposes.

<sup>4</sup>The parenthesis is to make sure Closure is compatible with the possibility of quantum indeterminism.

There is a standard distinction to be made between the two ways in which a property can be physical. One is that it belongs to the realm of fundamental physics. Call these *physical<sub>1</sub> properties*. Various branches of physics are called in to inform us of what that realm is like. It should be noted that physical science can also posit and be interested in non-fundamental physical properties. For example, astrophysics deals, amongst other phenomena, with non-basic objects such as stars and galaxies, which are governed by aggregates of arranged *physical<sub>1</sub> properties* and laws. Such arranged aggregates of the basic constituents are also physical. There is no part or aspect of them which is not physical. Distinctive properties of such arranged aggregates of fundamental elements, to mark the distinction, are *physical<sub>2</sub>*.

Jackson (2006) provides a useful illustration of this distinction by use of an analogy with geometry, and identifies the higher level sciences as gatherers of information about aspects of *physical<sub>2</sub> properties*.

We need... an extended sense (of the physical) because the patterns that economics, architecture, politics and very arguably psychology, pick out and theorise in terms of, include many that do not figure in the physical sciences. The reason is no mystery: it is that aggregation creates new properties... because aggregations fall under patterns, kinds, etc. that the items they are aggregations of do not fall under... Physicalists must allow that the world contains aggregations that have properties that are not *physical<sub>1</sub> properties* for the same reason, when all is said and done, that someone who holds, rightly, that a triangle is an aggregation of straight lines must allow that the triangle is not itself a straight line (Jackson, 2006: p. 234, parentheses added).

It appears attractive to hold that at least some of the properties and instances physics aims to discover are the most basic of all, and are regulated by a set of laws which allow things to be composed of them in various ways and not in others, and that those higher level objects and the causal patterns they fit into are tracked by higher level sciences. As Jackson says, with good reason, mental properties are of the higher level *physical<sub>2</sub> type*.

This is surely a physicalist view. Dualist arguments hinge on the supposed lack of *de re* logical entailment from *physical<sub>1</sub> properties* to mental properties, such as consciousness. For example, Chalmers (1996) argues that the very possibility of zombies, creatures which are physically identical with humans but do not have conscious experience, refutes physicalism. In such a case, consciousness would be neither intrinsically *physical<sub>1</sub>* nor *physical<sub>2</sub>*. Having the relevant physical properties a being with consciousness has would not be logically sufficient for being conscious. To duplicate the conscious experience of seeing bright red or feeling a sharp pain, one would need to add something not present in beings merely physically identical to beings who have the experiences. Non-reductive physicalism is incompatible with this view. It would say that indeed, any minimal physical duplicate of our world includes the conscious experience of seeing red and of feeling pain, to use Jackson's (1994) schema. Both the experiences of seeing red and of feeling pain are *bona fide* physical properties present in our world.

What the non-reductionist denies is that *physical<sub>2</sub> properties* (of which mental properties are an example) and *physical<sub>1</sub> prop-*

erties per se are identical. Reductionism requires mental and physical<sub>1</sub> properties per se to be identical. It is only when properties as such are identical that a property reduces to “another”; not when they merely coincide nor when they have a contingent identity relation. The minimal, but sufficient, claim of the non-reductionist is that the relation between mental properties and physical<sub>1</sub> properties as such is not identity. There is a family of ways this non-reductive relation can be. Baker (2002) and Pereboom (2002), for example, have claimed that the relation is constitutional coincidence without identity. A decision coincides constitutionally with the instantiation of certain physical<sub>1</sub> properties at a time, though they are not the same properties. This is in the same way in which, in terms of properties, being the Ship of Theseus coincides constitutionally at a time with having certain physical<sub>1</sub> properties. In terms of tokens, Pereboom (2002) argues that the fact that the Ship of Theseus and the basic physical entities of which it is composed at a particular time have different temporal and modal properties is enough to establish that the two items are not reductively the same. It should nevertheless be clear that the Ship of Theseus is a perfectly physical entity all of whose properties are completely *bona fide* physical properties in the context of physicalism. Similarly, physical<sub>1</sub> properties might coincide constitutionally with the property of being a decision, but have temporal and modal differences sufficient for them not to be identical with each other.

Constitutional coincidence without identity could be extended to events, like the sinking of the ship. That token event has many smaller lower-level events composing it: the many sinkings “experienced” by its parts. Are the part-sinkings identical with the ship’s sinking? According to the constitutionalist, they are not. For one, that same ship-sinking could have been composed of different part-sinkings, for example, if there were certain repairs done on it before leaving port, but which did not actually take place. Thus, the sinking of the ship and the part-sinkings have different modal properties. In this sense, the event that is the sinking of the ship is distinct from the events that are the part-sinkings. The same reasoning for non-identity that applies to the property and token case, applies to the event case.

However, one of the ways in which higher level objects are seen to be distinct from lower level objects is that they, at least sometimes, maintain their identity past a time when their former constitution does not, or vice versa. In the event case, this might seem to be pre-empted by the apparent fact that events are dated particulars and that the lower level events coinciding with the higher level event have the same temporal coordinates essentially.

This apparent fact, constitution theorists may contend however, is not really a fact. It is a live possibility that in fact the temporal coordinates of lower level constituting events do come apart from the coordinates of higher level constituted events in that sometimes the sinking of the ship and the part-sinkings coincide in time only partially. Such is the case of a ship that is being sunk and is fully submerged but has not reached the bottom. As the ship descends a powerful bomb detonates within it and its parts scatter, destroying the ship, but its parts survive and continue to sink. The part-sinkings have temporal coordinates that outstrip the temporal coordinates of the ship’s sinking. Thus, temporal properties for lower level and higher level events can differ as well, and the reason for advancing a relation between lower level and higher level events to be constitution and not identity would hold.

Another way in which the relation between mental and physical<sub>1</sub> properties might be non-reductive is through contingent identity. Reductive identities are stronger than contingent identities. A lump of clay might be identical with a statue (Gibbard, 1975), but being a statue does not reduce to being a lump of clay. A lump of clay and a statue, if identical, are contingently identical. Being a statue and being a lump of clay are two distinct properties, which at times are had by the same object, but not always, and sometimes the same statue can change composition and the same lump of clay can stop being a statue. One may wonder whether this is merely a perspective game. From the perspective of snake or bat, the lump of clay and the statue may not seem to be so different. This may well be true, but it does not prove anything. Bats and snakes are not sensitive to many facts humans are sensitive to (and we are probably insensitive to others they are sensitive to). For instance, Carbon and Iron may not seem so different from the perspective of a bat or a snake, but certainly these are two chemical kinds. The same goes for many such things human legitimately differentiate and other species do not. Of course, one could subsume statues under the category of lumps of clay. A statue might just be “a lump of clay that has been artistically worked upon.”<sup>5</sup> However, while it is true that some statues have this property, the point is that they have this property contingently. One such statue might have been, instead, a lump of metal that has been artistically worked on, for instance, and certainly, being a lump of clay that has been artistically worked on is a distinct property from being a lump of metal that has been artistically worked on. Not all statues are lumps of clay. The point is that these two properties could be had by identical objects, but since they can and sometimes do come apart, being an object with both these properties is a contingent fact—not one that reduces statues to lumps of clay.

To take another example, suppose the winner of the race is identical with David (Lewis, 1972). Does the property of being the winner of the race reduce to the property of being David? Are the two properties per se identical? To the non-reductionist, asserting that they are is untenable. These are two properties of the same object, and they are both physical in a sense compatible with physicalism. An identity between the properties as such does not have to be asserted in order to keep being a physicalist. It takes different things to be David and to be the winner of the race, though David satisfies both. Being David and being the winner of a race should not pose a problem for physicalist metaphysics.

There are also “closer” contingent identities, short of reduction, namely contingent property identities. Such cases fall into a generalized version of the **Supervenience** schema nicely. The general supervenience schema says:

**General Supervenience:** Physical<sub>2</sub> properties supervene on certain physical<sub>1</sub> properties. That is, if any system *s* instantiates a physical<sub>2</sub> property *M* at *t*, there necessarily exists a physical<sub>1</sub> property *P* such that *s* instantiates *P* at *t*, and necessarily anything instantiating *P* at any time instantiates *M* at that time.

For analogy, the property of weighing more than seven kilograms supervenes on the property of weighing ten kilograms, and even though these are two properties, they can be co-instantiated in a single token (Macdonald & Macdonald, 2006). Furthermore, one might say, *x*’s having the property of weigh-

<sup>5</sup>Thanks to a reviewer at the *Open Journal of Philosophy* for these objections.

ing more than seven kilograms is identical with  $x$ 's having the property of weighing ten kilograms. These are two different properties in themselves, but the properties might be identical in certain instances. Consider another example: an animal's being a mammal *is* its being a dog, though by no means can we say that being a dog and being a mammal is the same property *per se*, which is what reduction requires. Copeland (2000) develops a formalism for this Aristotelian conception of certain property relations. This kind of non-reductive identity can make sense of Kim's claims emphasizing that in particular instances, mental and physical properties might be identical, but shows that more is required for reduction.

Reductive identities require complete property coincidence. This is, at least in part, how one property *per se* is identical to "another", and this is a requirement the reductionist does not deal with very well because of the phenomenon of multiple realizability. Take Putnam's (1975a) twin earth example, that if someone came back from twin earth bringing the transparent liquid in twin earth's rivers with the XYZ chemical structure, this substance would not be water. The exciting original finding was that water seemed to be reductively identical with  $H_2O$  because it could not be identical with something without this chemical structure. Water cannot be XYZ (Funkhouser, 2007).

Or so it seemed. Of course, if XYZ is heavy water,  $D_3O$ , then water could be something other than  $H_2O$ . This, however, only shows that water is not necessarily identical to  $H_2O$ , and consequently not reductively identical to  $H_2O$ , even though, again, in particular instances, it might be that a particular sample's being water is its being  $H_2O$ .<sup>6</sup>

Of course, to this argument Kim responds with the "disjunctive move" (Kim, 1998, 2005; Jaworski, 2002). If  $M$  is multiply realizable by  $P_1$  and  $P_2$  (and only those) then  $M$  is reductively identical to  $(P_1 \text{ or } P_2)$ . The disjunctive move is that mental properties are reductively identical with the disjunction of their possible realizers. This is a remarkably *ad hoc* move in need of justification in my view.

One can say that having a negative charge is disjunctively identical with having a negative charge before yesterday or today or after, for instance. But a move such as this one is baseless. One wants to ask, what is the basis of the disjunctive move? The fact that it *may* theoretically be performed certainly does not justify that it *be* performed. Physical<sub>1</sub> properties, such as having a negative charge, are not disjunctive in a compelling way, and for the same reason physics<sub>2</sub> properties, of which mental properties are a subset, are not compellingly disjunctive. Though mental properties can be broken up in a disjunction just as the physics<sub>1</sub> properties can, there is little by way of motivation to think this way. On the contrary, such moves miss what is common to all the possible realizations in virtue of which they fall under the same kind, and in virtue of which they have similar effects.

The sufficient physical causes relevant to **Closure** are basic properties and instances physical science aims to uncover, and what they logically amount to (physical<sub>2</sub> properties). Thus, for example, we can know that triangles logically result from putting lines (putatively basic properties in geometry) together in certain admissible ways. This is where Putnam's (1975b: p. 296) well-known peg example is relevant: it identifies geometrical properties in a peg, let's say, being a triangle of particular size, which determine the peg's causal capacity to fit in certain holes

while other pegs of a different shape and size do not. The basic physical properties arranged in a certain way logically add up to the object's being a triangle (a physical<sub>2</sub> property), which corresponds to certain capacities which contribute to certain of its actual and potential activities.

So **Closure** is fine as long as we know it counts in physical<sub>2</sub> properties, and this opens the door for mental properties to be of this kind. Spin, charge, and the rest of the physical<sub>1</sub> properties make a contribution, but capacitate the peg to go through certain holes only when they add up to being a triangle within a range of certain sizes.

Kim frequently refers to Alexander's dictum to support his case (e.g. 1998: p. 119). This dictum says that only causal properties exist or should be believed to exist. But if we agree that physical<sub>2</sub> properties such as having a particular macroscopic shape exist, then we see that by Alexander's dictum we should believe in the causality of such properties. Of course, Kim says that he does but in a reductive way. If all this amounts to is saying that, to take the case of the peg for concreteness, that a triangular peg is identical to physical<sub>1</sub> properties arranged as a triangular peg of a certain size, then who could disagree? The condition that the physical<sub>1</sub> be arranged in certain ways, however, is exactly, according to the non-reductionist, what raises the situation to one where physical<sub>2</sub> properties are instanced, enabling new causal patterns of interaction.

There is famous complication here (Dretske, 1988). A singer shrieks "break" and a glass breaks. Now, the sufficient physical cause of the breaking includes the amplitude and frequency of the generated sound waves. However, in this case, the sufficient physical cause, by non-reductionist standards, appears to include the semantics of "break". But this clearly is not a property that makes a causal contribution to the breaking. The *qua* problem remains, it might seem.

But in my view this is not the case. The proposal of this paper involves the idea that sufficient physical causes are reliably tracked through the sciences using the experimental method. With the experimental method, we can manipulate independent variables to see their effects. Thus, take Dretske's case to be representative of the control group. A word, with a certain semantics, pitch and volume, is shrieked, and the glass breaks. In another case, we can vary the semantics (the independent variable) while keeping the pitch and volume fixed. For example, the singer might agree with her audience that when she shrieks "break" it will mean *that it is time to take a seat*, and she might faithfully intend to mean that proposition when she does. The glass would break without dependence on the independent variable (the semantics). Alternatively, she might decide to shriek "Amen" with the same pitch and volume. Again, the glass would break without dependence on the semantics of the word, thus confirming that this is not a part of the sufficient physical cause.

Analogous experiments yield different scientific results with respect to mental properties generally, however: their causal contributions are detected. Their manipulation results in observed differences. Consider Sally's intention to pick up the beer again. It is a physical<sub>2</sub> property, the logical result of certain physical<sub>1</sub> properties put together in a certain way. Furthermore, it is verifiable that without the intention, Sally would not have picked up the glass and that by varying physical<sub>1</sub> properties within the intention's range of possible realizations, the same result of picking up the glass comes about. However, if you vary the independent variable, that is, Sally's having the inten-

<sup>6</sup>For another argument that "water =  $H_2O$ " does not express a necessary identity see Barnett (2000).

tion to pick up the glass of beer, another effect will occur, thereby verifying the causal contribution of her mental properties.

This is the non-reductionist reading of the premises of the CEA, and, as we have seen, under such a reading there appears to be no contradiction and the argument seems, consequently, to not be logically valid. We might then conclude that since it is the non-reductive physicalist position which was supposed to be reduced to absurdity, it is the non-reductionist reading that counts. However, while this is true to a significant extent, it is not complete. For it might be that there is some other reading which we somehow nevertheless *ought* to give the theses. So it is worth seeing what other reading Kim suggests. One of his two suggestions involves Papineau's conception of Closure and the causal argument, and the second suggestion involves Edward's dictum, the additional unofficial assumption of the CEA, mentioned earlier. I consider these in turn.

### Papineau's Reductive Physicalism

Kim (2002: p. 671) indicates that reductive physicalism is the only consistent and robust form of physicalism there is. However, he considers many things physical which are not explicitly a part of the theoretical universe investigated by fundamental physical theory, such as neural activity, outward bodily movement, and so on. All these things are physical<sub>2</sub> phenomena, which the ideas raised thus far place on a par with mental phenomena.

So what could Kim mean by "sufficient physical causes"? Kim (2005: p. 43) tips us off by pointing to Papineau (2002). The first thing to notice is that Papineau (2002: Ch. 2) claims to be committed to a reductive form of physicalism, so it is questionable that Kim should be able to say that he is reducing premises representing non-reductive physicalism to absurdity, when his reading of the premises are exactly those of a self-proclaimed reductionist instead of a non-reductionist. There is a whiff of circularity in this move.

Nevertheless, Papineau's (2002) physicalism does not entail that "sufficient physical causes" must be non-mental.<sup>7</sup> We will see that some of the principles at work in Papineau's account allow for mental properties to do irreducibly mental causal work, and thus are able to be referred to as "sufficient physical causes". Papineau's (2002) physicalism asserts the two following components. First, the physical is to be understood as what is inorganically identifiable, or "identifiable non-mentally-and-non-biologically". That is, the physical just is what can be referred to "independently of this specifically mental conceptual apparatus" (Papineau, 2002: p. 41). If physicalism is true, this means that those concepts which do not operate in terms like *seeing* and *believing* are sufficient to refer to anything that exists in nature and participates causally in the world.

The second part is the supposedly reductive part, advancing

<sup>7</sup>In places other than Papineau (2002), he argues that the physical is the non-*sui-generis*-mental (Spurrett & Papineau, 1999; Montero & Papineau, 2005). It is worth pointing out here that this account is also compatible with the non-reductionist idea that mental properties are organized aggregates of basic physical properties, which can be referred to by "sufficient physical causes". For a property to be non-*sui-generis*-mental, and consequently physical by the standards of these theorists, it is sufficient that it be completely determined by basic physics. This is a condition that the mental satisfies. This is not a thesis I treat further here because it is different from the one Papineau advocates by himself and from the one Kim alludes to. Both theses are addressed by Restrepo (2012).

the idea that "materialism is to be understood as a matter of property identity" between mental and biological properties, on the one hand, and material properties, on the other. Mental properties, and particularly, "conscious properties are identical to material properties—that is, they are identical either to strictly physical properties, or to physically realized higher properties" (Papineau, 2002: p. 47). But this, of course, can easily be fit into the physical<sub>1</sub>/physical<sub>2</sub> scheme, with the non-reductionist result.

Recall that the causal argument says that since every physical effect has a sufficient physical cause (with the fine print about quantum indeterminism), every mental cause must itself be physical. By what we have seen in the previous section about the conception of the physical, this is not itself a reductionist position. But let's suppose we insert Papineau's two claims. Then, the causal argument looks like this: since every non-mentally-and-non-biologically identifiable effect has a non-mentally-and-non-biologically identifiable cause, mental causes are non-mentally-and-non-biologically identifiable. In fact, the reasoning continues, supposing such a mental cause to be a property, then that property is identical to one which is non-mentally-and-non-biologically identifiable.

This conception of the physical provides an interpretation of the premises of the CEA without generating a contradiction. Mental properties are non-mentally-and-non-biologically identifiable. Take Sally's intention, for example. It is identifiable as *a state resulting from an arrangement of fundamental particles in a particular space-time region*. But this fact certainly does not imply that the intention does not cause, which is what is required for a contradiction.

Furthermore, that the mental is identical with the physical, in his sense, is what Papineau thinks is of "great philosophical interest" (2002: p. 41). Given that conception of the physical, it is of course easy to derive the thesis that any property or cause, including mental ones, will be identical with properties or causes thus identifiable. However, this conception is not recommendable. For, how significant is it to say that one is a reductive physicalist when any world would qualify as such? Any putative property or entity I can think of, physical or not, I can identify non-mentally-and-non-biologically. Let us consider things which are clearly outside the physicalist worldview: God, emergent properties, Cartesian souls, and Platonic numbers. Supposing theism is true, God is non-mentally-and-non-biologically identifiable as *the thing that created the universe*. Supposing emergentism is true, non-physical emergent properties are non-mentally-and-non-biologically identifiable as *non-linear effects of certain arrangements of matter*. The immaterial Cartesian soul is non-mentally-and-non-biologically identifiable as *one of the things that interact causally with certain particles* (coincident with the pineal gland). The Platonic number eight is non-mentally-and-non-biologically identifiable as *the number of planets orbiting the Sun*.

By Papineau's standard, any possible world, including those containing core anti-physical things, are worlds where physicalism is true, since anything can be non-mentally-and-non-biologically identified. Thus, it follows that under his conception, "physicalism entails reductionism" is true, but it is trivially so. There is no philosophical interest in this assertion, since anything, including any property and any possible cause is non-mentally and non-biologically identifiable. Furthermore, given that any property can be identified in this way, then any world with any kind of property is a world where reductive physical-

ism is true. But that position faces a choice between having reductive physicalism be trivial and rejecting the proposed conception of the physical (Restrepo, 2012). The attempt to get reductionism from physicalism, by trivializing the notion of reductive physicalism, is Papineau's myth of psychophysical reductionism. This is one supposed avenue to reductionism that Kim endorses, but it is not the only one.

### Edward's Dictum

As mentioned before, Kim says that the idea that drives the CEA is Edward's dictum. This idea purports to generate a conception of sufficient physical causes that prevents **Closure** from including mental causes. Edward's dictum says that:

There is a tension between vertical determination and horizontal causation. In fact, vertical determination excludes horizontal causation (Kim, 2005: p. 36).

This is a very strong assertion. Assuming that supervenience is a vertical determination relation (as Kim does), together with **Supervenience**, Edward's dictum is sufficient to result in the conclusion that mental properties are causally excluded by the properties on which they supervene. The rest of the premises of the CEA would be superfluous. However, Kim does not formulate this assumption as a premise. Rather he acknowledges its strong and necessary role as something of an aside. But why formulate the rest of the premises of the CEA then, and not just have Edward's dictum together with **Supervenience** as the official argument?

The answer, I think, is that no non-reductive physicalist would endorse Edward's dictum—and not without reason—so it has to be smuggled in. It is worth noting that the dictum is not a proposition that looks attractive to, nor entailed by, non-reductive physicalism. Consequently, it appears that it cannot form a part of a successful *reductio* of non-reductive physicalism. In being a supposed *reductio ad absurdum* of non-reductive physicalism the argument was supposed to be driven by a tension arising from *within* that position, which we can now see it is not. But this is perhaps too quick a judgement to make. Perhaps this principle should in the final analysis be endorsed by non-reductive physicalists for some good reasons. Let's see if Kim's are.

### Being a Yellow Lump and Time

Suppose a lump is yellow at  $t$ . What determines this fact? One option is that the lump has a microstructural property  $M$  at  $t$  that vertically determines its being yellow. Another option is the lump's being yellow at  $t-\Delta$ , which we may suppose it was. Kim (2005) reflects that “[a]nything that happened before  $t$  seems irrelevant to the lump's being yellow at  $t$ ; its having  $M$  at  $t$  is fully sufficient to make it yellow at  $t$ ” (pp. 36-37), and concludes that because the lump's having  $M$  at  $t$  is sufficient for the lump's being yellow at  $t$ , that the lump's being yellow before  $t$  makes no causal contribution.

The initial problem with this idea is that Kim is considering determinative facts that happen at different times ( $t-\Delta$  and  $t$ ), and consequently the two options need not compete. In general, to each of two times in the history of a fact there will correspond a complete set of conditions that contribute causally to the determination of the fact in question. No conflict arises out of this supposition because we know from the beginning that

considering facts determinative of another do not necessarily compete if they happen at different times. Picture a horizontal column of standing dominoes. The domino falling at  $t-\Delta$  does not causally compete with the domino falling at  $t$ . It might be said that the falling domino at  $t$  is self-sufficient in the sense that if we erased its causal history, *while maintaining the falling domino at  $t$* , that time-slice of the domino would still exist. But this supposition should not lead us to believe that the falling of the domino at  $t$  was not caused by the prior event of another domino falling at  $t-\Delta$ . This is why the **Exclusion** premise in the CEA is time-relative—it says that there cannot be more than one sufficient physical cause at any one time, not in general.

To bring this out more, suppose a painter painted the lump yellow prior to  $t$  and you wish to know what caused the lump to be yellow at  $t$ . Then it would be quite causally relevant that a painter painted it prior to  $t$ , even if the lump's microstructural property  $M$  vertically determines this at the later time,  $t$ . Kim's view would say, to the contrary, that this is not really the case simply because the lump's having  $M$  at  $t$  is fully sufficient to make it yellow at  $t$ .

Kim's offered model has the implausible implication that a person's painting a lump yellow does not causally determine the lump's being yellow at a later time just because the lumps being yellow at the later time has a microstructural base. It should be noted that by this argument, there are no prior causes of the lump's being yellow at  $t$ . The lump's being yellow at  $t$  would be completely causally undetermined by prior facts just because it has  $M$  at  $t$ , which is implausible. By the same standard, the lump's having  $M$  at  $t$ , being self-sufficient at  $t$ , was not caused by prior facts. By extension, no physical fact at a time has prior causes. This, of course, is no longer a specific problem for mental causation, but expresses a general scepticism about causation, which is not supposed to be in question.

A common response to this argument is that the painter causes the lump to be yellow by causing a difference in its micro-structural base. While this may be true, the point is that the principle that Kim uses to justify Edward's dictum in order to drive the CEA implies that no properties of prior events are causally relevant to latter events. He just assumes, on the basis of somehow being characterizable as an isolated event at  $t$ , that no prior event caused  $x$ 's being  $B$  at  $t$ . If this applies to any event that is not the first event in the universe, then it applies all events. This causal eliminativist implication of Kim's reasoning is of course very unconvincing and therefore cannot provide support for Edward's dictum.

### Causing at the Same Time

Let's focus on a case that does not make the mistake of trying to pit two candidate causes from two times against each other. Take some physical effect  $P2$  which happens at  $t$  and the candidate physical cause  $P1$  which occurs at  $t-\Delta$  and the candidate vertically determined cause  $M1$  which happens at  $t-\Delta$ . Does  $P1$  exclude  $M1$  in the causation of  $P2$ ?

Let's test whether vertically determined properties are causally excluded by their realization bases. Consider the Boat: Suppose a boat has the property of being made out of metal arranged in a certain way ( $P1$ ). The boat's being made of solid material ( $M1$ ) is a supervenient property of its being made of metal. Being made of solid material is multiply realizable. The boat's property of being made of solid material could also be realized, for instance, by being made of wood. Notice the fact

that the boat's being made of solid material is a supervenient property with a realization base and that this does not rule this property out from being one of the causally contributing factors of floating. A causal power of being made of solid material is that it makes the boats that have this property float.

Is the boat's being made of solid material excluded by its being made of metal in the causation of floating? Intuitively, it is hard for it to even seem that it is or for one to think how or why one should believe it so. Reasserting Edward's dictum to support the exclusion at this point would be unconvincing, since it is this principle that needs explanation and justification in the first place.

One of the biggest "temptations", I think, in the debate about the CEA is to think that physical bases of supervenient properties are sufficient, *without the supervenient property*, to cause the putative effect of the mental: that P1, without M1, is causally sufficient for P2.

In order to evaluate this claim, one might consider the following. It is an indicator of X causing Z while excluding Y, that X be able to cause Z in the absence of Y. But this is exactly what X in this case is unable to do. The boat's being made of metal could not cause the boat to float unless the boat was made of solid material. It is beyond question that *ceteris paribus* the boat floats when it has the property made of metal and of solid material. However, suppose you minimally modified the boat so that you make the material of which it is made not solid. You could do this by subtracting the metal itself or melting it.

The result would be that the boat stops floating. Of course, this procedure would involve changing the microstructural base. Still, the point is that there is no scenario whereby the microstructural base is held fixed and does the causing of the event in question, while the supervenient property is not present. And, if you change the supervenience base (say, by substituting the metal with wood) while maintaining the supervenient property of being made of solid material, the boat would still float. This is evidence that the supervenient property is not causally superfluous. The base is not able to do any causing of the relevant effect, without the supervenient property. The fact that every time you subtract the supervenient properties from a cause, it fails to bring about the otherwise realized effects, together with the fact that you can change the base while keeping the supervenient property and generate the same effect, should lead one to think that supervenient properties have causal powers. This is how we generally detect causal powers: we see the difference between having certain factors present and not having them present.

The realization base was supposedly sufficient, but we see that if the supervenient vertically determined property is taken out, the subvenient physical cause is made insufficient. This provides evidence that supervenient properties make causal contributions, and Edward's dictum is false. This also indicates that there is no degenerate overdetermination. These factors are more aptly understood as contributing causes which form the sufficient physical cause of the boat's floating.

It is worth addressing at this point Gillet and Rives' (2005) belief that determinables (a variety of supervenient properties) with their own causal powers don't exist. Being red, being a determinable of being scarlet (or some other preferable physical base), does not exist. By the same reasoning, we might suppose that supervenient properties with their own causal powers don't exist. Gillet and Rives believe this based on what they call the "Parsimony Worry" and the "Causal Power Concern". The Par-

simony Worry is that:

[I]t is not clear that we should also take determinable properties such as being a mass and being charged to contribute powers in addition to the determinates that always accompany them. For to do so would be a kind of "double counting", to use David Lewis' phrase, of the causally efficacious properties (pp. 486-487).

The Causal Power Concern is that:

Once we distinguish determinates and their corresponding determinable properties the question is whether determinable properties really contribute any powers at all. By this we do not mean to again press the criticism that all the powers of individuals can be accounted for simply by positing determinates. Rather, the concern is whether being a mass or being charged, where the latter are not to be confused with some determinate mass or charge, actually contribute any causal powers to individuals. For example, ask yourself exactly what the property of being charged—again, not to be confused with some determinate charge—contributes to individuals by way of causal powers? The suspicion is that there are no such properties (p. 487).

Were Gillet's and Rives' thinking to apply to supervenient properties generally, their view would imply that causal supervenient properties don't exist. As Gillet and Rives recognize, however, the Subset View developed by Sydney Shoemaker (2003) provides a basis of an account that responds to these eliminativist worries, though they reject it. The Subset View says that determinable properties are distinct but partially overlapping with determinate properties. Causal powers of determinable properties coincide with a (non-empty) subset of causal powers of their determinate properties.

In an analogy with parts and wholes, it is worth pointing out that wholes don't eliminate their parts or the causal contributions of their parts on the basis of parsimony or because of the fact that parts do not outstrip wholes. The analogous claim is that causal powers of supervenient properties overlap with a subset of causal powers of their supervenience bases; but by no means does this entail their elimination.

To deal with their specific examples, I do think that having charge, any charge, has a causal power all positive and negative charges have: namely, the ability to interact through the electromagnetic force. All and only those things with charges have this ability, so it is not a trivial truth. Think about it this way. Sally has ten apples. Anybody who has ten apples has five apples. By implication, Sally has five apples. Sally should not say that because she has ten apples, that she does not have five apples. It is evident that she is not double-counting her apples if she asserts that she does have five apples when she has ten. She has ten apples, of which five are a subset. Now, those five apples surely have causal powers. For example, they might cause Sally to enjoy their taste, or be nurtured, or if she throws them down a building, they might break a car's windscreen. The fact that the five apples are a subset of ten apples does not eliminate this fact. The fact that having the ability to interact through the electromagnetic force is a subset of the causal powers of having some determinate charge does not eliminate that fact (see also Pereboom, 2002).

Now suppose that Sally did not have the five apples, then *ceteris paribus* the apples left to her are going to nurture her

less than the ten apples she might have had to eat when she had five apples. There will be less things she has which interact with the electromagnetic and gravitational forces, and so on, which makes prior sufficient causes for certain effects, insufficient. Having five apples is vertically determined by having ten apples. However, Gillet and Rives' eliminativism gets this wrong, I think, as does Edward's dictum. Eliminating Sally's having five apples, like eliminating her intentions, changes the world (see also Yablo, 1992).

### The Analogy with Fake Causation

Taking the lead from Jonathan Edwards, the second intuitive model Kim offers to support Edward's dictum is that of a mirror image which never causally depends for its existence on the existence of prior mirror images, but rather on being produced at each time by the thing it is an image of (Kim, 2005: pp. 36-37). The first thing to notice is that this model does not exhibit the truth of Edward's dictum, where vertical determination relations are at issue. Rather, the model consists of two things which might seem at first glance are such that one causally determines the other (the two images at different times), but this appearance is false. The model is an instance of "faux causation" (Kim, 2003: p. 171). But surely, this could not mean that just any putative causal relation is fake. So what is special about mental causation that makes it especially vulnerable or suspicious? Why does the existence of fake causation imply that vertically determined elements are always fake causes? Further, if the existence of fake causation implied that vertically determined causes must be fake, why would this not imply that all causes are fake? These questions are never addressed.

Just as there are plenty of examples of the fake causation that Kim points out, there are plenty of examples where there is real causation. By the same standard, these would show that vertically determined elements are real causes. Kim's argument here seems to be like saying that because Pluto is a fake planet (not a real planet in the current understanding anyways), that there are no real planets, which is evidently false.

General scepticism about causation is not at issue, nor is Kim supposed to be such a sceptic. That there is a distinctive threat for the mental is precisely what the CEA was supposed to show. But to point out that fake causation exists is insufficient for this. If it, by itself, were to work against the mental, it would work against any other kind of physical causation, which again results in general scepticism about causation.

### Conclusion

The CEA can only validly proceed by official incorporation of Edward's dictum. Without it, there is no reductive physicalist conclusion of philosophical interest. Once this fact is uncovered, it can be seen that the position validly reduced to absurdity is not non-reductive physicalism, but some other position no one should hold. If the premises are real representations of the non-reductive physicalist position, then there is no contradiction between them and the irreducible causal powers of the mental can retain their rightful ground in the physical world. Furthermore, if we give the premises the trivial understanding of "reductive physicalism" that Kim and Papineau endorse, we see that it is a trivial conclusion they are defending—one not worth calling by that name and standing up for, since it is unable to differentiate itself from a position asserting the exist-

ence of anti-physical things.

## REFERENCES

- Baker, L. (2002). On making things up: Constitution and its critics. *Philosophical Topics*, 30, 31-52.
- Barnett, D. (2000). Is water necessarily H<sub>2</sub>O? *Philosophical Studies*, 98, 99-112. doi:10.1023/A:1018636505551
- Chalmers, D. (1996). *The conscious mind: In search of a fundamental theory*. New York: Oxford University Press.
- Copeland, J. (2000). Indeterminate identity, contingent identity, property identity, Aristotelian-style. *Philosophical Topics*, 28, 11-25.
- Dretske, F. (1988). *Explaining behavior: Reasons in a world of causes*. Cambridge, MA: MIT Press.
- Funkhouser, E. (2007). A liberal conception of multiple realizability. *Philosophical Studies*, 132, 467-494. doi:10.1007/s11098-005-2530-z
- Gibbard, A. (1975). Contingent identity. *Journal of Philosophical Logic*, 4, 187-221. doi:10.1007/BF00693273
- Gillet, C., & Rives, B. (2005). The non-existence of determinables: Or, a world of absolute determinates as default hypothesis. *Nous*, 39, 483-504. doi:10.1111/j.0029-4624.2005.00510.x
- Jackson, F. (1994). Finding the mind in the natural world. In R. Casati, B. Smith, & G. White (Eds.), *Philosophy and the cognitive sciences*. Vienna: Holder Pichler-Tempsky.
- Jackson, F. (2006). On ensuring that *Physicalism* is not a dual attribute theory in *sheep's clothing*. *Philosophical Studies*, 131, 227-249. doi:10.1007/s11098-006-5989-3
- Jaworski, W. (2002). Multiple-realizability, explanation, and the disjunctive move. *Philosophical Studies*, 108, 289-308. doi:10.1023/A:1015799029176
- Kim, J. (1989). The myth of non-reductive materialism. *Proceedings and Addresses of the American Philosophical Association*, 63, 31-47. doi:10.2307/3130081
- Kim, J. (1998). *Mind in a physical world*. Cambridge: MIT Press.
- Kim, J. (2002). Responses. *Philosophy and Phenomenological Research*, 65, 670-679. doi:10.1111/j.1933-1592.2002.tb00231.x
- Kim, J. (2003). Blocking causal drainage and other maintenance chores to do with mental causation. *Philosophy and Phenomenological Research*, 67, 151-176. doi:10.1111/j.1933-1592.2003.tb00030.x
- Kim, J. (2005). *Physicalism, or something near enough*. Princeton: Princeton University Press.
- Kim, J. (2010). Thoughts on Sydney Shoemaker's *Physical Realization*. *Philosophical Studies*, 148, 101-112. doi:10.1007/s11098-010-9503-6
- Lewis, D. (1972, 1980). Psychophysical and theoretical identifications. *Australasian Journal of Philosophy*, 50, 249-258. doi:10.1080/00048407212341301
- Macdonald, C., & Macdonald, G. (2006). The metaphysics of mental causation. *The Journal of Philosophy*, 103, 539-576.
- Melnyk, A. (1995). Two cheers for reductionism: Or, the dim prospects of non-reductive materialism. *Philosophy of Science*, 62, 370-388. doi:10.1086/289873
- Melnyk, A. (2008). Can physicalism be non-reductive? *Philosophy Compass*, 3, 1281-1296. doi:10.1111/j.1747-9991.2008.00184.x
- Montero, B., & Papineau, D. (2005). A defense of the *via negativa* for physicalism. *Analysis*, 65, 233-237. doi:10.1093/analys/65.3.233
- Papineau, D. (2002). *Thinking about consciousness*. Oxford: Oxford University Press. doi:10.1093/0199243824.001.0001
- Pereboom, D. (2002). Robust nonreductive materialism. *Journal of Philosophy*, 99, 499-531. doi:10.2307/3655563
- Putnam, H. (1975a). The meaning of "meaning". In K. Gunderson, (Ed.), *Language, mind, and knowledge*. Minneapolis: University of Minnesota Press.
- Putnam, H. (1975b) *Mind, language and reality* (Vol. 2). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511625251
- Restrepo, R. (2012). Thinking about physicalism. *Open Journal of Philosophy*, 2.
- Shoemaker, S. (2003). Realization and mental causation. *Identity, cause,*



*and mind* (Expanded ed.). Oxford: Clarendon Press. 427-451.

Spurrett, D. & Papineau, D. (1999). A note on the completeness of  
“physics”. *Analysis*, 59, 25-29. [doi:10.1093/analys/59.1.25](https://doi.org/10.1093/analys/59.1.25)

Yablo, S. (1992). Mental causation. *Philosophical Review*, 101, 245-  
280. [doi:10.2307/2185535](https://doi.org/10.2307/2185535)