The Causal Closure of What? An Epistemological Critique of the Principle of Causal Closure

Carlo Gabbani

1. *The 'Physical World Causal Closure' Thesis*

The so-called 'causal closure of the physical world' thesis has played a major role (both in an implicit and explicit way) in recent philosophy of science and philosophy of mind, having also a substantial relevance for the relationship between the *manifest* and the *scientific image of man in the world* (W. Sellars).

A characterization of the thesis was given by Kim:

– "Pick any physical event […] and trace its causal ancestry or posterity as far as you would like; the principle of causal closure of the physical domain says that this will never take you outside the physical domain" (Kim 1996: 147).

Montero has proposed a 'softer' definition of the same thesis:

– "Every physical phenomenon that has a sufficient cause has a sufficient physical cause" (Montero 2003: 174).[1]

Another characterization was suggested by philosophers working on the 'completeness of physics:'

– "All physical events are determined (or have their chances determined) entirely by prior physical events according to physical laws" (Papineau 1990: 67; see also: Crane 2001: 45-46).[2]

[1] In principle, this (soft) formulation does not exclude (Montero 2003: 175; Lowe 2000) the possibility of an interaction between non-physical entities and physical effects (while this possibility seems to be excluded by Kim's 'strong' formulation), but it implies that it *always* has an over-determinant role in the physical world (in contrast with Robert C Bishop, I therefore read the thesis as a philosophical assertion *based on* physics, and not as one which states a mere "tipicality condition," see Bishop (2006). In this formulation, the principle is also compatible with the possibility that some physical events could be uncaused (Stoljar 2010: 211).

[2] Some authors draw a distinction between the thesis of the *completeness* of physics and (certain formulations of) the *causal-closure* thesis (like the first one we mentioned) (Montero 2005; Marcus 2005, § 1; Jones 2008).

I personally do not agree with the claim of causal closure and I maintain that the acceptance of such a thesis would have unfavourable epistemic consequences on our 'image' of man and the world. In this paper, I am mainly going to criticize the use of this thesis as a principle with empirical content which can be used as a premise of an argument, and not only as a mere research hypothesis (the term 'principle' is in literature).

2.1. *'Physical World' or 'the World of Physics'?*

I shall begin the analysis of the 'causal closure of the physical world' thesis with an examination of the words that constitute its name. Firstly, it is worth asking ourselves what the word 'physical' means here. And what does it mean to assert the 'physical' nature of every cause in our 'physical' world? This is relevant because the thesis does not only assert that there is a causal closure because all causes in our world simply belong to the same unique metaphysical realm, but specifies that this realm is the 'physical' one.

It has been remarked that the characterization of the 'physical' world given by many supporters of Causal Closure (that is, their answer to the so-called 'specification question')[3] often seems to be based on their delegating it to physics: the 'physical world' is simply 'the world of physics', no matter what this world is.[4] Any entity, process, property, or phenomenon which standard physics introduces (or will introduce) into its world-description will also belong to the physicalist philosopher's description of the 'physical' world. And he (or she) will also recognize a causal relevance to all the items to which physics recognizes this relevance, with the certainty of not violating causal closure in doing so. For, if the physical world is identified with the world of physics, whatever physics will introduce into its theories will, as a result, also be a part of the world as described by physics, that is to say (by definition) of the *physical* world[5].

[3] Crook, Gillet (2001: 334). As an exception, this essay avoids characterizing 'physicalism' on the basis of physics.

[4] "It looks as if the only way to identify facts as of one type rather than another – and therefore in particular as physical – is in terms of the vocabulary or concepts in which they are expressed. Purely physical facts are facts stated in a purely physical vocabulary, or with purely physical concepts" (Stroud 1987: 266; see also: Id. 2000: 50-52).

[5] As Crane notes: "A physicalist, traditionally, is someone who gives a certain kind of authoritative role to physics. This role is partly epistemological – physics has an authority in

The core of this strategy can be seen as an extension of what has been called the "Theory View" or the "Theory-Based Conception" of the 'physical' according to which (first part of the thesis): "F is a physical property if and only if F is expressed by a predicate of a physical theory."[6]

In principle, it appears that there aren't any insurmountable problems in delegating the definition of a dominion on which a philosophical thesis is asserted (or one that goes far beyond the proper limits of any scientific field) to an experimental discipline. In this case, however, it is not at all obvious whether this ontological authoritative role given to physics is warranted. I will analyse this point later. For the moment, I would like to underline that, by delegating physics to establish what counts as 'physical', a truly *epistemological* component is embedded within what is intended to be a form of *metaphysical* monism about (non-redundant) causality. Therefore, the claim that "in the *physical* world all influent causes belong to an ontological realm that is *physical* itself" becomes something like: "All causes in the physical world (or, at least, all influent, not over-determinant causes) also belong to *physics*." This implies that if the causal closure of the physical world thesis were true, then a certain kind of scientific knowledge could in principle account for every (non-redundant) causal process in our world. Note that, on the contrary, in the presence of a truly philosophical (and not only scientific) characterization of the 'physical', an ontological monism about causation would not entail an epistemological monism about causation, being consistent with the hypothesis that some causal nexus is 'physical' but, nevertheless, irreducible to 'physical' knowledge (see also: Strawson 2003: 49, 51-52).

To sum up, the thesis of 'causal closure' firstly identifies the causal closure of the world with a 'physical closure', and then identifies the 'physical' with the domain of 'physics'.

It could immediately be objected (and with good reasons) that, if assumed to apply to all causal processes in our natural world, a principle formulated in this way would seem false, it being impossible to reduce the descriptions and explanations of every causal link in our world to physics because of an *intra-scientific* irreducibility (see, for instance: Dupré 1993: 99-102; Id. 2001: ch. 1, 7). This amounts to say that genuine and non-redundant causal *nexa* introduced, say, into biology,

telling us what to believe – and partly ontological – physics has an authority in telling us what there is" (Crane 2001: 44; Id. 1994: 480).

[6] Stoljar (2010: 71; see also: 75, 236). He rightly describes this conception as the "standard view about these matters in contemporary philosophy" (2010: 72).

medicine, or geology, are simply irreducible to present or future fundamental physical theories. As a result, the principle does not apply to them (because we simply do not have 'physical' events), or the principle, when applied to them, is false.

However, Tim Crane differentiates the claim according to which "physics tells the whole story about the causation of *physical events*" from what David Lewis labelled "the explanatory adequacy of physics:" i.e. the idea that the conceptual framework of physics *alone* is enough in order to describe and explain every kind of existing phenomena. The causal closure, on the contrary, would be consistent with intra-scientific pluralism (that is with the notion that the different conceptual frameworks pertaining to the various experimental sciences are *irreducible* to that of physics), provided that all the entities they refer to have physical properties and that "whatever physical events occur have a physical causal history which fixes their occurrences" (Crane 2001: 45; Papineau 2001: 3).

I suppose that one could also understand the thesis we are examining as implying that what has a (not over-determinant) causal relevance in our world is entirely made by the primitives of physics, and that: (i) insofar as a world with physical properties is concerned, requirements and restrictions of the *physical* causation (that is, of causation as it appears in physics) properly apply to *every* (not over-determinant) causal process *in conformity to its scale*, even if this latter is *not* completely described and explained by physics *alone*; (ii) every (not over-determinant) causal process in that world *conforms* to requirements and restrictions of physical causation and *confirms* them, even if physics may not completely account for the causal process. Therefore, it is not only asserted that every (influent) causal process in that world would not violate a physical causal closure, but that it might *corroborate* it. In other words, according to this thesis, no (influent) causal process violates the standards of physical causation, in the sense that these standards do apply to every causal process in conformity to its scale, and that every influent causal process instantiates these requirements in conformity to its scale.

In any case, various supporters of scientific pluralism would also be likely to endorse a naturalistic (re)formulation of the causal closure thesis such as: for each x, if x holds an original and non-redundant causal role involving our physical world, then the influent causal properties of x are (at least in principle) fully described and explained by the different experimental sciences of nature. That is to say that every cause is a 'natural' cause, 'natural' meaning, here, "fully incorporated in a present or future naturalistic scientific-experimental theory." My critique is also ad-

dressed to this variant of the thesis. Thus, while I will examine the standard physicalistic formulation, my arguments should result in being equally valid against the 'naturalistic' version.

After these preliminary remarks, let me examine two epistemological problems concerning the formulation and the sense of the thesis. They have already emerged in literature: the first is in some way surmountable; the latter seems to cast a shadow of circularity on the use of the causal closure thesis as an empirical principle.

2.2. *The Argument from the "Open-ended Character" of Physics Against the Principle of Causal Closure*

Awareness of the fact that physicalists often delegate the characterization of the physical world (or of that which has a causal relevance within it) to physics has led to objections to physicalism in general. Insofar as these objections are effective, they are obviously also relevant to the causal closure thesis.[7]

The main objection is based on a dilemma firstly formulated by Carl Gustav Hempel.[8] The point is essentially the following: if physicalism claims that only that which belongs to our *present* physical theories exists, it seems to be false (because those theories appear to be at least largely incomplete); but, if physicalism claims that only that which belongs to *ideal*, accomplished, future physics exists, then it becomes an "unclear" thesis, which is open and undetermined (because we do not know what will belong to that physics). In both cases, 'physicalism' is in serious trouble.

Tim Crane and David Hugh Mellor, among others, have elaborated a

[7] As Stoljar observes (2010: 222): "Causal closure is subject to the same sort of objection as physicalism itself."

[8] "The physicalistic claim that the language of physics can serve as a unitary language of science is inherently obscure: the language of what physics is meant? Surely not that of, say, 18th century physics; […] The thesis of physicalism would seem to require a language in which a true theory of all physical phenomena can be formulated. But it is quite unclear what is to be understood here by a physical phenomenon" (Hempel 1980: 194-195). A similar point is made by Chomsky (Antony, Hornstein 2003); see also: Stroud (2000: 53-54). For a critical analysis of "Hempel's Dilemma" (and an alternative argument against physicalism), see Stoljar (2001 revised 2009: § 11.2); (2010: chap. 5); Melnyk (1997); Malatesti (2005); Vicente (2011). On the formulation of 'physicalism', see also the monographic issue of "Philosophical Studies", 131 (2006), edited by Janice L. Dowell. For a general analysis of the debate on the nature of the 'physical', see Montero (1999) and (2009).

similar argument in their influential essay "There is No Question of Physicalism" (Crane, Mellor 1990). They observe that physicalists generally assert that only what that belongs to physics, or is reducible "in principle" to physics, exists. But, plausibly, physicalists will not maintain that only what belongs, or is reducible to *present* physics (that is to the physics we already know) exists. The 'physics' that physicalists have in mind is not *our* physics, but an ideally accomplished physics of the *future*.

It is that physics that physicalists are inclined to recognize as the "measure of all things" (to use Sellars's motto). Crane and Mellor underline that there is an apparent threat of circularity here. How can this "principle of reducibility" for all existing things be understood and applied now, if, at least for the time being, we do not know what in principle may or may not belong to future, accomplished physics? Physics seems to have such an "open-ended character" that "if we apply the principle to an otherwise unspecified future physics, we shall not be able to say which sciences are physical until we know which of them that physics must cover – which is exactly what the principle was supposed to tell us […] So the physical cannot be defined as what is reducible in principle to physics, either present or future" (Crane, Mellor 1990: 188, see also: 191).

In other words, what can we know *now* about the characteristics of *future* physics, the entities, properties and processes that will belong to its future/final ontology, and the paradigmatic revolutions that it could undergo? Then – as Crane and Mellor urge – how might we ever exclude that the future physical world-description will be so different, up to the point of being based on very different elementary constituents, and up to the point of including new kinds of properties and entities as their proper objects? Therefore, once it has been asserted that everything is 'physical' (or that everything that has a genuine causal relevance in our world is 'physical'), i.e. belongs to the world-description of physics, it is unclear what exactly we have asserted. It may be something that is obviously *false*, if physics is our *present* physics, or something that is too *undetermined*, if physics is a *future*, ideally complete physics. And Crane adds that if we stipulate that this future physics has to be considered complete on the basis of its adequateness in explaining everything, then the physicalist claim will obviously be "true by definition" in its indeterminacy (or better, because of its indeterminacy).[9]

[9] Crane 1991: 35; Crane observed that "we cannot now rule out the possibility that some future complete physics will include mental properties and laws" (Crane 1991: 34); see also:

In recent years, Bas van Fraassen has also advanced similar objections independently, underlying that the claim according to which only that which is physical/material exists, would require "a genuine, and not just a verbal, distinction" between what is 'material' and what is not, in order to be a factual meaningful thesis (van Fraassen 1996: §§ 3-4; Id. 2002: 49-61; see also: Stoljar 2010: 176-180). Van Fraassen remarks that we should not expect this demarcation criterion to be found *within* physics itself, because "whenever philosophers take some general feature of physics and use it to identify what is material, what happens? Physics soon goes on to describe things that lack that feature and are altogether different" (van Fraassen 2002: 53) In this context van Fraassen has also advanced some significant meta-philosophical considerations. In his opinion, precisely for the reasons mentioned above, a physicalistic and materialistic perspective should be presented not as a "factual claim," but as a "scientific hypothesis" or as a "stance" in order to be philosophically honest (that does *not* mean acceptable for him). In other words, it should be presented as an "attitude or cluster of attitudes" (van Fraassen 2002: 53; see also: Melnyk 1997; Poland 2003; Ney 2008) which expresses a personal (even if motivated) option for a certain image of the world, and are far beyond the range of our factual knowledge. Specifically, the physicalistic component of a stance would be "the desire or commitment to have metaphysics guided by physics."

I will return to this theme later. Beforehand, I would like to consider the consequences that these kinds of objections based on the 'indeterminacy' of physics have on the option for causal closure.

2.3. *The 'Via Negativa'*

The objections examined up to now are of a very relevant kind, but they are possibly not insurmountable for the causal closure thesis, even in order to give it factual significance at the present time. This possibility supposes that, in spite of a still evolving 'physics', an effective characterization of what it entails to be a 'physical' property can be already given. A line of thought of this kind has been especially pursued in the so-called "via negativa" to physicalism. A version of this strategy was pro-

Crane (1994: 480); Montero (1999: 191-193); Ead. (2009: 182); van Fraassen (2002: 55-57); something analogous is asserted by Chomsky (see: Lycan 2003; Poland 2003; Strawson 2003, along with Chomsky's replies in Antony, Hornstein 2003).

posed by David Papineau (and David Spurrett) in his reply to Crane's (and Mellor's) remarks, and then in his (co-authored) research on the completeness of physics (see: Spurrett, Papineau 1999; Papineau 2002: 40-46; Id. 2009: § 2.3; also: Wilson 2006; Montero 2009: 184-186).[10]

What physicalists need is a characterization (it could even be a non-definitory one) of everything that can be considered as a basic 'physical' element. This task does not necessarily require that we find a *positive* characteristic that is both necessary and sufficient. It would be enough to have just a negative and (only) necessary characteristic of the primitives of the 'physical' world, that is to say, a characterization asserting that if something is a primitive constituent of physics, then it does *not* instantiate the property x and will never instantiate it.[11]

Such a criterion will allow physicalists to have some certainty about the basic ontology of physics and of the 'physical world', knowing at least what this certainly will *not* include. In fact, this seems to meet the minimal condition for the significance and the applicability of the causal closure thesis. What is essential, from this point of view, is not to establish what all the objects of the basic ontology of an accomplished physics exactly are, but rather to determine which, amongst the objects of our manifest image of the world, *cannot* in principle belong to the basic ontology of such physics: "it isn't crucial that you know exactly what a complete physics would include. Much more important is to know what it won't include" (Papineau 2002: 41). One could assume, for instance, that the basic ontology of physics is (and will be forever) characterized (negatively) by the absence of "intentional in-existence," which is a property that, on the contrary (as a matter of agreement), may be considered as a mark of the presence of mental phenomena or acts. In other words, the absence of intentional in-existence is considered here as a necessary (even though insufficient) condition for something to be 'physical'. This also means that not all candidates would result in being equal-

[10] For a critique to the 'via negativa,' see Gillett, Witmer (2001; and the reply: Montero, Papineau 2005); Crook, Gillett (2001: 349-350); Malatesti (2005: § 5); Ney (2008); Stoljar (2001 revised 2009: § 11.4); Id. (2010; ch. 4, 5), Vicente (2011: § 2), such a 'via' seems to be appreciated by Loewer (2001: 40). See also Strawson (2003). The debate between Papineau and Crane began with Papineau (1990) (where the remote possibility that 'mentalistic' categories will figure in a future, accomplished physics is not yet ruled out, see: p. 70) and then continued with Crane (1991), Papineau (1991). A different attempt to escape from the difficulties highlighted by Crane and Mellor is Pettit (1993), which is based on a 'positive' characterization of the 'physics' (the text is followed by Crane's reply).

[11] See also the 'contrast' element included by Stoljar (2010: ch. 4), within the "Theory View" of physicalism.

ly acceptable as basic constituents of a future physics (at least, as long as the methodological principles of this discipline remain the same). Therefore, one should not too much emphasize the "open-ended character" of physics (or of the physical) from this perspective.[12]

In any case, it seems that, together with such a kind of characterization, the thesis that what exists/has a genuine causal relevance belongs to/is reducible to physics may have, *even at the present*, significance and applicability. This is because it implies that everything that exists (or has causal relevance) can in principle be reduced to/explained by a basic ontology within which there is nothing with a certain kind of property, that is, nothing of a mental nature. From this point of view, as noted by Spurrett and Papineau, "the relevant completeness thesis is then the claim that the non-mental is complete," i. e.: "every non-mentally identifiable effect is fully determined by non-mentally identifiable antecedents."[13] In actual fact, according to Papineau, the correct negative characterization of the 'physical' would result in being "'identifiable non-mentally-and-nonbiologically', or inanimate" (Papineau 2002: 41).

What physicalists want to assert with such an understanding of the 'physical' is *not* that the *explanandum* of physics is restricted to what we now generally describe as "'identifiable non-mentally-and-nonbiologically', or inanimate." Otherwise, and paradoxically, their characterization would result in being very near to a classic antireductionist theory *à la* Brentano: we have mental, biological, intentional properties, but physics includes only non-intentional properties (or identifiable non-mentally-and-nonbiologically properties), so there are properties that cannot in principle belong to physics. What physicalists want to assert is that all those properties are in principle reducible to the basic ontology of

[12] Note, however, that this kind of negative characterization of the physical is still ultimately based on physics as a discipline, and on a methodological (legitimate) assumption concerning the type of primitives it can or it cannot call on. On the contrary, it is impossible to say that what is excluded from the basic ontology of physics/physical is excluded because physics' success in explaining *everything* without any need for reference to it is *already evident*. In this case the definition of 'physics'/'physical' (as the 'non-mental') would already *presuppose* the causal completeness of 'physics'. Also the choice to characterise physicalism simply as the thesis according to which the 'mental' is not basic and always depends on the 'non mental' is something that is conceivable, relevant and can be defended only on the assumption that physics does not assume, and will never assume, the mental (as such) amongst its primitives.

[13] See, respectively, Spurrett and Papineau (1999: 25) and Papineau (2002: 41). Wilson (2006) has labelled a similar negative criterion "no fundamental mentality constraint." Spurrett and Papineau also present another possible (and complementary) characterization: "the quantitative is complete" (Spurrett, Papineau 1999: 25).

physics, i.e. to the non-intentional, the inanimate, and the non-mentally-and-nonbiologically identifiable, etc. That is to say, according to this physicalistic perspective: "the mental must be identical with the non mental," or: "the mental is ultimately decomposable, or reducible to, or analyzable in terms of fundamentally nonmental phenomena" (see, respectively: Spurrett, Papineau 1999: 25; Montero 2003: 180). This point is worth underlining because, *prima facie* and without any adequate reduction theory, the 'via negativa' sounds like the exclusion (and not the inclusion) of something existent from the physical domain.

This proposal seems a relevant physicalistic attempt to escape from the problems arising from Hempel's dilemma.[14] From this perspective, I suppose that a reformulation of the causal closure thesis in the light of the 'via negativa' could be the following: "Every event involving the physical world (as negatively characterized above) is determined, or has its chance determined, by purely physical causes, that is, by causes belonging to the *world of physics*, as negatively characterized above. This means by causes which do not have a certain kind of property (such as being 'mental') which cannot be included, *as such*, within physics."

I do not want to discuss the validity and limits of the 'via negativa' strategy, here. Let us assume for the moment that, *prima facie*, we have a reformulation of the causal closure thesis as a coherent, empirical hypothesis, which is significant and also relevant to recent debates in philosophy.

3.1. *Empirical Hypothesis and/or Guiding Principle*

However, is the causal closure thesis really endorsed by its supporters as an empirical hypothesis? If so, it would simply represent an empirical (but not only experimental) hypothesis to be tested (van Fraassen 2002: 53-55) More specifically, it should be tested if the hypothesis according to which every event involving the physical world is caused by a physical cause (i.e., a cause belonging to physics) is, in fact, justified and validat-

[14] It is remarkable that, while *in theory* countless negative characterizations might be conceived in order to determine what is 'physical', *in fact*, if one is looking for a characterization that has an effective and practical relevance for us now, it must be grounded on a property that is now generally considered as instantiated in our world, at least in our common-sense image. Otherwise, one will exclude from the domain of physics only properties, and therefore entities, on which no existing causal explanation actually rests.

ed for all kinds of phenomena and events. It seems very hard to assert that physicalists do actually adopt the causal closure thesis as such an empirical hypothesis (even when they assure us that they do so). I particularly see two main objections to this merely 'hypothetical' endorsement of the thesis.

Firstly, if one accepts the causal closure thesis as a mere empirical hypothesis, it would result (at least at the moment) as a hypothesis which, even if not false, is certainly not yet justified, corroborated, or implemented.

For instance, we all use explanations that are currently indispensable, but that seem to be based on the attribution of a genuine and efficacious causal role to the conscious states individuated by common-sense psychology,[15] which is to say, intentional as well as qualitative states that we are, *now*, not in the condition of individuating, nor of reducing, nor of replacing and eliminating in any adequate way within the naturalistic or physicalistic conceptual framework.[16]

This does not imply that common-sense psychology merely functions as an explanatory framework or has a structure wholly "homogeneous" with that of experimental sciences of nature, but simply that our mentalistic discourse is sometimes able to provide us with genuine explanations, within which the conscious states it refers to play a (proximal) causal role towards our actions.

It is well known that this general perspective has been attacked in many ways and with different degrees of radicalism in philosophy. For instance, it has been claimed that the ordinary psychological discourse based on reasons does not have a causal character (it is not a kind of causal explanation of behaviour), and *does not aim* to confer a causal role to internal conscious states. Others, starting from a completely different perspective, have suggested that present folk psychology *would aim* to in-

[15] Speaking of a genuine and irreducible 'mental causality', I do not intend to endorse here a specific theory concerning 'causality.' 'Mental causality' is an appropriate expression for saying that we are justified in asserting that the occurrence of a certain conscious event plays a genuine and non-redundant causal role for observable events/actions in our world. For instance, we could apply Mackie's condition to a conscious state: "insufficient but non-redundant part of a condition which is itself unnecessary but sufficient for the occurrence of the effect."

[16] From the point of view of common sense, it is generally accepted that, when occurring in the mind and receiving the will's consent, certain conscious states may be among the causal antecedents of an action (also in the sense of efficient causation), and that, accordingly, what we usually individuate as 'reasons' may also play a role as 'causes'. One could object that reasons, considered as types, are not in space-time and cannot be causes; the point is that the occurrence of a reason in my conscious mind may nonetheless be part of an efficacious causal complex.

dividuate and categorize our conscious states and their causal role, but in fact it systematically fails, because it is so massively false that it cannot even have referential efficacy; or that it fails because the conscious states it refers to are not genuine causes of human behaviour, which cannot be known on a "personal level of analysis," like that of common-sense psychology. I do not want to discuss these objections here. I simply assume that my argument is valid for those who don't share these objections but, on the contrary, recognize that common-sense psychology legitimately individuates some conscious states as a causally influent (proximal) antecedent of some intentional actions,[17] and that our common-sense psychological discourse may therefore also contribute to causal explanations of our intentional actions.[18] Amongst those who share this latter point of view, there are today also many supporters of the causal closure thesis, who do not generally deny that common-sense psychology *wants* and *may in fact* individuate states of mind, which are causally influent in our world. On the contrary, this interpretation also seems to be of great relevance to the philosophical projects of some of them.[19] So, they generally assert that our common-sense psychology individuates occurrences of states and properties, which are causally influent, in a non naturalistic way. However, they also assert that the latter are identical/reducible to 'objects' (states, processes, entities etc.) which properly belong to the scientific image. Thus, the causal closure is perfectly respected because we are simply confronted here with a commonsensical, unscientific way to designate completely natural causes, which, at present or in the future, may be better individuated and described by natural science.

[17] As Alfred Mele observed: "some conscious intentions are among the causes of corresponding actions" (2009: 1) and: "There is a powerful evidence for the truth of the following thesis: the fact that an agent consciously decided to A or had a conscious intention to A sometimes has a place in a causal explanation of a corresponding overt intentional action" (*ibid*.: 144). For further details, see Dupré (1993: ch. 7).

[18] As Tylor Burge asserted: "No one has shown that mentalistic explanation is either non-causal or non-descriptive. Nor is either view plausible" (2007: 347).

[19] Kim's philosophy of mind, for instance, is explicitly based on the idea that reducibility is possible because and when the properties of our mental states (the referents of our mentalistic discourse) are "functionalizable," i.e. they can be "defined or characterized in terms of their causal work" (2005: 165; see also: 24). Here, then, reduction presupposes the referentiality of common-sense psychology and the causal relevance of its referents. From a different perspective, Papineau asserts a thesis of token identity with regard to physical and mental states (1990), and his "conceptual dualism" recognizes our everyday concepts as efficacious "ways of referring to conscious experience," and asserts that "an everyday term like 'pain' expresses both […] a concept of a state that feels a certain way, so to speak, and a psychological concept of pain, a concept that refers by association with a certain causal role" (2002: 96-98).

But will it happen now or in the future? This is an important question. It appears that a number of neutral observers and physicalists (see, for instance: Loewer 2001: § 5; Heil 2004: 215; Kim 2005: 165) recognize that, *at least at the moment,* we do not have such adequate physicalistic, reductive theories which fully account, for instance, for our intentional conscious states. It follows then that some categories and accounts of common-sense psychology are (at least at the moment) irreducible to science and are today the only way to correctly individuate conscious states that appear to be genuine causal factors for certain human actions.[20] Accordingly, if the causal closure were an empirical, factual hypothesis, it could simply be objected that at the moment we may individuate instances of a kind of causation which is irreducible to the physical (or even naturalistic) description of the world. Thus, since the 'physical' has been characterized as what belongs to the 'physical description of the world' (and 'nature' as what belongs to the naturalistic description of it), it seems that we may now exhibit a kind of 'non-physical' (or 'non-naturalistic') causation. So physicalists would still have to demonstrate, step by step, that all causal roles which our explanations confer to our conscious states, do in fact belong to something 'physical' (or do not really exist). And, as we have already remarked, physicalists themselves recognize that this job remains to be done.

So the causal closure thesis cannot be based on the existence of a successful, and fully completed, general reductionist theory: on the contrary, it seems to be one of the *principles* upon which the physicalist's expectations of this general reductionist theory are built.[21]

[20] It is worth pointing out that according to some theories of ontological commitment, this would in itself result in being a good reason for accepting the existence and the causal relevance of such conscious states in one's own basic ontology (Dupré, 1993: 94). This is because the more economic formulation of the theories one in fact adopts in order to describe and explain our experience introduces quantified variables of that kind (that is of a 'mentalistic' kind), conferring them a certain indispensable causal role (I'm not assuming this theory, at least not without restrictions). Crane and Mellor observed that "unless the sciences of the relatively large, including psychology, reduce to microphysics, we shall still need to quantify over entities described in those sciences term" (1990: 190).

[21] Note that, on the one hand, the physical world and the natural world are characterised in epistemological terms (that is on the basis of physics, or on the basis of natural sciences). On the other hand, it is asserted that also the referents of mentalistic language that we are not in condition to characterise and individuate within the physical or within the naturalistic frameworks alone are nevertheless (identical with) physical or naturalistic objects. But it seems that the 'physical world' and the 'natural world' cannot have both this *intension* and this *extension*; and one should not ascribe one item to a realm characterised on an epistemological basis, regardless of the present epistemological status of that item.

The second relevant point is based precisely on the fact that the causal closure thesis is not often used as an empirical *hypothesis*, but as an empirical *principle*. Obviously, this is not a nominalistic matter. If we had nothing but an empirical hypothesis, then the causal closure thesis could neither be used as a universal, true premise in an argument (as, in fact, sometimes happens), nor would it be possible to escape from the necessity of exhibiting an adequate reductive theory *before* any endorsement of the truthfulness of the Causal Closure. I certainly do not intend to assert that the causal closure thesis is proposed by physicalists as a self-evident, a priori, or logical truth. Nonetheless, once the meaning of words has been established, we often proceed as though we should not test its truth-value in each relevant case, but simply use it as a guiding principle for the explanation of our world, and explain how it must necessarily be valid in a world like *our* world.

More specifically, the causal closure thesis seems to be used in order to impose a preliminary constraint on what can be recognized as an adequate understanding of the causal processes concerning the phenomena we experience. Therefore, it will also be possible to use the causal closure as a principle on whose basis each causal explanation will be evaluated as legitimate or illegitimate, fundamental or reducible. This amounts to say that, in general, if an explanation we adopt appears to break the causal closure, this fact will not prove the falsity of the causal closure but the falsity or the reducibility of this explanation. This happens especially when the problem of mental causation is concerned. From this point of view, the amount of now inescapable and irreducible 'mentalistic' causal explanations is generally discarded as an argument for non-'physical' causality (i.e. for a causality that does not belong to physics). On the contrary, it seems that those who assert the genuine, irreducible causality of the mental (of what is identifiable only mentally) have the burden of proving this to the 'tribunal of physics'.

Thus, the use of the causal closure as a principle seems to lead to the following line of argument:

– If a 'mentalistic' explanation does in fact individuate a genuine causal nexus, then it must be reducible/identical to a 'physical' causal nexus (i.e., a causal nexus individuated by physics).
– If, on the contrary, the 'mentalistic', causal explanation really is (even in principle) irreducible to the physical description of the world, then this must be because it refers to ill-conceived and inexistent phenomena (eliminativism), or, in any case, to something which does not hold a genuine causal efficacy in our world (epiphenomenalism).

As a result, there is either no true irreducibility, or no true causation.[22]

It should now be clear that, in fact, the real issue here is not even the ultimate truth concerning the nature and reliability of common-sense psychology. Rather, it is a preliminary, methodological point: the physicalistic thesis according to which common-sense psychology cannot individuate genuine and irreducible kinds of causes often arises not from a neutral, unbiased analysis of its real nature and adequateness, but from an already physicalistic understanding of 'causation' and, therefore, from the adoption of the causal closure as a principle. Thus, the principle of causal closure becomes, so to speak, something similar to "the substance of things hoped for, the evidence of things not seen."

3.2. *Causality or Physical Causality?*

A crucial point emerging here is that the causal closure thesis eventually proves to be not an option concerning causality that is *actually* obtained in our world. As a matter of fact, it concerns the features and structure that something *must* possess in order to really have a genuine, influent, and irreducible causal role involving *our* world. In other words, causal closure as a principle seems to imply that the establishing of the characteristics and structure of genuine causal *nexa* in our world is delegated to physics. This choice is analogous and complementary to the one previously described in order to characterize what is 'physical'. While physics was then crucial to determine what counts as 'physical' (the distinctive features of the 'physical world' are in fact the distinctive features of 'the world of physics'), now seems to be up to it to establish what can be admitted as a genuine causal *nexus* in our world. This means that, in fact, physics is not simply asked to determine what a *physical* causation process (intended as a species of a conceptually broader *genus*) is, but rather to establish under what conditions something can be considered as a 'cause' *tout-court* in this world. While in the first case physics was

[22] A similar diagnosis can be found in Marcus (2005: 27): "We are left with the following dilemma: Either mental causation just is (ultimately) physical causation, or it's nothing at all" (cf. Dupré 2001: § 7.3). We have here something analogous to the materialistic attitude in conceiving mental phenomena described by John Searle: "Earlier materialists argued that there aren't any such things as separate mental phenomena, because mental phenomena are identical with brain states. More recent materialists argue that there aren't such things as separate mental phenomena because they are not identical with brain states. I find this pattern very revealing, and what it reveals is an urge to get rid of mental phenomena at any cost" (1992: 48-49).

asked to define a region of reality (the 'physical'), in this case physics is asked to define not only what 'physical' causation is, but what a true case of irreducible causation in our world is (I will call this 'one dimensional understanding of causality').

Yet, this seems to generate a kind of inconsistency. If it is implicitly assumed that the only effective kind of causality allowed in our world is causality as conceived within physics, then undoubtedly in our world only 'physical' causes will exist (since, in this context, to belong to physics is understood as being 'physical'). Consequently, this world will certainly result in being causally closed, but as a matter of definition and not as a matter of fact. This proves that non-physical causality is excluded from a conceptual framework, but not from our world. Delegating the understanding of causal processes in our world to physics seems therefore to be a real *punctum dolens* for causal closure intended as a *principle*.

To be honest, this inconsistency does *not* reside in explicitly asserting that which should be proved, i.e. that, in our world, irreducible and influent causality can only be physical (formally, the theoretical possibility of non-physical causation may still be admitted). This especially emerges from the objections levelled, in an implicit and, therefore, more dangerous way, by physicalists against the supporters of the existence of a true and influent non-physical causation in our world.[23] The 'conditions' that physicalists impose for accepting that something may constitute a genuine, irreducible occurrence of causation in our world often seem to be such that to fulfil them and to be 'physical' appear to be one and the same thing. Even the frequent request to show how the existence of non-physical causes can be 'compatible' with the laws of physics often takes the form of a request to show that such non-physical causes are consistent with the causal closure which is immanent to the laws of physics. However, if non-physical causation amounted to being subjected to the principles that (at a certain point of time) physics assumes as essential for having a causal process, and precisely as physical causation is subjected to them, then it would simply

[23] When analysing the causal argument for physicalism, Bishop, (2006: 51) suggested the presence of the "hidden premise – that the only efficacious states and causes are physical" and that it therefore begs the question. Looking at it from a different perspective, I would suggest that a very similar, but even more radical problem affects the causal closure principle itself. I suppose that this partial diversity is due to our different interpretation of the causal closure thesis: if the causal closure only states a "typicality condition," the hidden premise seems to be necessary *outside* of it in order to base the causal argument on it; if, on the contrary, it is a stronger *principle* (expressing the irrelevance of non-physical causes in the physical world), a hidden assumption similar to that may reveal itself as already present *within* it.

be a kind of 'physical' causation. This is because it is on the basis of those standards that the 'physical' dominion is characterized. In other words, my point is that, in order to be considered *real* in our world, a supposed causal link should not be required to posses exactly what is needed in order to belong to the physical realm (i.e., the domain of physics).

Physicalists might perhaps reply that this alleged impossibility of thinking of something as a genuine irreducible cause in our world, without already characterizing it as a physical cause, is a proper argument for physicalism (but if this is the case, it appears to be an argument for an a priori truth). However, my point is that, on the contrary, it seems that to conceive 'causation' in our world on the basis of physics (if not of mechanics alone) is in no way a necessary option; it is, indeed, even less justified than conceiving the 'physical world' as the 'world of physics'.

On the one hand, as Roderick Chisholm (in the wake of Thomas Reid) has remarked (Chrisholm 1964: § 9), the general concept of 'cause' seems to emerge from the conceptual framework of common-sense beliefs concerning our experience as *agents*, and not from the world of events described by the natural sciences. On the other hand, it is worth remarking that, on a conceptual level of analysis, genuine mental causation seems in principle (and by definition) only incompatible with general theories about the nature of a 'cause' which *already* presuppose the physical nature of every causal process (see also: Crane, Mellor 1990; Crane 2001: § 11). That is to say, it only appears to be at odds with theories which are valid for causal mechanisms pertaining to *certain* fields of physics (and, perhaps, to a few other scientific fields), but not even to science *in general* (like Mark Transmission Theory, or Conserved Quantity Theory, *deterministic* law-likeness).[24] On the contrary, it doesn't seem that there are *in principle* any barriers to mental causality in fulfilling the other general criteria advanced by 'non-chauvinistic' theories of causality (regularity of the nexus, probability of the link, counterfactual dependence, manipulability of effects, productive contribution to the effects, explicative function, evidential role, intrinsic nature of the relation, etc.) (see also: Schaffer 2004: 60), and, even less, in being consistent with various forms of a "contextual," or a "pluralistic" approach to causality.[25]

[24] Obviously, one can also try to construct a causal explanation based on those kinds of theories for consciousness, but I see no reason to suppose that causation in general has to be *defined* in those terms only.

[25] See: Hitchcock (2003); Id. (2007); Cartwright (2004); Campaner, Galavotti (2007: espec. § 5). On mental causation and counterfactual dependence, see Loewer (2007: 255-259).

A new sort of "open-ended character" of physics may emerge here in some way. It is unclear what properties or requirements (apart from those which simply characterize all that is 'physical') actually define something as a *causal* relation *within* physics. Which properties or requirements would define (specifically and in principle) each causal process within physics, while nothing mental can posses or fulfil them? And is it because of them that something really is a *causal* nexus in our world?[26] Are the conservation of energy, the relation between energy and work, measurability in quantitative terms, conditions concerning time or locality, etc., really good candidates for this general criterial-role?[27]

But, above all, as Tim Crane observed (1994: 481), it is precisely this choice of applying the same conception of causality to what is 'physical' and to what is 'mental' ("Homogeneity"), and then the idea that the two causal processes have the *same* job to do, that leads to the idea of a clash between the two forms. It is this type of one-dimensional understanding of causality, together with the belief that a linear additivity of the quantity of work of each causal component would be possible, that gives plausibility to the idea of a component being 'sufficient' for the observed outcome, as well as to the claim of a 'redundancy' of mental causation, as in Kim's "exclusion argument."[28] However, in dealing with 'mental' causality, all the apparatus of measurement and comparison, which are essential when we deal with the objects of physics, probably represent a 'category mistake'. Tyler Burge has convincingly described the problems of understanding

[26] As John Norton has pointed out: "centuries of failed attempts to formulate a principle of causality, robustly true under the introduction of new scientific theories, have left the notion of causation so plastic that virtually any new science can be made to conform to it. Such a plastic notion fails to restrict possibility and is physically empty" (2007: 12, see also: 21-22): "our notion of causality evolves in response to developments in the science" (*ibid.*: 32).

[27] The fact hotly debated today that some of these requirements are *in fact* violated in certain theories of our *present* physics shows that it is at least conceivable and possible (if not actual) that physics may renounce the universal validity of these requirements; that is to say, it is at least conceivable that theories properly belonging to physics do not assume those characteristics as defining the *explanandum* of physics, or the 'physical' causality. For the physicalistic analysis of this point, see Papineau (2002: 255), Montero (2006), Vicente (2011: 411-412).

[28] Barry Loewer (2007: 252-255), drawing upon Ned Hall's distinction between causation as "production," and causation as "dependence," has shown that the validity of Kim's exclusion argument seems to rest on an understanding of causation as 'production', or, even better, as a factor which is in itself *sufficient* for the production of the effect (see *ibid.*, p. 253), as in the folk-model for the mechanics of billiard balls. But this model would hardly be enough in physics, and in no way could be generalised: "The intuitive force of Kim's argument derives from the fact that we tend to think, mistakenly, that causation is a fundamental relation of production that connects relatively local events" (*ibid.*: 254-255); see also Marcus (2005: 43). On the causal argument for physicalism, see Stoljar (2010: ch. 11).

mental causation on the basis of physics.[29] He noted: "Why should mental causes of physical effects interfere with physical system if they do not consist in physical processes? Thinking that they must surely depends heavily on thinking of mental causes on a physical model – as providing an extra 'bump' or transfer of energy on the physical effect. In such a context, instances of 'overdetermination' – two cases having the same effect – must seem to be aberrations. But whether the physical model of mental causation is appropriate is part of what is at issue" (Burge 1992: 37).[30]

Even when, from a physicalistic perspective, there is a hint at the possibility of a genuine mental causation *in another possible world*, it always seems to be described according to something which is very similar to the "physical," or "mechanical" model. So, for instance, we are invited to "imagine a world where there are sui generis mental forces (bits of matter accelerate in otherwise inexplicable ways in the brain)." I am not sure that this represents the best way of conceiving our mental causation, i.e. as a quantifiable force which is entitled to produce the work of 'accelerating' (a typical parameter from physics) "bits of matter" for lack of other (physical) explanations (Spurrett, Papineau 1999: 28).[31]

[29] Regarding this point, see also Burge (1993: 358)," who notes: "Demanding that there be an account of mechanism in mind-body causation is tantamount to demanding a physical model for understanding such causation. It is far from obvious that such a model is appropriate. It is not even obvious why any model is needed. […] The demand for a mechanism is tantamount to an implicit demand for materialist solution;" see also the *Postscript* (2006: 368-369, 375-382, espec. 380): "Kim asks, given the physical cause is 'sufficient', what 'work' remains for the mental cause […] The notion of work has home both in physics and in talk of physical labor […] Here mental causation would be implicitly regarded as a form of physical causation. In such a role it can easily seem to be an intrusive, competing, physical-like cause." This is rather independent from other theses advanced by Burge (which I do not necessarily agree with) on mental causality and the (non)relationship between psychological and physical explanations.

[30] This is close to what Eric Marcus has chided as a "telekinetic *view of mental causation*" (2005: 42-46). In a similar vein, David Hodgson has pointed out that "the problem with all this is that it assumes that any functional or causal role, which conscious experience could have, could be fully realised by a mechanism […] But a (perhaps the) central question about the mind is whether or not conscious experiences have causal roles or functions, which cannot be performed or realised or even simulated by mechanisms of the type studied by the objective sciences" (1997: 126); Lowe makes an analogous point in (2000: 583). On the "mechanistic paradigm," see Dupré (2001).

[31] It is difficult to see what reason could ever urge us to consider this kind of causation both genuinely mental and genuinely irreducible *in principle* to physics. Note that what makes such "mental forces" *sui generis* and irreducible in principle to physics cannot be what makes something "mentally identifiable" in *our* world. This is because physicalists must assume that it is not a problem to reduce what makes something "mentally identifiable" in our world to physics, otherwise our experience would be full of irreducible phenomena. What kind of property should therefore make these *sui generis* forces both 'mental' and 'irreducible to physics'?

Nothing in my words is intended to deny that a lot of philosophical work is needed to render the hypothetical relationship between physical and non-physical causality (and also the proper nature of both) more intelligible. On the contrary, it is precisely for this reason that it is fundamental to underline that:

– It would be wrong to set up a research project on this topic in such a way that the understanding of a cause which operates in our world is seen as in principle equivalent to showing that it is isomorphic to a physical kind of cause (and so constitutes another instance of a physical cause). It would be even worse to assume that the *nexus* between *relata* with a causal relationship must have the same structure such as that of the mechanical causation as traditionally conceived in the past. This is an especially insidious and hidden risk because it blends easily with a common-sensical attitude to presume that, even if there are non-physical causes, the structure of all causal processes may always be objectivated, measured, and 'reified' according to the unwitting paradigm of the 'folk physics'.
– It would be equally wrong to set up research in such a way that it is already (more or less explicitly) assumed that the chains of physical causes are sufficient for there to be conscious, voluntary actions. What can be experimentally proved is that naturalistic causality may be sufficient to obtain some unconscious and unintentional events concerning our body. Those events, nevertheless, are in fact analogous to conscious, voluntary actions *only* for the portion of the behavioural outputs that natural sciences can deal with. Besides, supporters of genuine mental causation are not forced to assume that each body movement (even in a conscious condition) actually requires a role for non-physical causation: a mere muscular reflex due to a state of pain might have nothing to do with non-physical causation of the mind, even if the subject is conscious (see also: Bishop 2006). However, the absence of mental, non-physical causality seems to be exactly what distinguishes a reflex, i.e. a mere *movement*, from a conscious, intentional action, i.e. a real *gesture*. Perhaps this kind of distinction can contribute to throwing light on the role of mental causation.

To sum up, it is likely that we are confronted here with a relative autonomy between the understanding of *what* a cause is, and the understanding of *how* it actually operates in our world. We do not have any problem in understanding what we are saying when we assert that a certain conscious state is *a* cause of a certain physical event, and conscious states do have a role in many common-sense causal explanations. What

is inadequate is our understanding of how mental causes may work, interacting with the world of objects described by physics. However, this does not allow us to either impose physical standards for the legitimacy of mental causation in advance, or to identify the functioning of every causal process with the work of a physical mechanism.

It is also possible that we are faced here with a natural limit which precludes a satisfactory understanding of the conditions of possibility of there being mental causation in our world. However, it would be better to think that in this case, more than having evidence of the *causal closure* of our world, we have evidence of the "cognitive closure" of our minds (C. McGinn). And, in any case, it seems preferable to maintain a conceptual framework that allows for the possibility of an original mental causation (without a complete understanding of it), rather than to adopt a framework where it is eliminated (without any real reduction).

3.3. *All You Need Is Success?*

A misconception of mental causality also emerges in the more significant argument that is presented in favour of the causal closure thesis: the argument based on success. David Papineau has suggested that the major source of legitimacy for the causal closure thesis is the success and the generalised coverage of phenomena that characterize recent science (see, for instance: Papineau 2001; Id. 2002: Appendix; Montero 2003: §§ 3-5). Somehow, scientific progress will eventually make the principle of causal closure rather obvious (even if not compelling), as science neither has a need for an irreducible mental causality, nor leaves a place for it in its description of the world.

It is certainly true that the extraordinary advancements of experimental science have a deep influence on the conviction that physicalism is true and have also showed how experimental science can give a proper, fundamental contribution to the knowledge of countless phenomena. However, as we have noted, it seems that with all its success, science does not only give a framework of plausibility to certain theses here, but also substantially determines their meaning and many ways of thinking about reality and objects. In so doing, it also assumes a role that goes far beyond its range and its mere acceptance. It seems that calling on the success of modern science in order to legitimize the causal closure principle goes beyond what is justified. On the one hand, some scholars see the possibility of an argument for a true, irreducible mental causation

precisely in certain theories of contemporary physics.[32] On the other hand, and more importantly, the fact is that, precisely because of the success of physics, an improper role often seems to be assigned to science. Take, for instance, Papineau's inference from our mature (neuro)scientific explanations to the completeness of the non-mental (CNM). He argues that, "given that one would expect sui generis mental causes to turn up under physiological investigation, if they exist, the evidence for CNM is our failure, after extensive and prolonged physiological investigation, to come across any such mental causes" (Montero, Papineau 2005: 234-235; Papineau 2002: 253-254).[33]

So the fact that scientific disciplines have at the same time been successfully developed and have not found traces of such a mysterious mental causality would count as evidence against its existence. But what would happen if it 'turned up' in a specific way through research based on the method of experimental sciences? It seems that for this very reason we would face a new kind of 'physical' causality, since I presume that supporters of the 'via negativa' will not concede that something can be discovered by and incorporated within physics, while not being 'physical'. This means that, in both cases, mental causes do not have any space. Once again, a perspective according to which the naturalistic exploration of the world is considered as the best guide to its overall ontology predominates here up to the point that even the existence of what is in principle beyond it has to be proved, in some way, by this naturalistic exploration.[34] It might be stressed that the fact that mental causation is not identified through scientific exploration is not a proof of its inexistence. On the contrary, we could point out that if it had been especially detect-

[32] See Chalmers (2003), which refers to non linear collapses of the wave function. According to Chalmers, there is also another relevant argument from the history of physics against the physicalist's objection to original mental causation: "It is sometimes objected that physical and mental states could not interact, since there is no causal nexus between them. But one lesson from Hume and from modern science is that the same goes for any fundamental causal interactions, including those found in physics. Newtonian science reveals no causal nexus by which gravitation works, for example [...] And the same, presumably, applies to fundamental psychophysical laws: there is no need for a causal nexus distinct from the physical and mental properties themselves" (*ibid*.: 125).

[33] A more accurate discussion of the point can be found in Montero (2003: 182-186); see also Ney (2008: 10).

[34] Eric Marcus underlines a distortion that "comes in part as a result of thinking that mental causation, if it be causation at all, must work the way physical causation does. Under the influence of this distortion, we expect to find the same kind of evidence for the presence of mental causes as we find for the presence of physical causes" (2005: 42).

ed by science, then it would not be 'mental' causation in the proper sense. At the same time, if from the point of view of physics something were missing in scientific investigations and explanations, physics, by definition, would look for 'more of the same'; it would look for further 'physical' causality.

There is something to ponder here: if one *should not* expect mental causes to be detected in a specific way through scientific investigations, it is also because one should not imagine mental causation merely as a proper remedy for internal voids within the physical description of the world, as if it were simply an *analogous and homogeneous* substitute of physical causality.[35] It seems, once again, rather improper to conceive mental *causality* on the basis of the model of physical causality, and to imagine, for instance, that it can be added to or subtracted from it. The point is that the possibility of both properly conceiving mental causation, and calling on it, appears to be external and not internal to science: physics, *qua talis*, can neither detect it, nor call on it for its proper explanatory tasks. This is the reason why its absence in physical explanations should not be considered as evidence against it. After all, we should not expect that what manifests the external limits (of the method) of physics may appear as an internal object or evidence *within* physics.

However, one might reply that the argument based on the success of modern science is not primarily based on the fact that science has not found an original, irreducible mental causation, but that it is now able to offer a 'full coverage' of all phenomena we experience without any need of a reference to it. Consequently, it is likely that an influential, irreducible mental causation does not exist because a complete scientific image of the world neither makes reference to it, nor leaves any place for it.

Concerning this, it is worth lingering on what has, in part, already emerged. As we have seen, both the need for and the reference to genuine mental causation should not be placed within physics, but outside it and even beyond it. And this also means that the necessity or the redundancy of it has to be judged in light of the overall epistemic condition and needs of subjects and not with regard to physics (or science) only.

[35] A similar point has been made by Marcus (2005: § 3); at the same time, I do *not* share (among other things) his idea that the completeness of physics is possible, and that it "poses no problem" for the existence of irreducible mental causation. On the contrary, I assume that the existence of non-redundant mental causation renders all possible physicalistic (as well as naturalistic) accounts of our mind and of the causal history of our world *incomplete* (even if, from another point of view, this does not imply that physics is, because of this fact, incomplete as regards its possibilities and proper aims).

Now, from this point of view, it cannot be claimed that a redundancy of explanations based on irreducible mental causation has emerged. On the contrary, everyone commonly uses genuine explanations that are indispensable to their lives but are (at least at the moment) impossible to reduce to or surrogate with equipollent physicalistic or naturalistic explanations. It is also for this reason that many people are looking very passionately for a "synoptic" understanding of both.

While physicalists suggest that it is because of its exhaustive account of our experience that physics does not have an internal need to rely on non-physical causality, one may suggest, on the contrary, that physics does not have this *internal* need precisely because it gives us only a partial, regional account of our experience. Therefore, insofar as they are proper effects of mental causation, certain phenomena are not a proper part of the *explanandum* of physics. And only a preliminary option for physicalism might lead to the idea that what lies outside the physical domain (present or future) does not exist, or does not matter, or cannot interact with our world. This is as though the events of our world would result in being causally closed only because it has already been assumed that they really consist of only those aspects that physics may in principle describe and fully account for.[36]

Hopefully, there is no need to add that my analysis (§§ 3.1-3.3) represents an apology for irreducible, original non-physical causation, intended as *a true fact* of our world: a fact that exists independently from every theory concerning its role and that, if it exists only within subjects, it exists with no less 'reality' (whatever this means) than every other genuine phenomenon in our world. Original, influent, irreducible mental causation, therefore, is *neither* simply a way to rationalize our behaviour, nor the mere effect of a certain explanatory "stance" (Dennett) or of our need for teleological judgments. On the contrary, if original mental causation were absent, the history of the world would have been different: it would have been full of unwritten books, unwaged wars, and unperformed courageous, as well as cowardly, actions. Without non-physical causality, genuine conscious and voluntary actions would have no beginnings in our world. Furthermore, the fact that we may only describe and interpret what a voluntary gesture really is within the conceptual framework of intentional explanations is a consequence of the role of non-physical causation in our world, and *not* what non-physical causation

[36] "Universal physical explanation seems possible to the physicalist only because he has already got the idea that the physical world is the only world there is" (Stroud 1987: 276).

consists in. Therefore, physics (natural science) cannot give us a *complete* causal account of what happens in our world, since something that in principle lies outside its constitutive domain is significant in our world. Which is another way to say that all causes that physics can call upon are *not* enough to bring about all the effects that we observe in it.

All things considered, what may appear causally closed is the world of 'physics': and there is no reason to assert that it is our (entire) world.

4. *The Limits of our Physicalistic Language Do not Mean the Limits of our World*

To sum up, it seems that the causal closure thesis may assume four main roles:

[1] It could be an empirical hypothesis that has to be verified, step by step, in each case and therefore cannot be used as a guide to the analysis of each case or of each causal explanation. On the contrary, it will result as being justified only if it is confirmed by facts. From this perspective, however:

– (i) the causal closure thesis cannot be considered as a principle, and it would be unfair to use it as evidence against explanations that do violate it, or as a premise in an argument;
– (ii) besides, there doesn't seem to be much support for such an hypothesis at the moment; in fact, it cannot be implemented in our explanatory practices, because we systematically make use of non scientific causal explanations that we are not able to reduce or eliminate.

[2] Considered as an ontological principle, the causal closure thesis seems to presuppose a substantial part of what has to be proved. On the one hand, what is 'physical' is characterized (positively or negatively) in relationship to the properties of the basic ontology of 'physics' (present and/or future); on the other hand, genuine and influent causality in our world is already characterized by a need for the kind of characteristics that only 'physical' causality actually has.

[3] Considered as an epistemological principle, the causal closure thesis seems to be a methodological truism: *as far as* physics may legitimately establish, describe and explain causal *nexa*, they are for this very reason 'physical' *nexa* of causality. That is to say that the physical world (considered as the world fully described by physics) is characterized by the presence of physical causes only (that is of causes fully described by physics). But there is no reason to think that this 'physical' world is *our*

entire world, nor to assume that phenomena such as intentional actions are fully described by physics, and therefore only have causes fully described by physics. This simply means that the only kind of causality to which physics does and may refer is 'physical' causality.

[4] Considered as a meta-philosophical principle or a stance/attitude (see: Melnyk 1997; Poland 2003: § 2; Ney 2008), the causal closure thesis, as van Fraassen already suggested, seems to express a preference for a metaphysics (of causality) "guided by physics" (van Fraassen 2002: 59).[37] As noted by van Fraassen, there would be no "false consciousness" in this case: a stance is not a factual claim, therefore it is not adopted because of its recognized truthfulness, but rather as a personal orienting option going far beyond the corpus of available data and theories. Nevertheless, I would like to suggest that this stance is neither 'consubstantial' to the acceptance of contemporary physical theories, nor appropriate at all.

As we have already noted, the causal closure thesis, interpreted as a meta-philosophical stance, could be summarized with Sellars's famous dictum (1963: § 41): "in the dimension of describing and explaining the world, science is the measure of all things, of what is that it is, and of what is not that it is not" (that is, at least, the measure of all influential causes). Physics (or science in general) will therefore not only be adequate as a matter of fact for all causal phenomena in our world, but will be in principle the only criterion for deciding what is real: the substantive has irreversibly become an adjective, the *physical* world is the world of *physics* and, besides, this world is supposed to be *our* world, the only one. But this outcome seems to presuppose two very controversial conceptual passages:

– Firstly, what we can call the fallacy of hypostatizing the epistemic characteristics: a certain ultimate metaphysical nature (intrinsic and monadic property) is assigned to what results in being knowledgeable (or unknowledgeable) by us within a certain conceptual and methodological framework (relational and dyadic property). What is 'physical' is then what belongs or may belong to 'physics', and what is 'non-physical' is what does not belong and cannot belong to 'physics.'
– Secondly, what Thomas Nagel labelled as "an epistemological criterion of reality" according to which "only what can be understood in a certain way exists," that is only what is knowledgeable within the framework of a discipline effectively exists (Nagel 1986: 15).[38]

[37] Stroud (1987: 277) used the expression "scientistic faith."

[38] And see also Barry Stroud's pertinent remark: "even if we had an independently specifi-

On the contrary, epistemology may teach us that what from the internal point of view of physics (or of science in general) appear as insurmountable barriers, may result in being, from the point of view of humans, only the limits of *one* of their methodological resources and frameworks. And, paraphrasing Wittgenstein, the limits of our physical language (as well as those of our scientific language in general) do not mean the limits of our world.[39]

References

Antony, Louise M., Hornstein, Norbert, eds., 2003, *Chomsky and His Critics*, Blackwell, Oxford.

Bishop, Robert C., 2006, "The Hidden Premise in the Causal Argument for Physicalism", in *Analysis*, 66: 44-52.

Burge, Tyler, 1992, "Philosophy of Language and Mind: 1950-1990", in *The Philosophical Review*, 101: 3-51.

Burge, Tyler, 1993, "Mind-Body Causation and Explanatory Practice", in Id., 2007, *Foundations of Mind. Philosophical Essays*, vol. 2, Oxford, Clarendon Press: 344-362.

Burge, Tyler, 2006, "Postscript to "Mind-Body Causation and Explanatory Practice"", in Id., 2007, *Foundations of Mind. Philosophical Essays*, vol. 2, Clarendon Press, Oxford: 363-382.

Campaner, Rafaella, Galavotti, Maria Carla, 2007, "Plurality in Causality", in Machamer, Peter, Wolters, Gereon, eds., *Thinking about Causes. From Greek Philosophy to Modern Physics*, University of Pittsburgh Press, Pittsburgh: 178-199.

Cartwright, Nancy, 2004, "Causation: One Word, Many Things", in *Philosophy of Science*, 71: 805-819.

Chalmers, David, 2003, "Consciousness and its Place in Nature", in Stich, Stephen P., Warfield, Ted A., eds., *The Blackwell Guide to Philosophy of Mind*, Blackwell, Oxford.

Chisholm, Roderick, 1964, *Human Freedom and the Self* (The Lindley Lecture), University of Kansas Press, Lawrence.

able idea of the physical science, and so could say what makes a science a physical science, we would still need some reason to think that facts expressed in the terms of such physical science are the only facts there are" (1987: 277).

[39] I am very grateful to all the participants in the workshop. A special thanks to Robert Bishop and Raffaella Campaner for useful observations, and to Alexandra Shakespear for the linguistic revision.

Crane, Tim, 1991, "Why Indeed? Papineau on Supervenience", in *Analysis*, 51: 32-37.

Crane, Tim, 1994, "Against Physicalism", in Guttenplan Samuel, ed., *A Companion to the Philosophy of Mind*, Blackwell, Oxford: 459-483.

Crane, Tim, 2001, *Elements of Mind. An Introduction to the Philosophy of Mind*, Oxford University Press, Oxford.

Crane, Tim, Mellor, David Hugh, 1990, "There is No Question of Physicalism", in *Mind*, 99: 185-206.

Crook, Seth, Gillett, Carl, 2001, "Why Physics Alone Cannot Define the 'Physical': Materialism, Metaphysics, and the Formulation of Physicalism", in *Canadian Journal of Philosophy*, 31: 333-360.

Dowell, Janice L., ed. 2006, *Philosophical Studies*, 131, 1.

Dupré, John, 1993, *The Disorder of Things. Metaphysical Foundations of the Disunity of Science*, Harvard University Press, Cambridge, MA.

Dupré, John, 2001, *Human Nature and the Limits of Science*, Oxford University Press, Oxford.

Gillett, Carl, Witmer, Gene, 2001, "A Physical Need: Physicalism and the Via Negativa", in *Analysis*, 61, 4: 302-309.

Heil, John, 2004, *Philosophy of Mind. A Contemporary Introduction*, Routledge, London.

Hempel, Carl G., 1980, "Comments on Goodman's Ways of Worldmaking", in *Synthèse*, 45: 193-199.

Hitchcock, Christopher, 2003, "Of Humean Bondage", in *The British Journal for the Philosophy of Science*, 54: 1-25.

Hitchcock, Christopher, 2007, "How To Be a Causal Pluralist", in Machamer, Peter, Wolters, Gereon, eds., *Thinking about Causes. From Greek Philosophy to Modern Physics*, University of Pittsburgh Press, Pittsburgh: 200-221.

Hodgson, David, 1997, "The Easy Problems Ain't So Easy", in Shear, Jonathan, ed., *Explaining Consciousness: The Hard Problem*, MIT Press, Cambridge, MA: 125-131.

Jones, Kile, 2008, "The Causal Closure of Physics: An Explanation and Critique", in *World Futures*, 64, 3: 179-186.

Kim, Jaegwon, 1996, *Philosophy of Mind*, Westview Press, Boulder, CO.

Kim, Jaegwon, 2005, *Physicalism, or Something Near Enough*, Princeton University Press, Princeton.

Lowe, E. Jonathan, 2000, "Causal Closure Principles and Emergentism", in *Philosophy*, 75: 571-585.

Loewer, Barry, 2001, "From Physics to Physicalism", in Gillett, Carl, Loewer, Barry, *Physicalism and Its Discontents*, Cambridge, University Press Cambridge: 37-56.

Loewer, Barry, 2007, "Mental Causation, or Something Near Enough", in McLaughlin, Brian P., Cohen, Jonathan, eds., *Contemporary Debates in Philosophy of Mind*, Blackwell, Oxford: 243-264.

Marcus, Eric, 2005, "Mental Causation in a Physical World", in *Philosophical Studies*, 122: 27-50.

Malatesti, Luca, 2005, "Come formulare il fisicalismo in filosofia della mente", in *Arkete. Rivista di studi filosofici*, 1: 29-49.

Mele, Alfred, 2009, *Effective Intentions. The Power of Conscious Will*, Oxford University Press, Oxford.

Melnyk, Andrew, 1997, "How to Keep the 'Physical' in Physicalism", in *Journal of Philosophy*, 94: 622-637.

Montero, Barbara, 1999, "The Body Problem", in *Noûs*, 33: 183-200.

Montero, Barbara, 2003, "Varieties of Causal Closure", in Walter, Sven, Heckmann Heinz-Dieter, (eds.), *Physicalism and Mental Causation*, Imprint, Exeter: 173-190.

Montero, Barbara, 2006, "Physicalism and the Conservation of Energy", in *Dialectica*, 60: 383-396.

Montero, Barbara, 2009, "What is the Physical?", in McLaughlin, Brian P., Beckermann, Ansgar, Walter, Sven, eds., *The Oxford Handbook of Philosophy of Mind*, Clarendon Press, Oxford: 173-188.

Montero, Barbara, Papineau, David, 2005, "A Defence of the Via Negativa Argument for Physicalism", in *Analysis*, 65: 233-237.

Nagel, Thomas, 1986, *The View from Nowhere*, Oxford University Press, Oxford and New York.

Ney, Alyssa, 2008, "Physicalism as an Attitude", in *Philosophical Studies*, 138: 1-15.

Norton, John, 2007, "Causation as Folk Science", in Price, Huw, Corry, Richard, eds., *Causation, Physics, and the Constitution of Reality. Russell's Republic Revisited*, Clarendon Press, Oxford: 11-44.

Papineau, David, 1990, "Why Supervenience?", in *Analysis*, 50: 66-71.

Papineau, David, 1991, "The Reason Why: Response to Crane", in *Analysis*, 51: 37-40.

Papineau, David, 2001, "The Rise of Physicalism", in Gillett, Carl, Loewer, Barry, eds., *Physicalism and Its Discontents*, Cambridge University Press, Cambridge.

Papineau, David, 2002, *Thinking about Consciousness*, Clarendon Press, Oxford.

Papineau, David, 2009, "The Causal Closure of the Physical and Naturalism," in McLaughlin, Brian P., Beckermann, Ansgar, Walter, Sven, eds., *The Oxford Handbook of Philosophy of Mind*, Clarendon Press, Oxford: 53-65.

Pettit, Philip, 1993, "A Definition of Physicalism", in *Analysis*, 53: 213-223.

Poland, Jeffrey, 2003, "Chomsky's Challenge to Physicalism", in Antony, Louise

M., Hornstein, Norbert, eds., *Chomsky and His Critics*, Blackwell, Oxford: 29-48.

Sellars, Wilfrid, 1963, "Empiricism and the Philosophy of Mind", in Id., *Science, Perception and Reality*, Routledge, London.

Searle, John, 1992, *The Rediscovery of the Mind*, MIT Press, Cambridge, MA.

Schaffer, Jonathan, 2004, "Trumping Preemption", in Collins, John D., Hall, Edward J., Paul Laurie A., eds., *Causation and Counterfactuals*, MIT Press, Cambridge, MA: 59-74.

Spurrett, David, Papineau, David, 1999, "A Note on the Completeness of 'Physics'", in *Analysis*, 59: 25-29.

Stoljar, Daniel, 2001 (revised 2009), "Physicalism", in *Stanford Encyclopedia of Philosophy*, http://plato.stanford.edu/entries/physicalism/

Stoljar, Daniel, 2010, *Physicalism*, Routledge, London and New York.

Strawson, Galen, 2003, "Real Materialism", in Antony, Louise M., Hornstein, Norbert, eds., *Chomsky and His Critics*, Blackwell, Oxford: 49-88.

Stroud, Barry, 1987, "The Physical World", in *Proceedings of the Aristotelian Society*, 87: 263-277.

Stroud, Barry, 2000, *The Quest for Reality. Subjectivism & the Metaphysics of Colour*, Oxford University Press, Oxford.

Van Fraassen, Bas, 1996, "Science, Materialism, and False Consciousness", in Kvanvig, Jonathan L., ed., *Warrant in Contemporary Epistemology: Essays in Honor of Alvin Plantinga's Theory of Knowledge*, Rowman and Littlefield, Lanham: 149-182.

van Fraassen, Bas, 2002, *The Empirical Stance*, Yale University Press, New Haven.

Vicente, Augustin, 2011, "Current Physics and 'the Physical'", in *The British Journal for the Philosophy of Science*, 62: 393-416.

Wilson, Jessica, 2006, "On Characterizing the Physical", in *Philosophical Studies*, 131: 61-99.