

ON UNDERSTANDING PHYSICALISM*

*Julia Telles de Menezes***
juliatelles@gmail.com

ABSTRACT *This paper aims at exposing a strategy to organize the debate around physicalism. Our starting point (following Stoljar 2010) is the pre-philosophical notion of physicalism, which is typically formulated in the form of slogans. Indeed, philosophers debating metaphysics have paradigmatically introduced the subject with aid of slogans such as “there is nothing over and above the physical”, “once every physical aspect of the world is settled, every other aspect will follow”, “physicalism is the thesis that everything is physical”. These ideas are very intuitive but they are, of course, far from being a satisfactory metaphysical conception of Physicalism. For that end, we will begin with the definition of physicalism as the thesis that everything is physical, following Stoljar; we should be able to respond to one central question: how to interpret the physicalist claim that everything is physical.*

Keywords *Physicalism, Definition, Metaphysics.*

RESUMO *Este trabalho procura expor uma estratégia para organizar o debate em torno do Fisicismo. O ponto de partida (Stoljar, 2010) será a noção pré-filosófica de Fisicismo tipicamente formulada em forma de slogans. Formas paradigmáticas como: “Não há nada além do físico”, “Uma vez que todos os aspectos físicos do mundo foram estabelecidos, todos os outros aspectos seguirão”, “Fisicismo é a tese de que tudo é físico”. Tais ideias podem*

* Artigo submetido em 24/05/17. Aceito em 24/07/17.

** Universidade Federal do Rio de Janeiro. Rio de Janeiro – RJ – Brasil.

ser compreendidas intuitivamente, mas estão longe de caracterizarem uma concepção metafísica satisfatória de Fisicismo. Para tal fim, começaremos com a definição de Fisicismo como a tese de que tudo é físico para assim respondermos a uma questão central: como interpretar a afirmação de que tudo é físico.

Palavras-chave Fisicismo, definição, metafísica.

1 Prelude

The dominant world-view in the pre-modern era used to favor explanations of certain natural events that involved all kinds of immaterial entities such as gods, angels and magical creatures in general. Back then, the way to explain seizures, for example, was to consider it as a sign of demoniac possession; once, the cycle of the sun was attributed to Apollo and his horse moving across a flat Earth; what is now understood as a mental health issue was once related to curses, spells or something equally spooky. At a certain point in history, a scientific turn took place so that it progressively eliminated such explanations as inadequate accounts of the relevant phenomena. Such magical explanations were replaced by scientific ones. One requirement for the best explanation would be that it should not be easily replaced by other explanations without any loss in explanatory force. The scientific world-view that ties seizures with neurological dysfunction, the solar cycle with the earth's movement around the sun and around itself, or even that psychiatric issues should be explained in terms of genetic and environmental influences are far better suited to replace the magical explanations previously available. The scientific turn is what Weber called the phenomenon of 'disenchantment of the world' (*Entzauberung der Welt*). Modern science endorses, as background methodology, the existence of a physicalist world-view, which is established after a change of paradigm in our world, that is to say, after the disenchantment of the world. Clearly, this change in paradigm is not accomplished without resistance and controversy. In the past, many were burned at the cross for suggesting such a radical change of view. Nowadays the reaction is milder, yet the questions are still source of great dispute. Nevertheless, this 'new' physicalist world-view is still a pre-philosophical view about what there is in the world. The physicalist world-view is sometimes understood as a package of views distinct from the metaphysical thesis of physicalism. It involves, for example, the idea that the methodology employed in natural sciences will provide complete theoretical knowledge of the world and that this way of understanding the world will deliver a final theory

of everything. It also claims that this complete theory of everything will be, like natural sciences, objective; hence, it will reduce subjective perspectives to objective vocabulary. The view also involves the idea that all relevant explanation is physically reductive; that physics is general enough to reductively explain events that special sciences already explained, because every event is held to have a physical cause; and it is sometimes added to this the idea of atheism. Nevertheless, it would be a mistake to think that physicalism implies all of these views. For example, atheism is not implied nor implies physicalism, since it is consistent with physicalism and anti-physicalism. On the contrary, physicalism is, at least neutral, regarding many of the views. This is why we should distinguish the physicalist pre-philosophical world-view from the metaphysical thesis of physicalism.

One thing that the physicalist world-view and the metaphysical thesis of physicalism have in common is the strong intuition that physicalism generates a certain tension when confronted with some central features of our everyday life. This monist view is *prima facie* incompatible with features such as: abstraction, intentionality, phenomenality, normativity etc. If we grant that the tension is real, the core question of physicalism arises of how to accommodate such aspects of our everyday life (e.g. mentality and phenomenality) in this monist, physical world. This will be our starting point in the task of finding a more sophisticated definition of physicalism that is able to include features of mentality and phenomenality. If such definition is not available, the physicalist needs to provide good reasons to discard mentality and phenomenality. There are at least three routes out of this problem. One is to simply deny physicalism and, instead, recommend a form of dualism. This will resolve the tension, but it will generate new and serious problems as that of explaining non-physical influence in a physical world which is causally closed. Another possibility is to preserve physicalism but to abdicate from mentality by treating it as an illusion that should be eliminated, just like 'magical' explanations were once eliminated. The idea is that the vocabulary that we still use to refer to aspects of mentality can be replaced by a strict physicalist code. The problem with this proposal is that it asks us to give up something too essential to our common life, namely, mental talk. We do want to preserve the mental vocabulary as we do want to preserve mentality. The best way to preserve them is to define a *compromise* version of physicalism that accounts for mentality from a physicalist point of view. Before any of this work is done we have to find a preliminary definition of physicalism as a metaphysical thesis to which one can argue for and against. This is the aim of this paper. What kind of definition of physicalism should we have in order to account for mentality?

2 Defining Physicalism

Physicalism is roughly the metaphysical thesis that claims that the world is fundamentally physical. The term ‘physicalism’ was first introduced by Carnap and Neurath to designate instead a semantic thesis: every sentence describing the mental can be translated into sentences in a physical vocabulary. So originally, ‘physicalism’ was a semantic thesis, whereas ‘materialism’ was the term used to designate a metaphysical thesis, that is, a thesis about the nature of the world. Materialism was taken to be the doctrine which claimed that everything is matter, whereas the notion of ‘matter’ is historically understood as something that is extended, located in space and time. An old-school materialist is a philosopher who claims that everything is matter in this sense. However, modern physics renders this view of matter as false by acknowledging the existence of all sorts of physical entities which have no mass or are not extended in space. For old-school materialists, such physical entities would be, per definition, immaterial. Other philosophers do not see this as a problem, in fact they prefer to stick with the traditional way of characterizing the monist metaphysical thesis, regardless of the role played by progress in physical sciences. Be that as it may, the role that physical sciences play is one among many other reasons to prefer the term ‘physicalism’ over ‘materialism’ as the expression that designates the metaphysical claim about what there is in the world. It is important to notice that, while some philosophers prefer to name the monist metaphysical thesis in question ‘physicalism’ and others prefer ‘materialism’, there is still a third group, which chooses to talk in terms of ‘naturalism’ in an attempt to include other natural sciences besides Physics, like Biology, for instance. From this point on I choose to use ‘Physicalism’ to refer to the metaphysical claim in question.¹

This paper consists of finding a metaphysical conception of physicalism, that is, a more sophisticated definition than those available in the form of slogans. The strategy for defining physicalism consists of asking two questions; *the interpretation question* and the *truth question*. The interpretation question is concerned with how are we to understand the thesis of physicalism, whereas the truth question is concerned with the plausibility of physicalism. My focus will be only on the matters of how to understand physicalism, I will remain silent on the truth question. The interpretation question should be formulated so: What does it mean to say that everything is physical?

1 The way the debate is formulated in contemporary philosophy of mind, these terms may be used interchangeably.

3 The interpretation question: what does ‘everything is physical’ mean?

The question about how we should understand the thesis of physicalism unfolds into three different questions:²

- (a) Scope question: What does it mean to say that *everything* is physical?
- (b) Base question: What do we mean by *physical*?
- (c) Relation question: What is the relation between *everything* and *physical*?

The scope question leaves the specification of the conception of ‘physical’ open and inquires about the domain of ‘everything’ satisfying the condition of being physical, whereas the base question inquires about the conception of ‘physical’, clearly central to physicalist theories. The relation question asks about the connection between the base and the scope, that is, the relation between physical and everything. Is it identity? Is it supervenience? Or even emergence?

3.1 The scope question

What is it for *everything* to be physical? Although ‘everything’ works as universal quantifier, it certainly does not quantify over absolutely everything in the world like in the statement ‘Everything is identical to itself.’ To grant that ‘everything’ in our slogan has an unrestricted scope is to assume that *absolutely* everything (concrete, abstract, property, particular etc.) is physical. That could hardly be true.

The unrestricted quantification is problematic especially when we consider abstract entities like numbers. Numbers are not physical in the sense that chairs are, for instance. Just as institutions like universities or a court house are also not physical in the way that paradigmatic physical objects are (see Stoljar, 2010, p. 30). It is very hard to ascribe physicality to such complex entities, or, at least, it is highly implausible that the truth of physicalism depends on the truth of numbers and universities being physical. Some physicalists (who are also nominalists) may want to endorse the unrestricted quantification. They argue that since numbers cannot be physical, they do not exist. The problem with this suggestion is that one would have to erase from our ontology many other abstract entities such as complex entities. For this reason, some physicalists (who are not nominalists) think it is best to restrict the quantifier’s scope. So the scope question asks about the domain of the quantification operator: which things are properly physical?

2 Here I follow Stoljar’s suggestion (2009, 2010) of systematizing the discussion about physicalism.

One possibility is to restrict the domain to *concrete particulars*, where particulars are defined in opposition to properties. Properties are instantiated in particulars objects, whereas particulars are not instantiated at all. We can put it this way: things are particulars, whereas properties should be features of such things. Concrete entities are defined in opposition to abstract entities: concrete items are extended in time and space, whereas abstract things are not. Now, by restricting the quantification to concrete particulars we would have the following definition:

(1) *Concrete Particulars*: Physicalism is true if and only if every concrete particular is physical. (Stoljar, 2010, p. 33)

The problem with restricting physicalism to concrete particulars is that this definition overgenerates: it allows the inclusion of substance dualism in the scope of physicalism. A soul, for example, is usually defined as an unextended substance that could coexist with physical substances in the world. Restricting the domain of physicalism to concrete particulars does not prevent the inclusion of such paradigmatic non-physical entities. This would make physicalism true in a world populated with souls, which is not quite acceptable. Hence, restricting the quantificational operator of physicalism to concrete particulars overgenerates, it declares physical some properties which should be excluded.

Now what if we restrict our domain to properties? Then we would have something like this:

(2) *Properties*: Physicalism is true if and only if every property is physical. (Stoljar, 2010, p. 32)

This way of defining physicalism seems to avoid any kind of dualism. Since restricting the scope of physicalism to properties is inconsistent with property dualism, which is the claim that, although there is only one substance, which is physical, that physical substance might instantiate non-physical properties. And the substance dualism would also be ruled out, since, to think about non-physical substance is to think about it as instantiating non-physical properties. This is inconsistent with restricting physicalism to properties. Notwithstanding, this way of defining physicalism would make uninstantiated properties part of the physical realm. Some physicalists believe in the existence of uninstantiated properties. For example, 'being a perfect circle' is an uninstantiated property, since there are things that are circular, but not perfectly circular. Also the property of being the new Professor of Ethics in 2010 at Rio de Janeiro is a

property that could have been instantiated, but, as the selection of Professor failed in 2010, that remains as an uninstantiated property. Some philosophers believe that uninstantiated properties exist. Hence, if there are no souls, soul is an uninstantiated property, like the property of being a perfect circle. The property of being a soul would not be physical, but it would, nevertheless, according to the definition given above, be compatible with our definition of physicalism (2). That would make physicalism compatible with property dualism, still the definition of physicalism overgenerates again. It declares uninstantiated properties as part of the physical realm, which should be excluded. A slight modification would take care of the problem: it suffices that we restrict physicalism to *instantiated* properties:

(3) *Instantiated properties*: Physicalism is true if and only if every instantiated property is physical. (Stoljar, 2010, p. 33)

The formulation of physicalism in (3) avoids the problem of compatibility of property dualism with the existence of uninstantiated physical properties. Of course, we still need to specify the other parameters for our complete definition of physicalism.

3.2 *The base question: conception of 'physical'*

Perhaps, the most difficult problem, one would think, when engaging in the task of defining physicalism, is to define the embedded notion of 'physical'. The dominant conception of *physical* in the literature in the philosophy of mind ties the notion of instantiated physical property to an authoritative role of physical sciences. However, there are also other options to consider: define physical properties in terms of the properties instantiated in paradigmatic physical objects, or in terms of methods distinctive of the natural sciences, or by contrasting them with paradigmatic non-physical properties such as mental properties, spirits, souls, etc. We will take some time to explore a few of the possibilities to define 'physical property'.

3.2.1 Object-based definition

Object-based definition: A property is physical if and only if it is a property instantiated in *paradigmatic physical objects*.

This possibility consists in taking the classical route to define physical property in terms of paradigmatic physical objects. Of course, we are left with the task of defining paradigmatic physical objects. If we consider, for a moment,

what is the classical view of physical object we presumably arrive at a cluster of features like having certain form and size – being extended – and being located in space and time. An intuitive physical object must at least obey these criteria:

A thing is a material object if it occupies space and endures through time and can move about in space (literally move about, unlike a shadow or a wave or a reflection) and has a surface and has a mass and is made of certain stuff or stuffs. Or, at any rate, to the extent that one was reluctant to say of something that it had various of these features, to that extent one would be reluctant to describe it as a material object (Van Inwagen, 1990, p. 17).

However, modern Physics has established that not all physical objects can be measured by size and that not all physical objects have mass and extension. On that account, some physical objects are indeed located in space, but they lack mass. Modern physics renders the object-based conception of physical properties too naive, thus inapt to define physical properties. According to the object-based definition, many properties recognized by modern physics came out as non-physical, like being an electron, for example, a particle that has mass but is not extended. Since one of the features of paradigmatic physical objects is extendedness, an electron would be, by definition, a non-physical object.

Another problem with this definition of the physical is that it defines ‘physical property’ in a way that is compatible with property dualism. Property dualism is the claim that there is only one kind of substance, viz. the physical substance, which can be the bearer of physical and non-physical properties. Mental properties are instantiated by the physical substance, therefore, a physical property cannot be simply defined as a property of physical objects, since mental properties would then be, by definition, physical properties. Of course, if physicalism is true, mental properties *are* physical properties. The point is only that the result cannot be arrived at by sheer definition.

An additional objection often raised against the object-based conception of a physical property is called the problem of panpsychism. Roughly, panpsychism is the thesis that all physical objects are conscious beings, as well as all conscious properties are physical. Here is a passage where Lewis (1983) describes panpsychism:

It is often noted that psychophysical identity is a two-way street: if all mental properties are physical, then some physical properties are mental. But perhaps not just some but *all* physical properties might be mental as well; and indeed every property of anything might by at once physical and mental (Lewis, 1983, p. 362).

This extravagant idea is consistent with the truth of the object-based conception of physicalism, though its implausibility is quite clear: it is simply

strange to ascribe consciousness to simple physical objects like sofas or rocks.³ Though counter-intuitive, panpsychism is consistent with the truth of physicalism defined via reference to the object-based conception of physical property. The problem is similar to that concerning property dualism and the object-based definition of physicalism: if what defines a physical property is that it can be instantiated by a paradigmatic physical object; then being conscious is both a physical property and a mental property. This strategy overgenerates again, it includes as physical entities properties we should exclude.

3.2.2 Theory-based definition

To bypass the problems that arise with the object-based conception of the physical, an alternative definition of physical property is advanced by many philosophers: by ‘physical property’ we should mean what is within the range of the language of the physical sciences. Thus, physical sciences will play an authoritative role in defining the physical.

Theory-based conception: P is a physical property if and only if P is expressed by a predicate of a true physical theory.

Granting physical theories with the authority to determinate what a physical property is, is perhaps the most popular definition of ‘physical’. The idea is that we defer to physicists regarding the meaning of ‘physical’. According to the theory-based conception, statements usually made by physicists to explain some physical phenomena fix the reference of ‘physical property’. In this sense, the language of physics determines what the physical is. This sort of definition, although popular, raises serious problems of its own, many of them concerning the notion of physical sciences itself. What do we mean by ‘physical sciences’? Hempel is the first to launch the issue:

The language of *what* physics is meant? Surely not that of, say, 18th century physics; for it contains terms like ‘caloric fluid’, whose use is governed by theoretical assumptions now false. Nor can the language of contemporary physics claim the role of unitary language, since it will no doubt undergo further changes, too. The thesis of physicalism would seem to require a language in which a true story of all physical phenomena can be formulated. But it is quite unclear what is to be understood here by a physical phenomenon, especially in the context of a doctrine that has taken a determinedly linguistic turn (Hempel, 1980, p. 195).

3 More recently, philosophers who call themselves Panpsychists are not committed to such strong and extravagant doctrine. They usually defend the thesis that *some* fundamental physical entities are conscious. There is a variety of papers which seek to provide some plausibility to panpsychism (see Alter & Nagasawa, 2015).

Hempel's observations point to issues that are raised when we rely solely on the language of physical sciences to tell us what 'physical' means. First, choosing the language of a *particular theory* in physics does not allow us to capture the real spirit of physicalism, which is to claim a *general* thesis about the world. Many things that might be paradigmatic cases of physical objects or properties may fail to satisfy the condition of being physical in such a narrow sense (i.e. with regards to one particular physical theory).

What motivates the theory-based conception is the old idea of the 'unity of science' in which all theories will eventually be reduced to physics forming a complete unified science of everything, able to derive all scientific laws from one 'ever more adequate grand scheme'. Science's continuous change is presented as evidence for the implausibility of the so-called unity of science. Even within the physical sciences we have an amazing diversity of theoretical entities, properties and facts which require different methods of investigation. It is, therefore, quite implausible to think of unified science that integrates both astrophysics and genetics. Maybe this requirement of a unified language of science may be weakened so as to accommodate part of what we are looking for. But the main issue remains, that is the dilemma concerning the kind of physics presupposed in the attempted theory-based characterization of physicalism. Hempel objects that any theory-based definition of physicalism will be either trivially true or false. Is it present-day physics that holds this authoritative role? Or is it a future, complete physics? We know that current physics is subject to continuous change, since there is always the possibility of making progress by discovering new physical properties. If 'physical property' is defined by present-day physics, then properties discovered only by future physics would be, by definition, non-physical. So physicalism would be false. If we have in mind a future, complete physics, that is, a physical theory that explains everything, then genuine mental properties may have to be included in this final physics, making physicalism trivial. In sum, the first horn of the dilemma says that if physicalism is defined through present-day physics, then it is false. The second horn of the dilemma says that if physicalism is defined through future physics, then it is trivially true. Hence there is no possibility of coming up with a clear concept of physical that relies solely on the authoritative role of physical sciences. Thus, an adequate and non-trivial question of physicalism cannot even be formulated given the theory-based conception.

Hempel's dilemma is formulated as an objection to the general idea of physicalism. The dilemma is designed to yield the conclusion that the question of physicalism does not even make sense, for we cannot define a clear conception of the physical. Of course, one can avoid this objection by following another

course in the task of defining the physical. What the dilemma shows is a problem within the theory-based definition for physicalism, not physicalism itself. These remarks lead us to briefly glance at alternatives for defining ‘physical property’.

3.2.3 Method-based conception

An alternative to choosing between future or present-day physics would be to consider the possibility of defining physicalism by referring to the methodology of physical sciences: what if ‘physical property’ is defined by the language used in any science that applies the methodology of physical sciences? In this case, what would determine the meaning of ‘physical property’ is not physical sciences per se, but its methodology. Nevertheless, this approach is also subject to a dilemma similar to that posed by Hempel. Considering that methods in the physical sciences change over time, we might want to ask: when we refer to the method of physical sciences are we referring to present-day physical science or future physical science? We face the same dilemma as for the theory-based conception. In particular, if we fix methods of those presently adopted, we exclude items that should be ultimately recognized as physical.

3.2.4 *Via Negativa*

Yet another way to define physical in the context of the mind-body debate is to provide negative, contrastive definitions by referring to paradigmatic non-physical things: mentality, consciousness etc.. We might thus arrive at a list of mental properties and the like. The *via negativa* approach would look something like this:

Via Negativa: F is a physical property if, and only if, F is a non-mental property.

The major problem with this view is that it would imply eliminativism about mental properties. If a physical property is defined in terms of what is a paradigmatic non-physical property, such as mentality and phenomenality, there could not be a way of identifying physical properties and mental properties, since they would be, by definition, distinct. Consider that ‘pain’ is a paradigmatic mental state. Hence, the properties of ‘pain’ are by definition non-physical, since what defines physical is the fact that it is non-mental. If this is so, then we cannot even begin to make sense of the identification ‘pain is stimulation of c-fibers’, since ‘pain’ is non-physical and ‘stimulation of c-fibers’ is physical. Thus, the *via negativa* definition renders physicalism as incoherent.

So far we have seen that all extant attempts to define ‘physical’ fail. Each of them either overgenerates: they include properties that should be excluded

from the physical realm, or undergenerates: they exclude properties that should be included in the physical realm. The object-based conception seems to be too naive to be taken seriously, because it is based in ‘commonsense physics’ which is basically Newtonian mechanics applied to the megascopic world. Since there are physical properties that do not fall under the conception paradigmatic physical objects, the object-based definition of physicalism fail. Even if we overlook this first shortcoming, we are still left with a conception of physical that makes physicalism compatible with panpsychism and property dualism, since we might have mental properties figuring as properties of paradigmatic physical objects. The failure of this classical route to defining ‘physical’ leads to the search for alternatives. The most popular conception of ‘physical’ involves the authoritative role of physical sciences. Physics has complete authority in determining what, after all, is physical. This is initially a very attractive position since it relies on an important principle of the physicalist world view, i.e. that the body of physical sciences should be a complete doctrine. However, the theory-based conception of a physical property is susceptible to Hempel’s dilemma involving the conception of physical sciences we are presupposing in this definition: present-day physics makes physicalism false and future physics makes physicalism trivial. This objection is designed to yield drastic results for physicalism; since there is no coherent conception of the physical, physicalism cannot even be formulated. We have also explored the prospects of using the methodology of physical sciences as the central feature of the physical, but that is subject to a variant of Hempel’s dilemma. Finally, the *via negativa* which defines physical properties by contrasting it with mental properties has showed to be inadequate, for it has the consequence that the idea of physicalism is incoherent.

3.2.5 Revisiting the theory-based conception of the ‘physical’

My suggestion is to go back to the theory-based conception of physical, and examine some ways out of Hempel’s dilemma. One alternative for the physicalist is to resort to an indexical definition of ‘physical’: *that kind of thing* physics says there is. *That kind of thing* will change and develop with the progress of physics and so will our physicalist commitments. This will create an open-ended definition for physicalism in which ‘what is physical’ changes and makes progress along with the progress and changes of physical sciences. Consequently, physicalism becomes a floating doctrine. Indexical physicalism becomes a family of theses, each member individuated by an indexical. But then we turn into Hempel’s second horn of the dilemma: futures members of the family may render physicalism as a trivial thesis.

However, the open ended character allows the presence of disembodiment (minds without bodies) within a physicalist picture. Future physics might include properties which were once classified as non-physical properties, but are in the future classified as physical properties. Nevertheless, physicalists can resist the inclusion of such entities, since there is no strong empirical evidence for the existence of ghosts or parapsychological phenomena, it is very unlikely that someday they will be granted a physical status, for now physicalism should ignore such possibilities.

Another response close to the suggestion above is to insist that present-day physics is indeed complete or that it is at least rational to consider it as complete. This is proposed by Lewis:

It is a task of physics to provide an inventory of all the fundamental properties and relations that occur in the world. [...] We have no a priori guarantee of it, but we may reasonably think that present-day physics already goes a long way toward a complete and correct inventory. And we may reasonably hope that future physics can finish the job in the same distinctive style. [...] if we optimistically extrapolate the triumph of physics hitherto, we may provisionally accept that all fundamental properties and relations that actually occur are physical. This is the thesis of materialism (Lewis, 1994, pp. 51-2).

There is no structured argument to deny the first horn of Hempel's dilemma, rather, it is more of an intuition about the way we already treat physics: the intuition that it is rational to believe that present-physics is already complete. In fact, Lewis thinks that this is our attitude towards physical sciences, this is how we already proceed. Of course, there will be scientific progress which will lead to additions to the current physical science. However, the hypothesis is that no addition would be substantive enough to significantly change the face of physics. So it seems rational to preserve theory-based definitions. In the end, this is a pragmatic choice. True, there is an open-ended definition for 'physical', if physical is what is described by the ever changing physical sciences. But I do not see that as threatening physicalism. The response consists in taking the dilemma's first horn and denying its consequences: we grant that present physics is already a complete theory in the sense that new additions will not drastically change the theory.

3.3 *The relation question*

Now that we have settle on answers to the two first questions – the scope and the base question – we are getting closer to an adequate definition of physicalism. With the slogan 'everything is physical' we actually mean that *all instantiated properties* bear some ontologically important relation with

physical properties, understood, roughly, as the properties are determined by the language of the physical sciences. Now we ask what is the relation between instantiated properties in general and instantiated physical properties in particular? To respond to the relation question, we need to find the core commitments of physicalism, a *minimal physicalism* from which all versions of physicalism proceed. The dominant view among philosophers of mind is that *psychophysical supervenience* captures the most basic sense in which everything is claimed to be physical: everything is physical if and only if all properties supervene on physical properties.

3.3.1 Supervenience Physicalism

Supervenience is a dependence relation between low-level properties and high-level properties. To have an intuitive grasp of this relation, it is worth to look at how supervenience relations obtain beyond the mind and body interaction. Let us think of the global properties of a picture and the pixels that compose the picture. A picture that shows, say, the aurora borealis is composed of pixels, small dots arranged in a certain manner, so that when we stare at it from a certain distance, we see the aurora borealis. The image we see – many colors spreading through the sky – is the global property (high-level properties) of the picture, whereas the pixels are its base properties (low-level properties). Any change in the global properties (image) of the picture requires a change within the pixels of the picture, and not the other way around. The global properties supervene upon the pixels on the picture and they stand in an asymmetric relation: the former depends on the latter, but the latter does not depend on the former. In its slogan form: *there cannot be an A-difference without a B-difference*. Where A stands for the supervenient properties and B for the base-level properties. A copy of a painting will be identical to the original painting only if its lower-level properties are identical.⁴ If I am able to reproduce stroke by stroke, molecule by molecule one of Kandinsky's Compositions, that picture will be identical to the original. It is sometimes said that aesthetic properties are also supervenient properties. The arrangement of the dark and clear spots on the canvas is what makes the painting beautiful. The same relation is ascribed by moral naturalists to moral properties. Indeed, the notion of supervenience was first introduced in the context of the metaethical debate to explain a sort of normative naturalism, which argues that the normative properties supervene on the natural properties.

4 'Identical' in use here is in the sense of indiscernible instead of numerically different.

Now, in the philosophy of mind context, it is also claimed that physical properties and mental properties stand in a supervenience relation: the global properties (high-level) are mental properties that supervene on physical (low-level) properties. Supervenience physicalism is ‘the claim that if you duplicate our world in all physical respects and stop right there, you duplicate in all respects.’ (Jackson, 1998, p. 12) Following Jackson’s formulation of supervenience physicalism (1998):

(SP) Any world which is a *minimal* physical duplicate of *our* world is a duplicate *simpliciter* of our world.

The restriction of the supervenience thesis to *our* world (i.e. the actual world) is required because physicalism is a contingent thesis. Our world is physically determined, but things might have turned out differently: Cartesian worlds (worlds with non-physical properties-ghosts, spirits etc.) are not impossible. So the claim is that, at least in our world, physicalism is true: given the restriction to actuality, a physical duplicate of the *actual world* is necessarily a duplicate *simpliciter* of our world. This idea is also formulated by Lewis (2003, p. 88): ‘But we materialists usually think that materialism is a contingent truth. We grant that there are spooky possible worlds where materialism is false, but we insist that our actual world isn’t one of them.’ Once we restrict supervenience physicalism to actuality, we can say that the physical metaphysically necessitates the mental.⁵

There are reasons to believe that (SP) is the proper formulation of minimal physicalism, that is, all kinds of physicalism are committed to (SP). Supervenience physicalism defines the most basic physicalist position. To see this, we shall compare supervenience physicalism with two other positions which may be taken to be expressions of physicalism: token and type identity theories. Later we will consider some objections to supervenience physicalism. For now, however, what we want to ask is whether we can capture the intuitive idea of physicalism (that everything is physical) in terms of either of these two alternative theories. At this moment it is important to have in mind that supervenience physicalism, as the minimal requirement of any physicalist theory, is somewhat neutral regarding the mind-body theory in use, meaning that it is compatible with a couple of incompatible theories such as identity theory, emergentism, eliminativism etc.

5 The minimality requirement is introduced to prevent the duplicate of non-physical events such as miracles. It is not enough to consider a physical duplicate of the actual world because we risk duplicating a world physically like our phenomenally different. We want to duplicate only minimal physical aspects of the world.

3.3.2 Identity theories

According to the identity theory, our mental states are identical to our brain states, so believing that Lübeck is in Germany, that apples are red, or desiring apples are all brain states. Psychophysical identification is inspired by scientific identifications such as the discovery that water is H₂O. The idea is that there is one phenomenon described in two different ways. The identification is established in virtue of the transitivity of identity between the phenomenon in question, its causal role, and the occupant of that role: heat is whatever occupies a certain causal role R; molecular motion is the occupant of causal role R; so, by transitivity of identity, heat is molecular motion. *Mutatis Mutandis* for mental states and brain states: pain is the occupant of the causal role R, the occupant of causal role R is brain state B, so pain is the brain state B.

Type-token distinction: There are two kinds of psychophysical identification based on two ways of classifying individual things: consider a book and its copies. We can say we have read the same book by Thomas Mann, ‘Death in Venice’, although we have read different copies of the same title. ‘Death in Venice’ is the book-type and its copies are the tokens of the book-type. Or consider the question: How many letters are in the word ‘apple’? We can count the *tokens* of *types* of letters contained in the word: a, p, p, l, e (five letter-tokens) or we can count the *types* of letters: a, p, l, e (four letter-types). Tokens are occurrences of a certain type. In psychophysical identifications we may identify (i) states with tokens of physical states or (ii) types of such states. This distinction applied to mental states yields two ways in which states can be conceived: one can follow Davidson and conceive of states as concrete *events* (particulars/occurrences)⁶. Considering mental causes as events will generate token physicalism whilst considering mental causes as properties will result in type physicalism. Let us consider:

Token identity theory: For every actual particular (object, event or process) x, there is some physical particular y such that $x = y$.

The identification in token theory is between events (actual particulars) rather than properties. The main issue with token identity theory is that it is consistent with property dualism, thus not strong enough to be a proper physicalist thesis.

⁶ Events are roughly things that happen, or an occurrence of a process, such as births and deaths, thunder and lightning etc.

Token identity theory makes a claim about actual items only: it establishes no modal relation between the mental and the physical. So the whole truth is in the scope of an actuality operator (not just particulars). The token identity theory allows the possibility of a duplicate of our world with no mental or phenomenal properties instantiated, so token identity thesis could be true even when supervenience fails. The fact that two particular events actually have distinct mental properties does not rule out their physical identity in close-by possible worlds. But supervenience does rule out such close-by possibilities. Thus token physicalism does not entail supervenience. That token physicalism does not entail supervenience and that it is consistent with property dualism makes it an unsuitable candidate for expressing a physicalist theory.

Supervenience physicalism also does not entail token identity theory. Token identity theory claims that for every particular, there is some physical particular to which it is identical. And we have already seen certain problems that arise when we take the domain of physicalism to be particulars instead of properties. According to the token identity theory, there must be a particular physical object to which, say, a complex object like the Goethe University is identical. But it is very difficult to say what particular physical object is identical to the Goethe University, perhaps there is none. Supervenience, by itself, does not impose this sort of reductive requirement, rather it only claims that the university is dependent on or determined by physical properties. So, supervenience physicalism does not imply token physicalism.

The type identity theory, on the other hand, refers to identity not between events or processes (considered as particulars) but between types of events or processes.

Type identity theory: For every actually instantiated mental property F, there is some physical property G such that $F=G$. (Stoljar SEP)⁷

This formulation is evidently not consistent with property dualism. Then, contrary to the token identity theory, it implies the supervenience thesis: if every property instantiated in the actual world is identical with a physical property, then a world physically identical to our world will be identical to it *simpliciter*. Type physicalism entails supervenience but not the other way around. For supervenience is a contingent thesis; so it is consistent with the

⁷ 'Actually' indicates that the type identity theory, like physicalism in general, is meant to be a contingent thesis about our world.

(far-away) possibility of disembodied mental properties (Cartesian worlds), whereas type identity physicalism is not. Presumably, the world could have turned out differently such that there could be disembodied souls wondering around our universe. But the sort of psychophysical identity involved in type identity theories is of the necessary kind. For this reason, type identity theory is inconsistent with the possibility of disembodiment. Hence supervenience does not entail identity theory.

One problem with the type identity theory is that it does not cover cases in which very different physical states may occupy the same causal role characteristic for a certain mental state like pain. This is the *multiple realizability objection*. Consider pain in a horse. It is plausible (or so we may assume for the sake of argument) that the occupant of the pain-role in a horse is different from ours, given the significant difference between our organisms. Let us call the occupant of the causal role of pain in horses ‘stimulation of d-fibers’ whereas the occupant of the causal role of pain in humans is stimulation of c-fibers. If pain is stimulation of c-fibers and also stimulation of d-fibers, then (by reflexivity and transitivity) stimulation of c-fibers is stimulation of d-fibers, and that is false. Different types of state might occupy the pain-role in different organisms. Type identity cannot allow for that role to be multiply realized.⁸ Against this objection the type identity theorist could turn to the token identity theory, but we have seen that this version of identity theory does not yield an acceptable physicalist position. A solution to the multiple realizability objection is not to reject physicalism altogether, but rather to reject the identity theory.⁹ We then obtain theories that are subsumed under the heading of ‘non-reductive physicalism’ like functionalist and emergentist theories. The functionalist approach individuates mental phenomena according to their causal roles.

3.3.3 Supervenience as the standard relation

Supervenience is admittedly a weak thesis. Kim (1998) goes even further and says that besides weak, it does not provide a satisfactory account of the mind-body problem. It merely states a pattern of property covariation between

- 8 There is a better way to respond to the multiple realizability objection from the perspective of a type identity theory. One may want to finely-grained the *relata* in the identity relation. We should thus identify pain in humans with stimulation of c-fibers and pain-in-horses with stimulation of d-fibers instead of plain pain, and so on for other cases.
- 9 Kripke (1980) argues that any identity statement (where a mental state is said to be identical to a physical state) is false based solely on the structure of the statements. His argument shows that any statement that strictly equates mental events with physical events will be false. However, Kripke’s argument against identity theories will not be discussed here.

the mental and the physical and points to the existence of a dependence relation, being silent on matters like the nature of that relation, that is, it fails in explaining what the relation is. What favors for this deflationist account of supervenience is that supervenience itself seems to be a commitment of different and conflicting physicalist positions. Type identity theory implies supervenience, as well as realization physicalism – the view that the mental is physically realized – and epiphenomenalism – if two individuals differ in some mental respect, they have to differ in some physical respect – among other theories are all consistent with psychophysical supervenience. So, the supervenience thesis endorses views that have physical processes at the bottom of mental processes and rule out views that “allow the mental world to float freely, unconstrained by the physical domain” (Kim, 1998, p. 15). This certainly is a core commitment of physicalism. The supervenience thesis looks like the right candidate for *minimal physicalism*. Surely one can strengthen supervenience to obtain stronger physicalist theories. But the very neutrality of supervenience is what qualifies it as the key ingredient in a minimal physicalist answer the relation question.

4 Conclusive Remarks

The task of this work was to provide a more precise definition of physicalism. With aid of a background question suggested by Daniel Stoljar (2010) that structured the discussion about the definition of physicalism: ‘what does it mean to claim that everything is physical?’, the various attempts to respond consisted in formulating the minimal commitments a theory must meet to be a physicalist theory. To evaluate these attempts one may distinguish three questions: The base question, the scope question and the relation question. The question about the scope of the quantifier ‘everything’ in the preliminary slogan was how to restrict the quantification of the physical. As mentioned in the third section, it seems that the best option is to restrict the domain of the quantifier to instantiated properties. This way one avoids the problem of accidentally include immaterial beings, as well as immaterial properties, as part of the physical realm. With respect to the base question of what is a physical property, we have concluded that although the theory view presents some important shortcomings, it is still the most promising way to define the domain of the physical. As mentioned in section 3.2, there is no direct argument against Hempel’s dilemma, rather one may explain the preference for a theory-based conception of the physical as a pragmatic choice based on the intuition that it is rational to believe that present-day physics is already complete. Of course scientific progress is to be expected, but it is rational to expect that no future additions to the theory would

substantially change the face of physics. At last, regarding the relation question we found that supervenience physicalism is the most adequate account for physicalism. Supervenience physicalism is a dependence relation which claims that *any world which is a minimal physical duplicate of our world is a duplicate simpliciter of our world*. The fact that the supervenience is a weak thesis is considered a problem by Kim, it is its neutrality which makes supervenience the best candidate to meet the minimal requirements for physicalism. In the end we physicalism can be put as follows: Physicalism is true if and only if the instantiated properties in this world supervene upon the properties that are expressed by a predicate of a true physical theory.¹⁰

References

- ALTER, T., NAGASAWA, Y. "Consciousness in the Physical World: Perspectives in Russellian Monism". Oxford University Press, 2015.
- CRANE, T. "Elements of Mind". Oxford University Press, 2001.
- CRANE, T., MELLOR, H. "There is no question of physicalism". *Mind*, Vol. 99, Nr. 394, pp. 185-206, 1990.
- HEMPEL, C. G. "Comments on Goodman's 'Ways of Worldmaking'". *Synthese*, 45, pp. 193-199, 1980.
- JACKSON, F. "From Metaphysics to Ethics: A defence of conceptual analysis". Oxford University Press, 1998.
- KIM, J. "Mind in a Physical World: An Essay on the Mind-Body Problem and the Mental Causation". MIT Press: A Bradford Book, 1998.
- KRIPKE, S. A. "Naming and Necessity". Harvard University Press, 1980.
- LEWIS, D. "An argument for the identity theory". *Journal of Philosophy*, Vol. 63, Nr. 1, pp. 17-25, 1966.
- _____. "Mad pain and martian pain". *Readings in the Philosophy of Psychology*, Ed. N. Block, Vol. 1, pp. 216-222, 1980.
- _____. "New work for a theory of universals". *Australasian Journal of Philosophy*, 61, pp. 343-377, December 1983.
- _____. "What experience teaches". (pp. 29-57). In: W. G. Lycan (ed.). *Mind and Cognition*. Blackwell, 1990. (Reprinted in: P. Ludlow, D. Stoljar, Y. Nagasawa (eds.). *There's Something About Mary*, 2003).
- _____. "Reduction of mind". (pp. 412-431). In: S. Guttenplan (ed.). *Companion to the Philosophy of Mind*. Blackwell, 1994.
- McLAUGHLIN, B., BENNETT, K. "Supervenience". In: *The Stanford Encyclopedia of Philosophy*. 2014.

10 I wish to thank the anonymous referee for their valuable comments and suggestions.

- PAPINEAU, D. "Thinking About Consciousness". Oxford: Oxford University, 2002.
- STOLJAR, D. "Two conceptions of the physical". *Philosophy and Phenomenological Research*, 62, pp. 253-281, 2001.
- _____. "Physicalism". In: E. N. Zalta (ed.). *The Stanford Encyclopedia of Philosophy*. 2009.
- _____. "Physicalism (New Problems of Philosophy)". Routledge, 2010.
- VAN INWAGEN, P. "Material Beings". Cornell University Press, 1990.

