Organon F 24 (4) 2017: 487-510

Hempel's Dilemma and Research Programmes: Why Adding Stances Is Not a Boon

DUŠKO PRELEVIĆ[1]

ABSTRACT: Hempel's Dilemma is intended to force physicalists to make an unfavourable choice between the current physics and a future physical theory. The problem with the first horn of the dilemma is related to the fact that current physics is, strictly speaking, inconsistent, while the problem with its second horn is that we do not know how a future, completed physical theory will look like. In this paper, the two strategies of avoiding the dilemma are compared and assessed: the attitudinal approach, according to which physicalism is a stance or an attitude, and Lakatosian approach, according to which physicalism is best understood as a research programme. It is argued that the latter approach ought to be preferred over the former approach because, among other things, it better explains how some physicalists and their opponents sometimes switch the sides, as well as why different physicalists undertake different activities within a given time interval.

KEYWORDS: Attitudinal approach – Hempel's Dilemma – paradigm – physicalism – research programme – stance.

[1] Received: 31 January 2017 / Accepted: 5 August 2017

✉ Duško Prelević
Department of Philosophy
Faculty of Philosophy, University of Belgrade
Čika Ljubina 18-20, 11000 Belgrade, Serbia
e-mail: dusko.prelevic@f.bg.ac.rs; dprelevic@yahoo.com

1. Physicalism and Hempel's Dilemma

As one popular survey reports (see Chalmers & Bourget 2014, 476), most philosophers today are physicalists, that is, they think that everything is physical. They also believe that physics can explain the nature of the universe or at least that the fundamental level of reality is the subject-matter of physics. Physicalism can be spelled out in many ways, and, as pointed out by Robert Kirk, a less committing way is to say that the language of physics is (at least in principle) capable of describing all the facts about the universe, while the language of any other science is at best a re-description of the same reality.[2]

This means that even if one allows for non-physical facts and non-physical properties, these facts and properties, according to physicalists, do not belong to the fundamental level of reality: At best, non-physical properties might supervene on physical properties, which means that once all physical facts (plus the laws of physics) are fixed, everything else will be settled as well. If so, then even if some other scientific discipline (e.g., biology, psychology, economy, etc.) uses a vocabulary different from the one used in physics, all these disciplines would tell us something about one and the same reality: the physical reality.

It is likely that physicalism is a background for many/most scientists today. In physics, the debates over the correct interpretation of quantum mechanics and the validity of the causal closure of the physical, according to which all physical effects are fully determined by prior physical occurrences (and the laws of physics), might serve as an illustration. Although physics and physicalism are not the same,[3] many contemporary physicists are physicalists. For example, physicists typically endorse the principle of causal closure. However, there is more than one interpretation of quantum mechanics, one of which is the so-called Wigner's hypothesis, according to which consciousness might cause the wave function collapse. Such a hypothesis contravenes the principle of causal closure and goes in favor of

---

2 This is what Kirk calls "minimal physicalism"; see Kirk (2006) for more details.

3 That is because physicalism goes beyond physics by telling us that the fundamental level of reality can be fully described and explained by physics.

a dualistic ontology. Now, as Chalmers has pointed out,[4] physicists typically endorse the principle of causal closure (and, therefore, they reject Wigner's hypothesis) because they are physicalists, while, on the other hand, physicalists (for example, David Papineau when defending the causal argument for physicalism)[5] typically say that the principle of causal closure ought to be accepted because physicists typically endorse it.

In that respect, physicalism discourages work on certain theories (e.g. the work on the dualistic interpretation of quantum mechanics), while it encourages work on some other theories (e.g. the work on interpretations of quantum mechanics which aim to support physicalism).

Physicalism has also inspired establishing new disciplines. For example, Patricia Churchland's book *Neurophilosophy* (see Churchland 1986) had a great impact on establishing neuroethics (and perhaps later on neuroeconomics, neuroaesthetics, and so on), bringing together philosophers and scientists from various fields in order to discuss new problems. Also, many neuroscientists direct their research toward discovering neural mechanisms of yet unexplained mental processes just because they believe that all mental processes are physical. Otherwise, they would probably stop their research or they would redirect it, for example, toward dualistic or panpsychist sorts of explanation. So it is likely that physicalism is a background for many/most neuroscientists today too.

However, physicalism has been defended and characterized in more than one way. As is well known, physicalists respond to the main anti-physicalist arguments (such as the zombie argument, the knowledge argument, and so on) in different ways, and sometimes they even dispute among themselves over which responses are the most satisfactory ones. Physicalists also propose various positive arguments and accounts in order to justify their view.[6]

---

[4] See his talk (based on his collaboration with Kelvin McQueen) "Consciousness and the collapse of the wave function" (2014) which is available at: https://www.youtube.com/watch?v=DIBT6E2GtjA

[5] The causal argument runs as follows (cf. Papineau 2001, 9): All physical effects are fully determined by laws and prior physical occurrences; all mental occurrences have physical effects; the physical effects of mental causes are not fully overdetermined; therefore, mental occurrences must be identical with physical occurrences.

[6] Some of them will be sketched briefly in section 2.

Last but not least, the key notion physicalists use, the notion of "physical", underwent so many changes in the history of science. For example, the concept of matter has changed in light of new scientific discoveries (see, for example, Ney 2008a, 1034), and the same happened to some other basic notions in physics, such as space, time, mass, and the like. Further, science surprised us many times by positing new properties at the fundamental level, given that different physical theories considered different properties as fundamental. Bearing this in mind, one might be skeptical about the idea that there are necessary conditions for something to be a physical object. This opens the question on how far physicalists should go in accepting the changes of the key notions they use without ceasing to be physicalists.

This creates a tension between ontological and methodological commitments that arguably any physicalist should take. Namely, the ontological commitment binds physicalists to rule out a view that non-physical entities belong to the fundamental level of reality, while the methodological commitment binds them to accept everything physics says is true. Accepting the latter commitment threats to undermine the former, and *vice versa*. This, among other things, gives rise to a well-known dilemma for physicalists, posed by Carl Hempel (see Hempel 1980, 195), which is now called "Hempel's Dilemma".

The dilemma runs as follows: Physicalists, who claim that physics alone can explain the nature of the universe, should be more accurate and say exactly which physical theory they have in mind. At first glance, it seems that they have to choose between the current physics and a future physical theory,[7] which is rather an unpleasant choice: On the one hand, current physics is incomplete, and, strictly speaking, inconsistent, since the standard model of quantum mechanics, which is powerful in describing microphysical phenomena, is indeterministic, while general relativity, which accurately describes the universe on large scales, is deterministic (see, for example, Greene 2004, 333-335, for more details). Thus, taking the first horn of the dilemma (the so-called "currentism") is not attractive because it is irrational to believe in inconsistencies and take them as capable of providing a complete explanation of the universe. On the other hand, we

[7] Appealing to an already abandoned physical theory obviously would not be an option.

do not know how a future physical theory will look. This means that taking the second horn of the dilemma faces the "inappropriate extension worry" (see Wilson 2006, 68), which is based on the conceivability of a scenario in which a future physical theory posits irreducible non-physical entities (like phenomenal consciousness) at the fundamental level.[8] Such a scenario is likely the one in which physicalism is not true. Now, if physicalists deny in advance that such a scenario will happen, it would depart from ordinary scientific practice, to which physicalists appeal, since physics is, after all, an empirical science, and therefore it is possible, at least in principle, that it can surprise us (as it did many times in the history of science) by positing new properties at the fundamental level. On the other hand, if physicalists bite the bullet and claim that they will be ready to accept even the ideal physical theory that posits phenomenal consciousness at the fundamental level, then physicalism, according to the objection, turns out trivial and empty, because in that case anything goes (see, for example, Wilson 2006; Ney 2008a, 1037). The upshot of Hempel's Dilemma is that physicalism is either false or a trivial and contentless doctrine.

Hempel's Dilemma is a usual way to approach the problem of characterizing physicalism, and it serves as a fruitful guide that can help physicalists to spell out their view in a more precise way. In that respect, the dilemma primarily deals with the meaning of the physicalist claim, that is, it primarily concerns a *meta*-question about what physicalism is and, related to this, about what all physicalists have in common (see Prelević 2017, 5 for more details). Hempel's Dilemma deals with the question of truth of the physicalist claim too, since solutions that render physicalism false would not be considered plausible. Relatedly, competing solutions can be compared with regard to what extent they are realistic accounts of worth considering phenomena that will be addressed in due course.

Three strategies of dealing with Hempel's Dilemma have been proposed by now: defending currentism, defending futurism or trying to avoid the dilemma by claiming that physicalism is not a thesis that might be trivial or empty, but something else (e.g., a stance or a research programme). The

[8] Here, dualistic interpretation of quantum mechanics (Wigner's hypothesis), mentioned in this section, might serve as an illustration.

first two strategies have been widely defended and criticized.[9] In what follows, the focus will be on the third strategy.

2. Avoiding the dilemma

In the previous section, we have seen that Hempel's Dilemma, which is aimed to force physicalists to take an unfavourable choice between current physics and a future physical theory, presupposes that physicalism is a *thesis* that might be true, false, trivial or empty. Avoiding the dilemma consists in challenging such an assumption. Here, two ways of avoiding the dilemma will be addressed: the attitudinal approach and understanding physicalism as a research programme. In what follows, these two approaches above will be compared. It will be argued that the latter approach (presented in section 2.2) fares better than the former approach (presented in section 2.1) as to how some physicalists (and their opponents) sometimes switch the sides, as well as why different physicalists undertake different activities within a given time interval. These considerations, if correct, would license a view that the latter approach (properly understood) ought to be preferred over the former approach because it is a more realistic account of worth considering phenomena that are relevant for characterizing physicalism and resolving Hempel's Dilemma thereof.

2.1. The attitudinal approach

Let us start with the attitudinal approach, according to which physicalism is best understood as a stance (or an attitude). Alyssa Ney expresses such an attitude in the following slogan: "I hereby swear to go in my ontology everywhere and only where physics leads me" (Ney 2008, 11).

In philosophy of science, the notion of stance has been famously introduced by Bas van Fraassen (2002). He has done so because, among other things, he wanted to resolve the problem of justifying empiricism. Namely, if empiricism is the claim that experience is the one and only source of factual information, then there is a problem of how to justify the empiricist

[9] For arguments against currentism, see, e.g., Wilson (2006, 64-66); and Prelević (2017); for the disputes among futurists, see Wilson (2006); and Dowell (2006); for critiques of futurism, see, for example, Ney (2008a); and Prelević (2017).

claim itself, since such a claim cannot be supported by experience. Hence, adopting empiricism as a thesis would be self-defeating. For that reason, van Fraassen understands empiricism as a stance that commits its adherents to act in a certain way and, at the same time, being aware that adopting such a stance is not justified by providing an algorithm or something of that sort. By the same token, van Fraassen thinks that problems like Hempel's Dilemma can be avoided once we understand physicalism as a stance, and not as a thesis (see van Fraassen 2002, 49 for more details)

It has already been pointed out that van Fraassen's account does not match well with the standard classifications in the history of philosophy, since it allows us to count philosophers like Descartes, Leibniz and Chalmers – who arguably tried (or could have tried) to reconcile their ontologies with their preferred physical theories[10] – as physicalists, which is rather implausible (see Prelević 2017). Perhaps one way of dealing with this problem would be to include some metaphysical commitments in characterizing physicalism along the lines of James Ladyman's defence of what he calls the "scientistic stance" (see Ladyman 2011). Although this would depart from what van Fraassen originally had in mind – after all, van Fraassen's empirical stance was purported to be anti-metaphysical – it would still be in line with the view that physicalism is a stance rather than something else.

By having or taking a stance, van Fraassen means "having or adopting a cluster of attitudes, including a number of propositional attitudes which will generally include some beliefs" (van Fraassen 2004, 175; see also van Fraassen 2002, 47-48). Here, the main point is that stances are not theses (although they typically contain them) as well as that stances permit someone to endorse a belief *without* pretensions to claim that such a belief is

[10] For example, Descartes considered conservation laws (the "quantity of motion") a nondirectional (scalar) quantity (mass times speed; see, for example, Descartes 1985, 83–84; see also Woolhouse 1985; and Papineau 2001, 14-15), which made it possible for him to claim that mind can alter the direction of body's motion leaving the conservation laws intact. Leibniz famously criticised him on these matters (see, for example, Leibniz 1997), but given that he, like Descartes, endorsed the causal closure of physics, he proposed the doctrine of pre-established harmony instead of interactionist dualism. On the other hand, a dualistic interpretation of quantum mechanics, to which Chalmers sometimes appeal, contravenes the causal closure of the physical world (see section 1).

rationally mandated (cf. Teller 2004, 161). As Paul Teller suggests, the notion of stance can be clarified by using the analogy with adopting a policy:[11] truth values are not assigned to policies, policies commit us to act in a certain way or to make certain decisions, they may be overridden by some other criteria or policies, they may be interpreted or applied in more than one way, and so on.

Given that, as stressed above, stances permit someone to endorse a belief *without* pretensions to claim that such a belief is *rationally mandated*, van Fraassen's approach is confronted with the problem of "stance voluntarism", which refers to "the thesis that one can intentionally acquire or sustain a stance in the absence of any epistemic reasons for that stance" (Baumann 2011, 29). Such a thesis implies that contrary stances are rationally permissible (see Chakravartty 2011).[12]

In that respect, it is not surprising that van Fraassen's conception of stance is often compared with Kuhnian view of paradigms, since Kuhn (1962) famously argued that, during scientific revolutions, "paradigm shifts" occur in a way in which replaced and newly established paradigms are incommensurable. Paradigms are, simply put, frameworks within which scientific communities work. In his "Second Thoughts on Paradigms" (see Kuhn 1974), Kuhn understood paradigms as disciplinary matrices that consist in "a constellation of group commitments" which, among other things, include examplars (shared examples) that suggest new puzzles, approaches to resolving them, and serve as standards that enable those who do the research within the paradigm to measure the quality of the proposed solutions (cf. Rowbottom 2011, 115)

As Darrell Rowbottom has pointed out, stances are very similar yet not identical to paradigms. According to him, stances should not be understood as paradigms writ large, since paradigms, unlike stances, include examplars. Rowbottom thinks that introducing stances should not be understood merely as spelling out a known idea in a new fashion, but as appraising it as a boon. He thinks that the distinction between stances and paradigms

---

[11] Van Fraassen agrees with him on that by telling that it clarifies the epistemological aspects of the notion (see van Fraassen 2004, 179).

[12] I will stay neutral in due course on whether van Fraassen's view of stance voluntarism leads to latent irrationality or not (this objection can be found, for example, in Baumann 2011).

enables us to explain why different scientists undertake different activities, that is, "how and why there is a measure of dissent within the boundaries of the disciplinary matrix" (Rowbottom 2011, 115). Rowbottom's solution to this problem runs as follows: "My basic idea is that a disciplinary matrix implies a *set of permissible stances*, and that the difference in stances of individual scientists explains how and why a broad range of activities occur" (Rowbottom 2011, 117). At the end of his paper, Rowbottom conjectures that van Fraassen's notion of stance may be also used to explain Kuhnian conversions in science, yet he finishes his paper without developing such an idea.

2.1.1. Physicalism and conversions

In the previous section, it was stressed that both Kuhnian view of paradigm shifts and van Fraassenian view of stance voluntarism are aimed to support the thesis that conversions in science are not rationally mandated. However, in the context of the debate over the possibility of characterizing physicalism, these accounts are hardly acceptable.[13] After all, the fact that so many arguments have been proposed for or against physicalism (and alternative views as well) suggests that a rational choice between physicalism and the alternative views can be made within a given time interval, contrary to what Kuhn's incommensurability thesis and van Fraassen's stance voluntarism presuppose.

Here, it is worth mentioning that even if some physicalists appeal to the Kuhnian view of scientific revolutions, it would still not follow that they themselves experience paradigm shifts whenever they introduce their theories. For example, eliminativists like Daniel Dennett[14] and Paul and Patricia Churchland typically claim that phenomenal consciousness will be

---

[13] As is well known, Kuhn's incommensurability thesis as such has been criticized many times (see, for example, Newton-Smith 1981 for more details). However, the main point here is that even if such a thesis can help us get a better grasp of some interesting episodes in the history of science, it would still not be of any use for our understanding of the nature of physicalism. The same holds, *mutatis mutandis*, for van Fraassen's account.

[14] As for Dennett, many times he has challenged anti-physicalist arguments, such as the zombie argument and the knowledge argument, by arguing that they are bad thought

explained away within a future physical theory in almost the same way as it happened with some other theoretical terms in science, such as phlogiston, luminiferous aether, and the like (see, e.g., Churchland 1996). Given that Kuhn interpreted episodes like these as the cases of paradigm shifts, a natural guess is that at least some eliminativists think (or could have thought) that a corresponding paradigm shift will dissolve phenomenal consciousness too. Yet this would at best show that philosophers who appeal to Kuhnian insights on how revolutions in science occur do that in order to provide a *rational* support for their view rather than because of experiencing a paradigm shift. Here the structure of their arguments would be almost the same as of those used by some identity theorists or analytic functionalists who appeal to theoretical identifications established in natural sciences (such as that water is H_2O, that genes are DNA, and the like) in order to justify the claim that consciousness is a brain process, and the like. Such optimism is far from not being rationally mandated[15] at least from the perspective of philosophers who share it and in the absence of counter-arguments. So it is likely that physicalistic views like eliminativism are not incommensurable with anti-physicalistic views.

In addition, let us recall a few representative cases of conversion in philosophy of mind. One such example is Frank Jackson's conversion, whose version of the knowledge argument is widely discussed in contemporary philosophy of mind.[16] Here is what Jackson says on this issue in one interview:[17]

---

experiments (he calls them "intuition pumps"; see Dennett 1991, 282 for more details). This also reveals that his defence of physicalism is rationally mandated.

[15] Here, as well as in cases below, I just present briefly some well-known arguments of various physicalists and their opponents in order to shed a better light on the nature of their debates and enterprises. I do not commit myself to holding their arguments valid.

[16] Jackson's knowledge argument is intended to show that knowledge of completed physics (chemistry and neurophysiology) does not enable us to know everything about the world, since one who knows everything about a completed science of colour vision could still be, for instance, ignorant of what is it like to see red.

[17] See the interview: "Frank Jackson, Later Day Physicalist" (2011), which is available at: http://www.philosophersmag.com/index.php/tpm-mag-articles/14-interviews/22-frank-jackson-latter-day-physicalist.

> In 'Epiphenomenal Qualia' I explain why it's not such a disaster being an epiphenomenalist, but I came to think of this as a triumph of philosophical ingenuity over common sense. This is what someone who's done a good philosophy degree can somehow make seem all right, but if you look at it in a more commonsensical way it's actually pretty implausible. So the epiphenomenal stuff was just very hard to believe.

However, Jackson himself changed his mind definitely *after* realizing that a representationalist theory of consciousness (a version of intentionalism that goes in favor of physicalism) is a viable doctrine. Actually, he detected the key intuition behind the knowledge argument and tried to show how such an intuition conflicts with an attractive view of the nature of phenomenal concepts that can be defended on independent grounds (see Jackson 2007 for more details). He has also provided some reasons why, for example, he believes that alternative responses to the knowledge argument, such as the "missing-concept reply", are not convincing.[18] So it is likely that Jackson's conversion to physicalism was rationally mandated, contrary to what van Fraassen and Rowbottom would say in similar cases.

It is also interesting to notice that some main figures in the debate over the validity of the zombie argument have completely changed their views on these matters.[19] On the one hand, Robert Kirk, who introduced the zombie argument in 1970s (see Kirk 1974), has changed sides and started to argue that zombies are not just impossible, but inconceivable as well (see, for example, Kirk 2007), while on the other hand, David Chalmers, whose version of the zombie argument against physicalism has been in focus for

---

[18] This reply consists in claiming that inside her black-and-white room (in Jackson's thought experiment) Mary is unable to acquire phenomenal concepts, which does not entail by itself that phenomenal truths are not a priori deducible from corresponding totality of micro-physical truths (plus the laws of physics).

[19] The zombie argument, roughly, starts with the premise that zombies – our physical duplicates who, unlike us, do not have phenomenal consciousness – are conceivable, continues with the principle that conceivability entails metaphysical possibility, ending up with the conclusion that metaphysical possibility of zombies undermines physicalism, in one way or another.

last twenty years or so (see, for example, Chalmers 2010), originally had thought that zombies are impossible, albeit conceivable.

However, these conversions can hardly be regarded as the cases of Kuhnian paradigm shifts. Robert Kirk has tried to show that the zombie scenario implies a sort of epiphenomenalism that involves a contradiction (cf. Kirk 2007). As for Chalmers, here is what he says in a recently held interview about his conversion:[20]

> I wanted to write a big-picture treatment of consciousness in philosophy and science and at the same time put forward a positive theory of consciousness. In my first couple of years at Indiana I wrote two long articles (still unpublished except on the web) pursuing the connection between consciousness and the way we talk about consciousness, but I also gradually got drawn into issues about materialism and dualism. I had come to graduate school thinking of myself as a materialist (albeit one who was very impressed by the problem of consciousness), but I gradually realized that commitments I already had meant that materialism couldn't work, and I should be some sort of dualist or perhaps panpsychist.

The passage above suggests that Chalmers has changed his view after a more careful reflection on the commitments he already had accepted, and realizing that those commitments are incompatible with physicalism (materialism). A natural guess is that he realized that his views on the relation between modality and apriority, semantics of phenomenal and micro-physical concepts, quantum mechanics, and the like, do not match well with physicalism.

These representative cases of conversion suggest that it is more likely that they are rationally mandated. They neither justify Kuhnian view of paradigm shifts, nor van Fraassenian stance voluntarism, which is considered to be a hallmark of the attitudinal approach.

---

[20] See the interview: "What Is It Like to Be a Philosopher?" (2016), which is available at: http://www.whatisitliketobeaphilosopher.com/#/david-chalmers/.

2.1.2. Varieties of physicalism

Now, let us check whether van Fraassen's attitudinal approach can explain why different physicalists undertake different activities. In order to show that this is not the case, let us start with noticing that the history of physicalism is to a great extent parallel with the history of analytic philosophy, primarily with respect to the question about how philosophers see the relationship between philosophy, science and metaphysics. Namely, when Otto Neurath coined the term "physicalism" in 1930s (see Neurath 1983), he thought, like other members of the Vienna Circle who were influenced by the work of the early Wittgenstein, that there are no meaningful propositions in philosophy (in traditional metaphysics, in particular), and also that philosophy is a quite different activity from science. Generally, in the age of the "linguistic turn" (Gustav Bergmann's phrase), philosophers who endorsed physicalism in one way or another typically tried to provide a reductive analysis of the mental (for example, by means of a dispositional analyses of mental states; see, for example, Carnap 1959; Ryle 1949) or to show that there is no room for the subjective aspects of conscious experience (*qualia*) in corresponding language-games (this was the upshot of Wittgenstein's the-beetle-in-a-box thought experiment; see Wittgenstein 1958, § 295), and the like.

Quine's critiques of the main ideas defended by philosophers of the Vienna Circle[21] inspired many philosophers of that time and led them to think that philosophy and science should not be separated, and that metaphysics (modal discourse and essentialism, in particular) ought to be rejected. In view of the last fact, it is not surprising at all that the proponents of the identity theory, such as Place (1956), famously claimed that their theory "is a reasonable scientific hypothesis". They also believed that statements like "Consciousness is a process in the brain" are contingently true, and that the past successes in providing physical explanations of biological and chemical phenomena give rise to a belief that corresponding theoretical identifications in psychology are available.

---

[21] See, e.g., Quine (1951) for his famous criticism of the analytic/synthetic distinction.

However, Barcan's and Kripke's insights on identity, modality and essence became influential, and increased philosophers' interest to take metaphysics seriously. Although Kripke has famously argued against physicalism (see Kripke 1980 for more details), very soon physicalists tried to reconcile their own views with Kripke's compelling examples of necessary a posteriori statements and his explanation of modal illusions. For example, some physicalists claimed that views like "token physicalism" are even strengthened by the Kripkean insights on the necessary a posteriori statements (see, for example, McGinn 1977), while some others tried to show that terms like "pain" are not rigid designators that pick out their objects of reference through the use of essential modes of presentations (see Lewis 1983; and, more recently, Grahek 2007).[22]

In 1990s Chalmers famously amended conceivability arguments against physicalism, such as the zombie argument, in order to show that his view is compatible with the standard Kripkean cases of the necessary a posteriori. He has elaborated the key notions used in the argument, applied the epistemic version of the two-dimensional semantics, setting up his argument to the effect that the burden of proof has been shifted to physicalists.

Physicalists react to Chalmers's zombie argument in various ways. Some think that phenomenal consciousness can be explained a priori in terms of the physical, while others think that, although there is an explanatory gap between the physical and the mental, this gap still does not entail that there is an ontological gap between the physical and the mental. In other words, the latter argue that the conceivability of zombies does not entail that they are metaphysically possible. There are also physicalists who are ready to redefine physicalism in order to save the day (see, for example, Leuenberger 2008).

This very brief and incomplete outline of some representative physicalists' strategies of dealing with the zombie argument illustrates that it is, contrary to the attitudinal approach, highly unlikely that physicalists voluntarily undertake different activities due to the stances they adopt.

---

[22] In "Mad Pain and Martian Pain", Lewis constructed thought experiments purported to show that "pain" is not rigid designator. On the other hand, Nikola Grahek argued that some interesting cases in neuroscience, such as pain asymbolia, suggest that feeling pain (painfulness) and being in pain can be departed from each other.

Further, it is not unusual that physicalists dispute among themselves over which responses to the anti-physicalistic arguments, such as the zombie argument, are the best. Here, the disputes over the validity of the phenomenal concept strategy can serve as good illustrations.[23] Let us recall that this strategy consists in providing an account that would support the claim that, due to a specific nature of phenomenal concepts, physicalism can be true despite the explanatory gap. Various accounts of that sort have been proposed by now: indexical account, recognitional account, quotational account, and so on (see, for example, Alter & Walter 2006 for more details). On the other hand, setting aside the criticisms coming from anti-physicalists, the phenomenal concept strategy has been criticized by some physicalists more than once. For example, Daniel Stoljar[24] argues that the proponents of the phenomenal concept strategy at best can show that psychophysical conditionals, in which it is claimed that a complete description of the world in physical terms necessitates a complete description of the world in phenomenal terms, are not a priori synthesizable, yet they are not capable of explaining why those conditionals are not a priori.[25] The proponents of phenomenal concept strategy typically try to handle such an objection by providing examples and arguing that the psychophysical conditionals are analogous with some other conditionals that are likely not a priori (see, e.g., Diaz-Leon 2008). It is evident that such a dispute is rationally mandated, contrary to what the attitudinal approach would predict.

---

[23] Intentionalism in philosophy of mind, which is sometimes taken to support physicalism (see, for example, Cutter & Tye 2011; Grahek 2007; Klein 2007), can also serve as a good illustration here, since its proponents often dispute among themselves over which version of intentionalism better explains interesting phenomena. At the same time, there are physicalists, such as Ned Block (see, e.g., Block 1997), who reject intentionalism, typically by claiming that such a theory cannot explain some interesting phenomena (such as blindsight, and the like). This suggests that the debates over the validity of intentionalism are rationally mandated too.

[24] Another critique of the phenomenal concept strategy, posed by a physicalist, comes from Tye (2009).

[25] According to Stoljar (2005, 478), a sentence is a priori synthezisable when "a sufficiently logically acute person who possessed only the concepts required to understand *its antecedent*, is in a position to know that it is true," while a sentence is is a priori when "a sufficiently logically acute person who possessed only the concepts required to understand it, is in a position to know that it is true."

Now, one might think that these considerations turn into a sociological analysis of the physicalist debate, and that it is not clear if such an analysis helps us to address Hempel's Dilemma.[26] As a response to this worry, it should be stressed that these considerations are just partly devoted to a sociological (or a historical) analysis of the physicalist debate: They are primarily aimed to shed a better light on the *rationality* lying behind the willingness of various physicalists to undertake different activities within the same research programme.[27] I hold it is a common practice in philosophy of science to compare competing accounts (for example, Popperian, Kuhnian, Lakatosian accounts, and the like) of the nature of science and scientific rationality by taking into account representative episodes in the history of science, and evaluating to what extent those accounts are realistic in explaining them. This method has been applied outside philosophy of science as well. For example, in his *The Philosophy of Philosophy*, Timothy Williamson writes: "The primary task of the philosophy of science is to understand science, not to give scientists advice. Likewise, the primary task of the philosophy of philosophy is to understand philosophy – although I have not rigorously abstained from the latter" (Williamson 2007, ix). So, I think it is legitimate to apply the same method in assessing competing solutions to Hempel's Dilemma. This means that the considerations above are relevant for assessing competing solutions to Hempel's Dilemma, and that, as it stands, they do not go in favor of the attitudinal approach.

2.2. The Lakatosian approach

Now, let us turn to another strategy of avoiding Hempel's Dilemma, namely that of understanding physicalism as a research programme. In philosophy of science the term "research programme" was famously introduced by Imre Lakatos (see Lakatos 1978), who thought that the units of evaluation in science are not theories but research programmes, within which particular theories and models are produced. According to Lakatos, research programmes guide one's research, and they consist in the "hard-

---

[26] I would like to thank an anonymous referee for drawing my attention to this issue.

[27] The same holds, *mutatis mutandis*, for the cases of conversion, presented in section 2.1.1.

core", positive heuristic, and negative heuristic (cf. Lakatos 1978, 47). The hard-core of a research programme contains basic claims (for example, the three principles of motion in Newton's mechanics), and it is always protected by negative heuristic that redirects potential counterevidence to inessential parts of the programme (to auxiliary hypotheses, etc.). Positive heuristic suggests paths worth of being pursued, the order of investigation, ways to construct models and theories, and so on (see Lakatos 1978, 50). While negative heuristic discourages work on certain theories and models, positive heuristic encourages work on some other theories and models. Also, many competing theories and models might be produced within a single research programme (Newton's mechanics and Darwinism might serve as good examples).

Although it might seem that many philosophers do not explicitly admit of being engaged in a research programme, the fact that they typically try to amend tenaciously their arguments from critiques is a good evidence that they actually are. Of course, philosophers sometimes switch to another research programme, quite the opposite to the one they endorsed earlier (some representative cases of conversion were presented in section 2.1.1).

Recently, a view that physicalism is a research programme has been proposed independently by Guy Dove and Duško Prelević.[28] According to Dove (2016, 5), physicalism is an "ongoing interdisciplinary research programme", the core theses of which are, respectively, that current physics inspired physicalists to count certain entities as physical, and that past suc-

[28] A view that physicalism is a specific *theory* through which materialist (metaphysical) research programme expresses itself at various times is defended by Seth Crook and Carl Gillet (see Crook & Gillett 2001, § 3). However, although physicalism is usually regarded as a descendant of the materialist worldview, it is still rich enough to be understood as a separate research programme: After all, commitment to physics as fundamental science is not a necessary part of materialist metaphysics, whereas, on the other hand, physicalism is arguably incompatible with some materialist views about material substance, like those that were famously criticized by George Berkeley, and the like.

Of course, a natural guess is that Lakatos himself, had he been asked, would have said that physicalism is a research programme: after all, he understood science and ideologies in the same way.

cess in providing physical explanations of biological and chemical phenomena may serve as positive exemplars of how mental phenomena should be explained.

However, putting current physics and past exemplars into the hard-core of physicalist research programme is unsatisfactory for two reasons. First, such an account does not say too much about the nature of current physics. Here, let us recall that today there is more than one interpretation of quantum mechanics, one of which is a dualistic interpretation (see section 1) that is by no means acceptable to our day physicalists. If so, then Dove's account is too permissive. On the other hand, it seems that by positive exemplars in biology and chemistry Dove means well known cases of reductive explanations (functional reductions), to which physicalists frequently appeal in philosophy of mind, such as the explanation of why water is H_2O, why heat is the motion of molecules, why genes are DNA, and the like. Yet, in our times, explanatory pluralism in philosophy of science is a more viable doctrine, which means that physicalists need not be constrained by just one sort of explanation. For example, many phenomena in biology are explained in a non-reductive (and even in a non-causal) way, by appealing to the same level phenomena or even to the higher-level phenomena, as is the case, for instance, with the statistical explanations in theoretical population biology (see Walsh 2015 for further details), and the topological explanations that are used to explain, for example, metabolic economy, synchronicity, stability, robustness, resilience, and the like (see Kostić 2016). Furthermore, there are physicalistic views in philosophy of mind, such as the higher-order theory of consciousness (see, for example, Rosenthal 2011), which aim to provide a (second-order) representational account of consciousness, in which physics or any lower-level theory plays no role. Thus, it seems that Dove's solution is too restrictive concerning the sorts of explanations available to our day physicalists in dealing with mental phenomena.

In contrast to the solution above, Prelević (2017) understands physicalism as a research programme by putting some positive aspects of the term "physical" into the hard-core, such as the claim that a necessary condition for something to be a physical object is to be located in space and time, that is, that what physics generally deals with is, as Chalmers puts it, "structure and dynamics of the world throughout space and time" (Chalmers 1996,

36).[29] According to Prelević, this is possible because the hard-core of the physicalist research programme need not be *fully* specified: further specifications of the core claims of the physicalist research programme belong to its *positive heuristic*, in which many physical models and theories have been proposed.[30] This way, it would be possible to handle the standard problems related to *via negativa*,[31] which concern the (im)possibility of delineating physicalism from views such as the Russellian monism, which, historically speaking, were not counted as physicalistic.[32] Another virtue of the proposal just sketched is that it avoids the problems typical of Dove's solution above, since, on the one hand, it does not rely upon accepting current physics as such, while, on the other hand, it is quite compatible with explanatory pluralism.

Now, let us assess the explanatory power of the Lakatosian solution (as proposed in Prelević 2017) to Hempel's Dilemma. First, it is easy to notice that understanding physicalism as a research programme matches well with the standard classifications in the history of philosophy. Within the Lakatosian account, philosophers like Descartes, Leibniz and Chalmers might easily be classified as philosophers who develop research programmes involving core theses different from the core these of physicalist research programme.

As for the cases of conversion, mentioned in section 2.1.1., they are quite in accordance with the Lakatosian approach too. Lakatos's view of

---

[29] Physicalist research programme also includes, according to Prelević, a view that listing the furnishings of the universe is the subject-matter of physics.

[30] In that respect, one's views about the real nature of space and time (for example, whether space is three-dimensional or configurational, whether space and time are independent of each other or it is better to speak about space-time, and the like), about how many properties belong to the fundamental level, about the nature of the laws of physics, and so on, depend on a physical theory one adopts (see Chalmers 1996, 119 for more details), which is a part of the positive heuristic of physicalist research programme.

[31] Here, *via negativa* is a view that "physical" is best defined negatively, like the "non-fundamentally mental". This view is originally introduced as a version of futurism (see Montero 1999), but it can be also incorporated into the hard-core of physicalist research programme.

[32] That is because, according to Russellian monism, neither physical properties nor mental properties are counted as fundamental; see Judisch (2006) for more details.

research programmes is in many respects akin to Kuhn's view of paradigms (see, for example, Kuhn 1970, 238), but one crucial difference among them was that Lakatos rejected Kuhn's incommensurability thesis, arguing that research programmes can be compared within a given time interval (for example, by comparing to what extent they are fruitful or degenerate). In that respect, the cases of conversion above can be understood as those in which philosophers simply switched to another research programme, and got started developing arguments and accounts in light of new challenges.[33] Finally, the Lakatosian approach alone can explain why different physicalists undertake different activities. As stressed above, this account can easily explain the differences between various physicalists by positing their novelties and disputes in *positive heuristic* (see Prelević 2017, footnote 11). Bearing this in mind, there is no need for introducing stances (in van Fraassen's sense) in order to explain why different physicalists undertake different activities within one and the same research programme. Representative examples, addressed in section 2.1.2, suggest that physicalists do that not because they voluntarily adopt different stances, but due to some arguments they find convincing.

3. Conclusion

Previous considerations suggest that research programmes can be pursued without invoking van Fraassenian stances, and that there is no need for adding stances in order to explain the nature of physicalism and resolve Hempel's Dilemma thereof. Choices that physicalists make are far from not being rationally mandated, contrary to what Rowbottom's view of the role of stances would predict. On the other hand, these practices are compatible with a view that physicalism is a research programme within which different solutions are proposed and compared. In view of the last fact, it is highly unlikely that anything should prompt us to introduce van Fraassenian notion of stance in order to understand what physicalism is, what all physicalists have in common, and how to explain the differences among

[33] This is evident, for example, in the case of Chalmers, who updates his amendments of the zombie argument from time to time, taking into consideration new criticisms (see, for example, Chalmers 2010).

them. In other words, adding stances, which is a hallmark of the attitudinal approach, is not a boon.

Acknowledgments

An earlier draft of this paper was presented at the international conference EENPS 2016, which took place in Sofia (24–26 June, 2016). I would like to thank the organizers and the audience for their input. I am also very much indebted to Daniel Kostić for stimulating discussion of these matters, and the anonymous reviewer for *Organon F* for valuable and very helpful comments and suggestions provided.

This research was supported by Ministry of Education, Science and Technological Development of the Republic of Serbia (project: *Logico-epistemological foundations of science and metaphysics*, No. 179067).

References

ALTER, T. & WALTER, S. (eds.) (2006): *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*. Oxford: Oxford University Press.

BAUMANN, P. (2011): Empiricism, Stances, and the Problem of Voluntarism. *Synthese* 178, No. 1, 27-36.

BLOCK, N. (1997): On a Confusion about a Function of Consciousness. In: Block, N., Flanagan, O. & Güzeldere, G. (eds.): *The Nature of Consciousness: Philosophical Debates*. Cambridge (Mass.): MIT Press.

BOURGET, D. & CHALMERS, D. (2014): What Do Philosophers Believe? *Philosophical Studies* 170, No. 3, 465-500.

CARNAP, R. (1959): Psychology in Physical Language. In: Ayer, A. (ed.): *Logical Positivism*. New York: The Free Press, 165-198.

CHAKRAVARTTY, A. (2011): A Puzzle about Voluntarism about Rational Epistemic Stances. *Synthese* 178, No. 1, 37-48.

CHALMERS, D. (1996): *The Conscious Mind*. Oxford: Oxford University Press.

CHALMERS, D. (2010): The Two-Dimensional Argument against Materialism. In: Chalmers, D. (ed.): *The Character of Consciousness*. Oxford: Oxford University Press, 141-205.

CHURCHLAND, P. S. (1986): *Neurophilosophy*. Cambridge (Mass.): MIT Press.

CHURCHLAND, P. (1996): The Rediscovery of the Light. *The Journal of Philosophy* 93, No. 5, 211-228.

CROOK, S. & GILLETT, C. (2001): Why Physics Alone Cannot Define 'Physical': Materialism, Metaphysics, and the Formulation of Physicalism. *Canadian Journal of Philosophy* 31, No. 3, 333-359.

CUTTER, B. & TYE, M. (2011): Tracking Representationalism and the Painfulness of Pain. *Philosophical Issues* 21, No. 1, 90-109.

DENNETT, D. (1991): *Consciousness Explained*. Boston: Little, Brown and Co.

DESCARTES, R. (1985): The World or Treatise on Light. In: Cottingham, J., Stoothoff, R. & Murdoch, D. (eds.): *The Philosophical Writings of Descartes*. Vol. 1. Cambridge: Cambridge University Press, 81-98.

DIAZ-LEON, E. (2008): Defending the Phenomenal Concept Strategy. *Australasian Journal of Philosophy* 86, No. 4, 597-610.

DOVE, G. (2016): Redefining Physicalism. *Topoi*, online first. DOI 10.1007/s11245-016-9405-0.

DOWELL, J. (2006): The Physical: Empirical, Not Metaphysical. *Philosophical Studies* 131, No. 1, 25-60.

GRAHEK, N. (2007): *Feeling Pain and Being in Pain*. 2nd ed. Cambridge (Mass.): MIT Press.

GREENE, B. (2004): *The Fabric of the Cosmos*. New York: Vintage.

HEMPEL, C. (1980): Comments on Goodman's Ways of Worldmaking. *Synthese* 45, No. 2, 193-200.

JACKSON, F. (2006): The Knowledge Argument, Diaphanousness, Representationalism. In: Alter, T. & Walter, S. (eds.): *Phenomenal Concepts and Phenomenal Knowledge*. Oxford: Oxford University Press, 52-64.

JUDISCH, N. (2008): Why 'non-mental' won't work: on Hempel's dilemma and the characterization of the 'physical'. *Philosophical Studies* 140, No. 3, 299-318.

KIRK, R. (1974): Sentience and Behaviour. *Mind* 83, No. 329, 43-60.

KIRK, R. (2006): Physicalism and Strict Implication. *Synthese* 151, No. 3, 523-536.

KIRK, R. (2008): The Inconceivability of Zombies. *Philosophical Studies* 139, No. 1, 73-89.

KLEIN, C. (2007): An Imperative Theory of Pain. *Journal of Philosophy* 104, No. 10, 517-532.

KOSTIĆ, D. (2016): The Topological Realization. *Synthese*, online first. DOI 10.1007/s11229-016-1248-0.

KRIPKE, S. (1980): *Naming and Necessity*. Cambridge (Mass.): Harvard University Press.

KUHN, T. (1962): *The Structure of Scientific Revolutions*. Chicago: The University of Chicago Press.

KUHN, T. (1970): Reflections on my Critics. In: Lakatos, I. & Musgrave, A. (eds.): *Criticism and the Growth of Knowledge*. Cambridge: Cambridge University Press, 231-278.

KUHN, T. (1974): Second Thoughts on Paradigms. In: Suppe, F. (ed.): *The Structure of Scientific Theories*. Urbana: Urbana Illinois Press, 459-482.

LADYMAN, J. (2011): The Scientistic Stance: the Empirical and Materialist Stances Reconciled. *Synthese* 178, No. 1, 87-98.

LAKATOS, I. (1978): Falsification and the Methodology of Scientific Research Programmes. In: Worrall, J. & Currie, G. (eds.): *The Methodology of Scientific Research Programmes*: Philosophical Papers. Vol. I. Cambridge: Cambridge University Press, 8-102.

LEIBNIZ, G. W. (1997): [First] explanation of the New system of the communication between substances, in Reply to what was said of it in the Journal for 12 September 1695. In: Woolhouse, R. & Francks, R. (eds.): *Leibniz's 'New System' and Associated Contemporary Texts*. Oxford: Clarendon Press, 47-52.

LEUENBERGER, S. (2008): Ceteris Absentibus Physicalism. In: Zimmerman, D. W. (ed.): *Oxford Studies in Metaphysics*. Oxford: Oxford University Press, 4-145.

LEWIS, D. (1983): Mad Pain and Martian Pain. In: *Philosophical Papers*. Vol 1. Oxford: Oxford University Press, 122-130.

MCGINN, C. (1977): Anomalous Monism and Kripke's Cartesian Intuitions. *Analysis* 37, No. 2, 78-80.

MONTERO, B. (1999): The Body Problem. *Noûs* 33, No. 2, 183-200.

NEURATH, O. (1983): Physicalism: The Philosophy of the Viennese Circle. In: Cohen, R. & Neurath, M. (eds.): *Otto Neurath: Philosophical Papers 1913-1946*. Dordrecht: D. Riedel Publishing Company, 48-51.

NEWTON-SMITH, W. (1981): *The Rationality of Science*. London: Routledge & Kegan Paul, Ltd.

NEY, A. (2008a): Defining Physicalism. *Philosophy Compass* 3, No. 5, 1033-1048.

NEY, A. (2008b): Physicalism as an Attitude. *Philosophical Studies* 138, No. 1, 1-15.

PAPINEAU, D. (2001): The Raise of Physicalism. In: Gillett, C & Loewer, B. (eds.): *Physicalism and Its Discontents*. Cambridge: Cambridge University Press, 3-36.

PLACE, U. T. (1956): Is Consciousness a Drain Process? *British Journal of Psychology* 47, No. 1, 44-50.

PRELEVIĆ, D. (2017): Physicalism as a Research Programme. *Grazer Philosophische Studien*, online first. DOI 10.1163/18756735-000023.

QUINE, W. V. O. (1951): Two Dogmas of Empiricism. *The Philosophical Review* 60, No. 1, 20-43.

RYLE, G. (1949): *The Concept of Mind*. New York: Barnes and Noble.

ROSENTHAL, D. (2011): Exaggerated reports: reply to Block. *Analysis* 71, No. 3, 431-437.

ROWBOTTOM, D. (2011): Stances and Paradigms: A Reflection. *Synthese* 178, No. 1, 111-119.

SMART, J. J. (1959): Sensations and Brain Processes. *The Philosophical Review* 68, No. 2, 141-156.

STOLJAR, D. (2005): Physicalism and Phenomenal Concepts. *Mind & Language* 20, No. 2, 469-494.

TELLER, P. (2004): Discussion: What Is a Stance? *Philosophical Studies* 121, No. 2, 159-170.

TYE, M. (2009): *Consciousness Revisited: Materialism without Phenomenal Concepts*. Cambridge (Mass.): MIT Press.

VAN FRAASSEN, B. (2002): *The Empirical Stance*. New Haven & London: Yale University Press.

VAN FRAASSEN, B. (2004): Replies to the Discussion on The Empirical Stance. *Philosophical Studies* 121, No. 2, 171-192.

WALSH, D. M. (2015): Variance, Invariance and Statistical Explanation. *Erkenntnis* 80, Supplement 3, 469-489.

WILLIAMSON, T. (2007): *The Philosophy of Philosophy*. Oxford: Blackwell Publishing.

WILSON, J. (2006): On Characterizing the Physical. *Philosophical Studies* 131, No. 1, 61-99.

WITTGENSTEIN, L. (1958): *Philosophical Investigations*. Oxford: Basil Blackwell.

WOOLHOUSE, R. (1985): Leibniz's Reaction to Cartesian Interaction. *Proceedings of the Aristotelian Society* 86, 69-82.