9

On Characterizing Metaphysical Naturalism

Lok-Chi Chan

1. Introduction

Despite the fact that metaphysical naturalism (henceforth MN) and physicalism are widely discussed topics in, or even the metaphysical foundations of, many philosophical areas—e.g. the philosophy of mind, contemporary phenomenology, the philosophy of religion, metaethics, and the philosophy of science—there are several distinct and conflicting definitions of the two doctrines. The view that everything in the actual world is natural or physical requires the distinction between the metaphysically natural and non-natural, or the distinction between the physical and non-physical. It is commonly assumed (by both naturalists and non-naturalists) that tables, trees, and electrons are natural and physical and that disembodied minds, God, ghosts, and magic are non-natural and non-physical. However, there is disagreement about the explanation for this common assumption.[1]

Of the various proposed definitions of MN and physicalism, the *disciplinary characterization* (henceforth DC)[2] is the most popular. According to this characterization, the natural is defined with reference to whatever natural sciences posit; and the physical is defined with reference to whatever physical theories posit. This chapter is an attempt to offer a novel version of DC. With this in mind, it is important to note that the interest of this chapter is not whether or not MN (or physicalism) is true.

[1] I should note that metaphysical naturalism is not the same as *methodological* naturalism. Methodological naturalism is a doctrine according to which philosophy should be seen as an extension of science and should use similar methods. This chapter is concerned with *metaphysical* naturalism and considers it to be an independent doctrine. After all, there is a standard division between MN and methodological naturalism (see Devitt 1998, p. 46; De Caro & Macarthur 2010, p. 4; Papineau 2014, p. 116).

[2] This is Copp's (2012, p. 28) term.

In what follows I first characterize DC as it has been understood and the main objections to it. I then examine a seminal version of DC that I call the *similarity approach* (SA), which is defended by Lewis (1983), Ravenscroft (1997), Braddon-Mitchell and Jackson (2007), and McPherson (2015). The approach is to allow an extension of current science, but that this extension be constrained by an adequate similarity to such science. I argue that while SA can avoid most of the main objections to DC, it has a fatal shortcoming. I then develop a novel version of DC using the strategy of SA that allows the extension of current science. I call this the *historical paths approach* (HPA). The idea is, very roughly, that MN can be defined with reference to the historical ideas that current scientific theories descend from. HPA can, I argue, avoid the fatal shortcoming of other implementations of SA. Finally, I show that a definition of MN developed from HPA is a useful definition and can provide a useful framework for the naturalization of the philosophy of mind and phenomenology (if possible).

2. The Disciplinary Characterization (DC)

Roughly, the DCs of MN and physicalism are:

DC (MN): Every entity or property instantiated in the actual world is natural in the sense of being a posited entity or property of natural sciences or being exclusively constituted by those posited entities or properties.

DC (physicalism): Every entity or property instantiated in the actual world is physical in the sense of being a posited entity or property of physics or being exclusively constituted by those posited entities or properties.[3]

[3] I use the term 'constitute' to mean the disjunction of a large inclusive family of ideas including constitution, reduction, realization, identity, and so forth, many of which are considered as rivals, and each of which, much like naturalism, is fairly imprecise and has distinct and conflicting conceptions (Melnyk 2003; Kim 2006). Such an ambiguity is unavoidable if we take into account the variety of metaphysical frameworks adopted by different naturalists, and if we want to avoid taking a stance in this internal dispute. While it is quite impossible to offer a comprehensive survey of these ideas here, I find it useful to follow Kim in using Smart's expression "nothing over and above" to capture the core idea and commonality of the family of ideas (Kim 2006, p. 275). Following Kim's interpretation, the idea is roughly that if Xs are nothing over and above Ys, then Ys are all we need for there to be Xs, and that no Xs can be considered as something in addition to Ys (Kim 2006, pp. 275–6).

There is also significant disagreement in metaphysics and philosophy of science concerning the nature of entities and properties. For our purposes, we may just take it that the entities here are the kind of thing that can be instantiated, and that the properties here are the kind of feature of such things that can be objectively conceived. But these are rough conceptions rather than

This chapter is concerned with MN, not physicalism. However, I take it that many arguments apply equally to both doctrines, so I will consider arguments by authors concerning the definition of physicalism.

DC is intuitively attractive to many, myself included. On the one hand, it appears to be consistent with most presumptions shared among philosophers about what MN should and should not be like. For one thing, as would be the case were DC true, we seem to determine whether or not something is natural by its relation to natural sciences. For another, it appears that DC matches well with widespread philosophical presumptions about the naturalness of particular entities or properties. Consider the paradigmatic cases that are often taken into consideration, such as tables, trees, Cartesian minds, God, and so on. All these cases can all be assigned their agreed upon categories by DC. Furthermore, the enlightenment materialists' and logical empiricists' attempts at a unified account of everything were found to be failures: the world is pluralistic and multifaceted, and thus cannot possibly be accounted for simply by appealing to certain metaphysical properties (e.g. those of 'matter' like spatiotemporality and solidity) or epistemic qualities (e.g. quantifiability and observability). Hence, while it is impossible to compare DC with other attractive approaches to characterizing MN here, DC seems to be more attractive than many of those which equate naturalness with a few metaphysical properties or epistemic qualities. This is because DC allows for a certain plurality of natural things, which is more in line with our current understanding of the world. The above considerations, then, provide us with sufficient reasons to develop DC to its fullest potential.

3. Problems Faced by DC

We have seen the intuitive pull of DC, but the approach also has several shortcomings. Only by being able to solve the problems, can a modified version of DC be successful. I will discuss them in turn.

formal definitions, and nothing in this chapter hangs on these particular conceptions. It is worth noting that some metaphysicians and philosophers of science believe that there are more fundamental and irreducible ontological categories such as capacities and activities, which can as well occur or be instantiated (for a nice survey, see Chen 2017). I see no reason why the DCs of MN and physicalism cannot expand to include such things, but for the sake of simplicity, I shall not take them into consideration.

3.1 Hempel's Dilemma

Hempel's dilemma was first formulated in Hempel (1969). If, on the one hand, MN is defined with reference to current science, then it is almost guaranteed to be incomplete and false, for it is almost certain that current science is an incomplete description of the world. If, on the other hand, MN is instead defined with reference to ideal science, then it is trivially true. That is, if ideal science is supposed to correctly describe everything (in the actual world), then MN is trivially true because everything (in the actual world) is automatically counted as natural.

3.2 The Problem of the Naturalistically Respectable in History

The second problem facing DC is whether historical ideas (or figures) can possibly be counted as being naturalistically respectable (Montero 2009). In history, many ideas (and theorists who believed in them) were typically considered naturalistically respectable. However, if we define MN with reference to current or ideal science, then most naturalistically respectable historical ideas must be counted as naturalistically unacceptable. For few ideas posit entities and properties that are also posits of current or ideal science (or are constituted by them). Hobbesian physics, for example, takes matter to be something like billiard balls: an object is describable only in terms of its spatiotemporal locations and motions; and it acquires a new motion only when another object is taking (and thereby pushing) it away from its current spatiotemporal location. But this view is rejected by contemporary physics. Unless a view like Hobbesian physics gets revived in ideal science, which seems unlikely, DC cannot count Hobbesian physics as naturalistically respectable.

3.3 The Problem of Naturalistically Unacceptable Scientific Theories

The third problem is the diversity of scientific theories. There are some current scientific theories that are presumably naturalistically unacceptable, but if MN is defined with reference to current scientific theories then these theories will be by definition naturalistically respectable. As Quine notes in

his criticism of MN, "Quantum mechanics today, indeed, in its neoclassical or Copenhagen interpretation, has a distinctly mentalistic ring" (Quine 1995, p. 257). The Copenhagen interpretation includes the observer effect, according to which the way in which a measurement is made determines part of the results. The mentalistic interpretation of the Copenhagen interpretation goes further and suggests that the observer effect is best explained by the consciousness of the observer. If the mentalistic interpretation turns out to be a legitimate scientific theory, then current science is positing some mental properties at the quantum level of the world. Similarly, eminent neuroscientist John Eccles (1994), together with physicist Friedrich Beck, proposed a mind-body dualism which posits psychons, a kind of mental 'particle'. Both of these theories posit fundamentally mental properties that are in conflict with what many philosophers consider to be the standard understanding of MN (see, e.g. Kim 2003; Montero & Papineau 2016; Oppy 2018a).[4]

3.4 The Problem of the Scientifically Inaccessible

The fourth problem facing DC is the worry that there might be natural entities or properties that science cannot ever access due to, say, some sort of causal isolation or due to our cognitive limitations (Montero 2009). If there are things science cannot access, and if they are not constituted by things that are of scientifically accessible kinds, then DC cannot count them as natural. Defining the natural and non-natural solely with reference to the posited entities and properties of our theories is beside the point, for not all limitations of our knowledge or epistemic capability are relevant to the distinction between the natural and non-natural.

3.5 The Problem of Non-Actualized Properties and Worlds Containing Them

The fifth problem facing DC concerns whether so-called 'alien properties'—i.e. properties not actually instantiated—and worlds containing them can be counted as natural (Stoljar 2010). The worry is that, simply because our

[4] It is worth nothing that some philosophers argue that MN can tolerate some fundamentally mental properties (e.g. Chalmers 1996; Zhong 2016). See Section 4 for further discussion.

actual science cannot study other possible worlds and matters within them, (almost) all alien properties cannot possibly be natural, and MN will turn out to be false in (almost) every possible world that contains alien properties.[5] While our standard view suggests that some alien properties are natural and that MN is true in some worlds containing them, DC cannot allow this.

4. The Similarity Approach (SA)

Some suggest that a simple modification of DC can solve the problems above. The suggestion is to allow extensions of current science, but that this extension be constrained by an adequate similarity to such science (Lewis 1983; Ravenscroft 1997; Braddon-Mitchell & Jackson 2007; McPherson 2015). This is known as the similarity approach (SA). Even though I do not endorse the approach, I consider it to be fairly attractive, for it effectively solves four out of the aforementioned five problems. It leads to an extension of the inclusive range of DC, and thereby avoids the narrowness in the first horn of Hempel's dilemma (i.e. defining MN with reference to current scientific theories only). Future and historical scientific theories can be counted as naturalistically respectable through being adequately similar to current science. Alien properties and scientifically inaccessible properties can also be accounted for if they are likewise similar to some posited properties of current science. The problem of naturalistically unacceptable scientific theories is the only remaining problem, for MN is defined with reference to similarity to current science that by definition includes all current scientific theories, whether naturalistically respectable or not.

There is, however, a more fatal problem with SA, which I call the *dilemma of the right kind of similarity*. To have a level of similarity that acts as a boundary between the natural and non-natural, we need to have the right kind of similarity to act as the measure. The options are either (1) an overall similarity or (2) a similarity in a certain respect. However, on the one hand, (1) cannot work, for any boundary drawn by using overall similarity cannot work as the boundary we expect for MN. More specifically, it seems hopeless to expect that all presumably non-natural things are less similar to things in

[5] I add the qualification 'almost' because there are in fact exceptions. Due to mistakes, instrumental concerns, and so forth, some properties are posited by our scientific theories even though they exist only in worlds other than ours.

current science than presumably natural things in terms of overall similarity. I will discuss this in detail below. On the other hand, we might instead follow (2) and consider all presumptively natural things to be similar in some specific respect—such as being spatiotemporal or being governed by the laws of nature (to use Armstrong's 1978, p. 261 and McDowell's 1994, pp. 71–6 characterizations of naturalism as examples). Unfortunately, in this case the natural is, in effect, defined in terms of those specific respects, not in terms of similarity. Hence, if we were to accept this option, we are in effect accepting some other approach to characterizing MN and thus giving up SA (and, very possibly, DC too).

The concern is whether some presumably natural things are, with respect to overall similarity, less similar to the posited entities of current science than some presumably non-natural things. Consider the following analogy. Compare an ordinary cat, a cursed black cat, and some imaginary mythological animals like Yetis, Leviathans, and werewolves. It appears that, on the one hand, at least some versions of Yetis, Leviathans, and werewolves, namely those without magical properties, could be natural entities. On the other hand, the black cats as conceived by Medieval Europeans, which were cursed by the devil and were supposed to bring forth bad luck, are obviously non-natural entities. However, it appears that a conception of naturalness that appeals to mere overall similarity can hardly make sense of these categorizations. The cursed black cat is identical to an ordinary cat which is posited by natural sciences in every respect except the curse, but Yetis, Leviathans, werewolves, and the like are posited to have different physiologies from animals recognized by natural sciences.

While one may argue that a single curse makes a more significant dissimilarity than plenty of radical physiological differences, it is difficult to see why this is the case, unless we construe similarity in a specific respect that is not that of overall similarity. One may also argue that the cursed black cat, the werewolves, and the Yetis are *all* dissimilar to natural things, but that werewolves and Yetis can be counted as natural if they are exclusively constituted by natural things like physical particles. However, the notion of overall similarity used in this case to measure naturalness would be very narrow—too narrow for the purpose of SA which is to expand the inclusive range of naturalness.

It might be argued that in the case above the real issue is not overall similarity of entities but rather the overall similarity of properties. That is, the property of being cursed is radically different from ordinary biological and physiological properties. This might be true, but shifting the concern to

properties does not resolve the problem. Consider the locational properties of strings and Hobbesian matter: it has been suggested that strings are not spatiotemporally located like Hobbesian matter but are instead instantiated in a non-local manner in eleven dimensions. While it appears that both of these locational properties are natural, the question remains as to the way in which we can say that their similarity is greater than, say, that between being vitalistic and being organic, or that between being karmic and (the ordinary sense of) being causal. I doubt that there is a satisfactory answer to this question. Since we are here comparing the degree of overall similarity of things that are dramatically different in nature, such a comparison risks committing a category mistake, or at the very least the relevant degrees are too vague to be compared unless we use a notion of specific similarity rather than overall similarity.

In the face of this problem, those who are sympathetic to SA and thus do not wish to repudiate it have two options. The first option is to accept that the inclusive range of MN is different from what most believe. After all, some philosophers argue for a revision of the inclusive range of naturalness. For example, it has been argued MN should tolerate some fundamentally mental properties.[6] Though this option is attractive, it is costly, for if we revise naturalness strictly according to degree of similarity, our presumptions concerning the naturalness or non-naturalness of many things (and the naturalistic respectability of many theories) might have to be radically revised. Furthermore, there is a certain unpredictability about the results the revision would produce. For instance, it might be the case that panpsychism and pantheism would need to be counted as naturalistically respectable, whereas Hobbesian physics and Newtonian physics would need to be counted as naturalistically unacceptable. If this turns out to be true, many discourses concerning MN in philosophy, for or against the doctrine, would have to be considered wrongheaded. This consequence may be too costly even for many of those who are sympathetic to the revisions.

The second option is to keep our assumptions about MN and the naturalness of things, and to develop or revise SA to bring forth a regimentation of its inclusive range which could be consistent with those assumptions. I believe that there is no non-radical way of development that can escape the dilemma of the right kind of similarity. My option is an intermediate proposal between repudiating SA and the second option: to *radically* revise

[6] See, e.g. Chalmers 1996; Zhong 2016. Note that these people do not deal with SA and its problems.

SA. More precisely, I will develop a novel version of DC with the same strategy that SA uses, whereby *extensions* of current science are allowed.

5. The Historical Paths Approach (HPA)

I call my approach the *historical paths approach* (HPA). HPA is based on a concept I use to replace the similarity relation used by SA, which is what I call the *historical paths* (HPs) of contemporary science.[7] HPs are the processes of modifying and developing scientific concepts over the course of history, along which scientific concepts are modified and finally developed into the versions in contemporary natural sciences. (Note that old paths that are disconnected from contemporary scientific concepts do *not* count as HPs.) Of course, the relevant 'concepts' here are theoretical contents and not methodological factors: for example, the concept <momentum> within Newtonian physics, the concept <positron> within contemporary particle physics, and the physicist's preference for mathematically quantifiable theoretical posits. What is suggested here assumes a distinction between concepts and theories, with the latter including Newtonian physics, the theory of relativity, string theory, the Darwinian theory of evolution, and so on; and HPs are about concepts, not theories. Although similar distinctions are seen in the works of many influential theorists, such as in Nagel 1961, Lewis 1972, Laudan 1977, Kuhn 1996, and Jackson 1998, and, not every philosopher will make such a distinction because those concepts themselves might be, as some theories of concepts suggest, small, component theories. I take no stance on this controversy; and I suggest that those who dislike the distinction can replace the notion of concepts here with the notion of small component theories which are components of bigger theories or theoretical frameworks. Some might wonder why I intend HPs to be about concepts but not theories. The motivation for my position is that only when HPs are about concepts can it resist the *problem of overinclusion*, which is the worry that HPs will be overly inclusive and thus useless. I will return to the problem in Section 5.2.

Let us return to the idea of HPs. Even though the concepts within Newtonian physics might be very different from those within string theory,

[7] In the case of physicalism, there might be some HPs of contemporary physics. Nonetheless, I will focus on MN and not assess whether or not an HPA to characterizing physicalism is attractive.

they have some substantive historical connections. For example, <momentum> is developed from <impetus> in medieval science; <gravity> in Einsteinian theory of relativity is developed from its counterpart in Newtonian physics, and it, in turn, has developed into <quantum gravity> in string theory. Of course, lots of scientific concepts have no older counterparts. For example, it is unclear that <positron> and <Higgs boson> have older counterparts. But they can be said to have been developed in close relation to concepts we already have, such as <electron> or <photon>, or the concept of physical particle in general. It is also important to note that we are talking about historical *paths*, not *a* historical path: there is more than one path, even within the same scientific discipline; and a single concept might be traced back to multiple paths. For instance, <light>, to which we now attribute the notion of wave-particle duality, can be traced back to its two counterparts in the particle theory and the wave theory, each of which can, in turn, be traced back to different paths.

On the other hand, some theoretical contents and methodologies in other scientific theories—such as Aristotelian teleology and some versions of vitalism—are (relatively more) disconnected from HPs, even though there may have been good reasons in the past for using these concepts. This is because many of the central concepts of those theories were wholly abandoned rather than modified and absorbed by new theories in a historical progression. For example, the concept of telos was abandoned rather than modified in physics. In other words, many of their central concepts fail to have descendants in current science.

If my hypothesis about the existence and nature of HPs is correct, then we may use the idea of HPs to formulate a definition of MN—call it the *historical paths definition* (HPD). The basic formulation is as follows:

HPD1: Every entity or property instantiated in the actual world is natural in the sense of being a posited entity or property of some theoretical concept that is on some HP.[8]

Four remarks are in order. Firstly, not all scientific concepts posit entities or properties. What is important is that if an entity or property is natural, then there should be at least one concept that posits it. Secondly, the concepts of

[8] I omit here the clause in the original version of DC 'or being exclusively constituted by (i.e. being reducible to or realized by) those posits'. Since I am going to develop some fairly complex variants of HPD1, I purposefully set the clause aside to avoid unnecessary confusion.

some natural effects—by which I mean something like the causing of asphyxia by rail travel at high speed—might not be on the HPs. But these effects might be considered as the expected causal behaviours of some natural entities and thereby as natural. Thirdly, of course not all naturalistically respectable scientific concepts in history have descendants and are thereby on some HPs. I will get back to this problem in Section 7.

Fourthly, it is not impossible for there to be some very different HPs from the perspective of future science. For instance, it is possible to return to Aristotelian teleology, and a future scientist might therefore count our current scientific concepts to be ones that fail to have descendants. But this is not a problem. It would be problematic only if we assumed that our conception of MN ought to be modally and chronologically neutral in the sense that it leaves open which possible world the actual world is and when the present is, and could thereby account for counterfactual and future judgments. As we all have learned from Putnam and Kripke, not even natural kind terms like 'water' are modally and chronologically neutral: the meaning of the term 'water' depends on the empirically discovered fact that the watery stuff in the actual world is H2O, but this is not true for otherworld inhabitants, nor is it the way our ancestors understood the term. So we should be open to the possibility that future generations may have different conceptions of MN, just as the possibility they may discover that the watery stuff in the actual world is not H2O but something else.

At this point, my approach is yet to be completed—most problems facing DC are yet to be solved. However, before I further develop it, two problems must be addressed. They are (1) the *problem of disconnection* and (2) the *problem of overinclusion*. Only when these two problems are addressed, can the nature of HPs be made clear, and can we see a large part of how HPA can resist the worry motivating the dilemma of the right kind of similarity. Remember, the worry is there in the case of SA because no level of overall similarity can act as the required boundary which allows us to count things as natural or non-natural as we expect MN to do. Addressing these two problems allows us to see how the inclusive range of the HPs is shaped and can be used by MN as a boundary, which I believe to be more useful than the boundary drawn by SA.

5.1 The Problem of Disconnection

The problem of disconnection is the worry that HPs are not spread over substantial lengths of time. For instance, some followers of Kuhn's (1996)

may claim that major developments in the history of science occur with paradigm shifts (i.e. with radical changes of the most fundamental assumptions within short periods of time). With these radical changes, it is unclear that there is any joining path because there might be too many disconnections between scientific theories.

An in-depth discussion of the philosophy and history of science (and of Kuhn's views) is outside the scope of this chapter, but let me outline my response briefly. It is difficult to see why the sceptic is correct. On the one hand, Kuhn's theory of paradigm shifts is very controversial and is not without alternatives (e.g. Laudan 1977). On the other hand, even granted that the theory is correct, it undermines only the view that concepts such as <gravity> and <space> are defined or understood in exactly the same ways before and after the paradigm shifts. However, this does not imply that there cannot be some kind of HP, on which old concepts are modified and absorbed by the new theories. In fact, Kuhn himself believes there are "borrowed elements" from the old paradigm, "both conceptual and manipulative", that are incorporated into the new ones (1996, p. 149). The examples he provides include <planet> and <space> (pp. 128, 149). It is true that he believes the borrowed elements are more radically revised than many believe: they form new relationships with other concepts in the new paradigms, and are, in addition, incommensurable with their older counterparts. But it is hard to see why this alone conflicts with the idea of HPs, for the idea does not commit itself to commensurability.

5.2 The Problem of Overinclusion

The other problem that follows is the problem of overinclusion. Some might, contrary to the adherents of the problem of disconnection, consider HPs to be overly inclusive and thus useless, since it is possible to understand our current science as descending from belief systems that are largely naturalistically unacceptable. For instance, some suggest that modern chemistry is a descendant of alchemy, which is full of magical concepts.

To respond to this objection, two conceptions have to be distinguished. Recall that HPs are about *concepts (or small, component theories)*, not *belief systems (or bigger theories)*. For example, HPs are not about whether alchemy as a belief system is the ancestor of modern chemistry but, rather, which particular concepts in alchemy are the ancestors of those in modern chemistry. In fact, as I have mentioned earlier, the problem of overinclusion is why I intend HPs to be about concepts—because HPs about theoretical

frameworks, belief systems, or big theories cannot avoid being overly inclusive, but HPs about concepts can. Alchemy, as a discipline, is a belief system containing many concepts, including theoretical and methodological concepts about magic, classifications of chemicals, measurements, and so on. As history progressed, some of them were developed and absorbed by chemistry while others were eliminated or abandoned entirely. HPA has no problem taking some concepts in alchemy, such as its observations of chemical reactions, to be on the HPs, thereby taking them to be naturalistically respectable. Yet it notes that the theoretical concepts about magic within alchemy have been entirely abandoned and should therefore be considered naturalistically unacceptable. Put simply, only some but not all of the concepts within alchemy are on the HPs, and the presumptively naturalistically unacceptable ones are not on the HPs.

A worry is that, say, the alchemist's concept of gold and the vitalist's concept of muscle should not count as naturalistically respectable—the alchemist might believe that gold is a spiritual substance, and the vitalist certainly believed that muscles are powered by vital forces—but these concepts are doubtlessly on the HPs. The worry can be addressed by slightly modifying HPD1:

HPD2: Every entity or property instantiated in the actual world is natural in the sense of being a posited entity or property of some theoretical concept(s) (1) that is on some HP and (2) every constituent concept of which (if any) is also on some HP.

The (2) clause is added to the definition as a condition of being naturalistically respectable. The idea of a constituent concept is as follows. There are complex concepts that are partially 'made up of' constituent concepts. For instance, <bachelor> is made up of <unmarried> and <male>. Let us apply HPD2 to the case of muscles powered by vital forces. While the vitalist believed that muscles are powered by vital forces, her concept of muscles may or may not be independent of her vitalist beliefs—for her concept of muscles might be based solely on, say, empirical observations and have nothing to do with her vitalism. If, on the one hand, Vitalist *X*'s concept of muscles is not made up of vitalist concepts, then we may count the concept as naturalistically respectable and its posits as natural. If, on the other hand, Vitalist *Y*'s concept of muscles is partially made up of her vitalist concepts, then even if the concept is on HP, we may count it as naturalistically unacceptable and its posited entity as non-natural. This strategy of

analysing concepts allows us to count spiritual gold and muscles powered by vital forces as non-natural.

It might be objected that HPD2 is too strong. For instance, Isaac Newton was a theist. He certainly believed that everything was created by God, which we count as non-natural. Let us assume that he was so pious that whenever he formed a concept of a thing, he always bore in mind that it is a divine creation; call him the *pious Newton*. Does HPD2 imply that the pious Newton, and many other influential theist scientists like him, had no naturalistically respectable concepts of things at all, and that every posited entity of their concepts is not natural?

My answer is mixed. On the one hand, we can indeed say that the pious Newton's concepts of mass, gravity, and chairs are naturalistically unacceptable, and that every posited entity of his concepts is non-natural. This is not a significant problem, though, for what is important for us is not whether the pious *Newton's* concepts are naturalistically respectable, but whether some *Newtonian* concepts can be understood or interpreted in a naturalistically respectable way which is independent of his theism. Obviously, we do not need Newton's theism (or any other version of theism, for that matter) to understand, say, Newtonian mass and gravity. On the other hand, we can analyse the pious Newton's concepts in a more flexible way, and be more generous in granting naturalistic respectability. The strategy is to make use of counterfactual conditionals. Let us assume there is an idealized perfectly rational agent *R* sharing all beliefs of the pious Newton. It seems unlikely that, were *R* presented with compelling evidence that God does not exist, *R* would have abandoned his concepts of mass and chairs, or substantively or radically revised them. The reason is that *R* could have just slightly revised those concepts by removing the idea that they are created by God. Nothing concerning the general conceptions of what those entities or properties are like would have had to be modified.[9] With this in mind, we can take it that the pious Newton's concepts of mass, gravity, chairs, and tables are by and large naturalistically respectable, and that these concepts, understood in this way, can thereby posit natural entities and properties.

For the sake of clarity, and to avoid them becoming too complex and wordy, I shall not add the additional clause of HPD2 'and (2) every constituent concept of which (if any) is also on some HP' to the other variants of

[9] For a similar approach, see Lewis's (1972) theory of near-realization. Note that he is not dealing with the problem I am considering here.

HPD1 I am going to propose. Nonetheless, it is important to note that they can all be modified in the same way.

With the problem of disconnection and the problem of overinclusion solved, and the nature of HPs made clear, we see how the inclusive range of the HPs is shaped, and that the worry motivating the dilemma of the right kind of similarity dissolves. For cases like Hobbesian physics, even though they might not be that similar to current science, most if not all of their main concepts have been continuously modified and finally developed into those in current science. On the other hand, the concept of cases like fundamental mental properties, vital forces, karma, cursed animals, and the like are not on the HPs. Certainly, it can still be questioned whether the HPs can exclusively and exhaustively cover every natural entity or property that is posited in the history of science. The possibility that there are some other counterexamples cannot possibly be expunged. Nonetheless, as far as I am concerned there are no examples that cannot be handled by my developments of HPA below. The problem is, I believe, significantly less severe than it was in the case of SA.

6. Future Scientific Theories

I have offered HPD1 as the basic formulation of the historical paths definition, but it does not allow for future scientific theories to be counted as naturalistically respectable and it thus falls prey to Hempel's dilemma. This is because future theories are very likely to have new concepts that are not ancestors of current concepts. A modified version of HPD1 can solve this problem:

HPD3: Every entity or property instantiated in the actual world is natural in the sense of being a posited entity or property of some theoretical concept(s) that (i) is on some HP, or (ii) on some possible future path that is a reasonable extension of our current HPs.

Condition (i) is added to the definition as a sufficient but not necessary condition alternative to condition (ii). When presented with a concept, we form reasonable expectations of whether it could be a descendant of some our current scientific concepts developed via scientific advancement. For example, a physical particle that quarks are composed of is something we could expect to discover, and, in addition, we can treat its concept as a

possible descendant of our current concept of physical particles.[10] On the contrary, a psychic universe described by some New Age theories is no such thing—while its concept was developed from its scientific counterpart, no future science is going to tell us its existence. Some might wonder whether the judgment concerning reasonable extensions here is a matter of similarity to current theories and hence a retreat to SA. I do not think so. At least, this is not merely a matter of similarity. Many other considerations can be invoked to bring forth a more sophisticated judgment than what mere similarity can provide: e.g. tendencies of what new discoveries are like, the scientific method in general (which includes principles such as the need for empirical observations or experiments, the preference for repeatability, and so on), theoretical virtues in general (such as simplicity, explanatory scope, and so forth), technological limitations, research strategies, and methodologies used by experts in different branches of science.[11]

Admittedly, the criteria of these judgments have to be fairly imprecise because they are based on abstract conceptions of what current science is like. The kind of judgment needed might be based on, say, what most reasonable laymen who are sufficiently informed about current science would accept as a possible science of a century later. These judgments are not the precise technical expectations a scientist would have with regard to possible discoveries based on current findings (say the expectation of a nanophysicist that silver might be used in the future as a conductor for a particular purpose). If the expectations are limited to precise technical expectations then revolutions in science cannot be accounted for. A set of expectations that is wider in scope is needed.

Some would be unsatisfied with the imprecision involved, but it is not a failing of the approach. As Ravenscroft argues in his defence of SA, on the

[10] No doubt, some possible future paths might be expected because of social or psychological factors, such as corruption in the scientific community, political upheaval, cultural influence, and the extinction of human beings, which most of us consider to be irrelevant to the discussion here. By the clause 'some possible future path that is expected through extending our current HPs', I only mean those HPs that are expected to develop because of unbiased and undisrupted scientific inquires, which are not intervened by the above factors. This move involves what Godfrey-Smith calls an "idealised simplification" (Godfrey-Smith 2014, p. 21; see also Musgrave 1981), in which a theorist deliberately imagines things to be simpler than they actually are, in order to construct a theory of significance.

[11] It is worth nothing that there is an asymmetry between HPs and possible future paths discussed here: it is a historical fact as to what HPs there are, but not all possible future paths will turn out to be true, and we often cannot tell which will and which will not. This is not a problem, however, for it is beneficial rather than detrimental to be inclusive of the possible future paths that will not turn out to be true. See the next paragraph for the reason for this.

one hand, what is most important is to be able to categorize most paradigmatic cases of the natural and non-natural; on the other hand, our conception of the natural is indeed imprecise and vague: if a characterization of the natural is committed to vagueness, the commitment might be an explanation of the vagueness within our conception of the natural, which is a virtue rather than a failing (Ravenscroft 1997, pp. 425–6). I think Ravenscroft is correct, at least when the two kinds of vagueness correspond to each other. The same can be said about HPD3. On the one hand, it is obvious that the kind of expectation HPD3 makes use of enables us to categorize most paradigmatic cases, such as the fundamentally mental, God, undiscovered particles, tables, and chairs, as we expect of MN. None of these falls in the marginal area between the natural and non-natural where vagueness is an issue. On the other hand, the naturalness (or naturalistic respectability) of some peculiar cases, such as qualia, free will, the self, and some of the posited entities of European phenomenology, is indeed difficult to determine;[12] and these cases also seem to be where the vagueness of the boundaries of HPD3 lies. Hence, the vagueness may be helpful in explaining why such uncertainties exist. I will discuss an additional advantage of this view in Section 9.

7. Naturalistically Unacceptable Scientific Theories

The next problem is that of naturalistically unacceptable scientific theories, according to which there are some current scientific theories that are presumably naturalistically unacceptable and which a standard DC would by definition count as naturalistically respectable. In response to this problem, one might argue that there is something like a *sudden divergence* from the general directions of the paths. That is, theories such as the mentalistic interpretation of the Copenhagen interpretation and Eccles's theory of psychons, by positing concepts such as the role of consciousness in the mechanics of fundamental physics, are not going along the general directions of gradual changes, towards which most concepts on the relevant paths are moving.

The idea of a sudden divergence is not ad hoc because it not only allows us to account for why some scientific theories are presumably naturalistically

[12] For a good overview of these debates, see Gallagher et al. (2015). For an example of the posited entities of European phenomenology whose naturalness is under debate, consider Merleau-Ponty's embodiment and embeddedness. For some contemporary contributions to the debate, see Varela, Thompson & Rosch 1991; Pollard 2014; Gallagher 2018.

unacceptable, but it is also useful in its own right, for it can (partially) explain the research directions of some scientists and philosophers. For example, some scientists are motivated to propose new interpretations of quantum mechanics that are alternatives to the Copenhagen interpretation in order to avoid its use of the (possibly mentalistic) observer effect. Consider the example of theoretical physicist Cramer, the proposer of the transactional interpretation. He writes:

> The knowledge interpretation's account of state vector collapse and nonlocality is internally consistent but is regarded by some (including the author) as subjective and intellectually unappealing. It is the source of much of the recent dissatisfaction with the Copenhagen interpretation. The author has proposed an alternative and more objective interpretation of the quantum mechanics formalism called the transactional interpretation. (Cramer 1988, p. 228)

In addition, some scientists and philosophers particularly expect and welcome such proposals (e.g. Price 1996). HPA offers a possible explanation of their motivation: they might, explicitly or tacitly, believe that the Copenhagen interpretation is a sudden divergence (though this belief is contestable) and thus want to avoid accepting it.[13] If this explanation is correct, then there can be a rational basis for this kind of motivation other than a mere intuitive discomfort about introducing the fundamentally mental. After all, it might be rational to have a conservative attitude. And since what I am describing here are probably *tacit* attitudes, this is compatible with the possibility that these theorists might have explicitly accepted some conceptions or characterizations of MN different than those developed by HPA.

8. The Alien and the Scientifically Inaccessible

We have seen how HPA survives objections from Hempel's dilemma, the problem of the naturalistically respectable in history, the problem of

[13] It is worth noting that I am not arguing that the observer effect in the Copenhagen interpretation is non-natural, nor that its concept is a sudden divergence. I am, in fact, inclined to think that the observer effect lies in the vague boundary area between the natural and non-natural. What I am suggesting is that those who (explicitly or tacitly) understand its concept as a sudden divergence might be rational to look for alternative interpretations; I am not arguing that their view is correct.

naturalistically unacceptable scientific theories, and even the dilemma of the right kind of similarity. What are left are three problems: the problem of the scientifically inaccessible, the problem of non-actualized properties and worlds containing them, and the problem concerning naturalistically respectable concepts in history that have no descendants. HPD3 cannot solve the former two problems because alien properties and scientifically inaccessible properties are not things we can discover in future science. Something similar can be said about cases like the concept <optical aether>: even though it is not logically impossible that we may rediscover the optical aether in future science, this is not expected by any sufficiently reasonable and informed person. But, again, a modification of the definition can solve the problems. The idea is to allow some variations of the natural things recognized by the previous formulations of HPD. The precise new formulation of HPD is as follows:

HPD4: Every entity or property instantiated in the actual world is natural either (1) in the sense of being a posited entity or property of some theoretical concept that (i) is on some HP or (ii) on some possible future path that is a reasonable extension of our current HPs, or (2) in the sense of being a robust intradisciplinary recombination of the theoretical characters of some natural entities or properties identified by condition (1).[14]

Condition (2) is added to the definition as a sufficient but not necessary condition of naturalness. Let me first outline the theoretical characters concerned here, before I move on to the more complex idea of their robust intradisciplinary recombinations. By the term 'theoretical characters', I mean the framing of the intrinsic or relational characters of the relevant entity or property in a scientific theory or theoretical framework. There is substantial research in our best philosophy of science on such theoretical characters, and some major kinds of theoretical characters have been

[14] Stoljar (2010, p. 88) argues against the use of cluster definitions of physicalism, where each of which consists of a cluster of independent, unrelated conditions. The worry is that it is arbitrary to consider various such conditions as a single doctrine when physicalism is expected to be a non-arbitrary and systematic metaphysical doctrine. This objection does not apply to HPD4, which consists of the disjunction of several conditions and is fairly complex. For condition (2) is a natural extension of condition (1) to deal with peculiar cases, but is not conceptually independent of it. Hence, despite the complexity of its definition, an MN characterized by HPD4 is a systematic metaphysical doctrine.

identified. For instance, some theoretical characters are surveyed in Craver and Tabery (2015):[15]

A. *Generation of phenomena*. The behaviour of the thing concerned as a whole, which can be classified into:
   1. *Production.* The thing concerned brings about some product.
   2. *Underlayer.* The thing concerned underlies the instantiation of something of another kind.
   3. *Maintenance.* The thing concerned holds some state of affairs or some range of states of affairs in place.

B. *Parts*. The parthood required of the composition of the thing concerned.

C. *Organization*. The organization of the parts of the thing concerned which is characteristic of it, and which can be classified into:
   1. *Spatial organization.* Location, size, shape, etc.
   2. *Temporal organization.* The order, rate, duration of the component activities.

D. *Levels.* The location of the thing concerned in the hierarchy of levels of things.

E. *Natural kinds*. The kinds of things to which the thing concerned belongs and whose members share a substantive, scientifically recognized similarity to each other.

This borrowed list is not meant to be exhaustive, but it sheds light on what theoretical characters there are.[16]

So far, so good. Let us get back to our discussion of HPD4. The additional condition (2) is not only about theoretical characters, but also about robust intradisciplinary recombinations of them. To have a recombination of

[15] It is worth noting that Craver and Tabery belong to the new mechanist school of philosophy of science, and thus their aim is to survey the characters of 'mechanisms', which they take to be the protagonist of particular special sciences. Nevertheless, I believe that those theoretical characters generalize to other posited entities and properties of natural sciences. It is also important to note that since I am taking the list of theoretical characters out of Craver and Tabery's new mechanist context, I have altered the descriptions of a few of them in order to make them more general.

[16] Not every reader will agree with my understanding of theoretical characters—for example, Lewis (1972, 2009) has a more monistic account of theoretical characters according to which these characters are largely about causal-nomological roles. This is not a problem, however, for HPD4 is neutral in this regard: such a reader could replace my conception of theoretical characters with her own; the definition is still useful. My conception is, after all, merely a recommended way of understanding theoretical characters; nothing hangs on it.

theoretical characters is to put them together in a new way. A toy example is that a unicorn is a recombination of a horse's theoretical characters and a horn, a parthood character shared by many horned animals such as rhinoceroses. The robustness here, then, concerns (1) the internal logical and conceptual consistency of the recombination and (2) its consistency with the theoretical frameworks in which we extract the relevant theoretical characters. (1) demands that the recombined thing cannot be a logically or conceptually impossible entity or property like a square circle; (2) demands that the recombined thing cannot be an incomplete entity or property according to the relevant theoretical frameworks: for example, no theoretical frameworks in animal studies can make sense of an animal without a physical body. An intradisciplinary recombination, then, is a recombination that makes use of theoretical characters merely from a single scientific discipline, such as physiology, molecular biology, or particle physics. A recombination of theoretical characters found in, say, cognitive science and particle physics (e.g. an electron that has emotions) does not count as an intradisciplinary recombination.

Let me elaborate on the application of HPD4 in light of a few examples which range from easy to difficult. We mentioned above the case of a unicorn, which is a recombination of the theoretical characters of a horse and some horned animals. It appears that it is a natural entity, unless it possesses some mysterious magical powers as described in some of the relevant folktales, for the relevant theoretical characters are all extracted from natural entities of some HPs, namely the horse and the horned animals. Similarly, a werewolf is a recombination of theoretical characters of human beings, wolves, and metamorphosis.[17] With the same approach, we can also deal with the case of the optical aether. Despite its being massless and transparent, the optical aether was believed to be a physical substance which is much like other physical substances posited by classical physics and which shares almost the same theoretical characters as them, such its being microphysical particles and its capacity to have and transmit motion. It is in fact in virtue of its possession of these theoretical characters that it was claimed to act as the medium of light waves, in the same way that water acts as the medium of water waves.

[17] Of course, an easier way of counting these animals as natural is, as mentioned in Section 6, to take into account the possibility that they are exclusively constituted by natural things like physical particles. Even so, I believe that consideration of naturalness at the level of special sciences is nonetheless useful in the methodological application of MN. See Section 9 for a related discussion.

HPD4 applies to properties as well. First, consider twin-mass, twin-charge, and twin-spin, which are considered by Stoljar (2010, p. 79) in his discussion of the problem of non-actualized properties. Suppose that an entity with twin-mass produces no gravitational forces but instead produces twin-gravitational forces. Twin-mass can nonetheless be counted as natural. For even though gravitational forces and twin-gravitational forces are different forces, they are identical in terms of their *format of operation*. Hence, twin-mass brings about a product that is *identical in format* to that of mass, and can be considered as a recombination of the theoretical characters of the latter. The same can be said about twin-charge and twin-spin. Consider, then, an example of a natural property that was posited by science in the past and whose concepts have no descendants left nowadays: levity. Scientists in history explained light substances like air and fire going upward by positing that such substance have levity in them. In addition, some chemists in history like Lavoisier believed a fluid called phlogiston left a body when it burned, and since things were heavier after burning—we now know this was due to oxidation—phlogiston was also believed to have levity. Levity can be considered as a negative version of weight (as understood in history), for it produces phenomena in a very similar manner to weight, albeit having a revised vector. In this light, it can also be considered as a recombination of the theoretical characters of weight.

Let me now show that the idea of robust intradisciplinary recombination in HPD4 does not lead to an overinclusion, namely the counting of presumptively non-natural things as natural. For our purposes here, I shall focus on the example of fundamental mental properties—which Kim, following Roy Wood Sellers, takes to be a crucial test case for characterizations of MN (Kim 2003, p. 96). The problem concerns how theoretical characters of things in, say, cognitive science are recombined: can the recombined thing be both mental and fundamental? I do not think so. Such recombinations can only be done in three ways, all of which are either unsuccessful or are not permitted by HPD4. Specifically, the recombination could involve (1) combining mental properties with fundamental physical things like electrons, (2) having ontological fundamentality as one of its theoretical characters, or (3) not having intrinsic characters concerning parthood. (1) is not permitted by HPD4 for the obvious reason that it involves interdisciplinary recombination, whereas HPD4 only allows for intradisciplinary recombination. (2) is not permitted for the same reason: it is not within the scope of cognitive science to posit ontological fundamentality; this theoretical character has to be found in particle physics. (3) will not be successful in making the

recombined thing ontologically fundamental, for, on the one hand, not having intrinsic characters concerning parthood merely means that the descriptions of the thing concerned in the relevant theory(ies) 'bottom out' in the sense that descriptions of lower-level components come to an end. Such descriptive bottoming out does not imply ontological fundamentality, though it is compatible with it (Machamer, Darden & Craver 2000, pp. 13–14). On the other hand, it appears that the demand of robustness from contemporary cognitive science requires its posited things, bottoming out or not, to have some particular hierarchical characters. More specifically, its posited things should be located at some particular hierarchical level of things in the sense of constantly co-instantiating with some neurological processes, artificial intelligence, or the like, and being roughly compatible with the possibility that they be grounded in them.[18] This effectively eliminates the possibility of the recombined cognitive thing being ontologically fundamental.

9. The Usefulness of HPA

I believe that HPA is not only defensible as I have argued above, but also useful in several ways. Consider the way I handled the problem of naturalistically unacceptable scientific theories in Section 7 where I appealed to their sudden divergences. I argued that it not only allows us to account for the reason why some scientific theories are presumably naturalistically unacceptable, but that it is also useful in explaining the research directions of some scientists and philosophers who attempt to remain naturalistically respectable and, in addition, providing a rational basis for these directions.

There is another reason why HPA is useful. While the naturalness of some entities, properties, and states of affairs—such as qualia, free will, the self, some of the posited entities of European phenomenology, and the observer effect in the Copenhagen interpretation—is intensely debated, HPA offers us a useful framework to resolve these debates. HPA does not, as some other characterizations of MN do, offer answers to those debates by simply saying that those things are by definition natural or non-natural—which may imply that one group of philosophers (or scientists) are inexplicably

[18] Of course, cognitive science is not metaphysics, and thus the relationship between the relevant hierarchical levels are typically not restricted to particular metaphysical framings such as a reductionist one; they often are compatible with views like emergentism.

misguided.[19] Instead, following Ravenscroft's argument, HPA acknowledges that those debates are substantive and cannot be easily settled: the naturalness of those things is indeed difficult to determine, and might even be currently impossible to determine; for it is ambiguous whether they relate to the HPs in the way required by HPD4. This provides an elegant framework for different parties to settle their debates and reach some consensus: while we can all agree that some cases lie in the vague boundary area between the natural and non-natural, progress can be made by inquiring into the relevant scientific disciplines and thereby improving our ideas of their paths, which include our understanding of their historical paths and our expectations of their future paths. For instance, when two parties disagree about the naturalness of qualia or some of the posited entities of European phenomenology, an inquiry into recent neuroscience or cognitive science might change our view on what the historical and expected future paths of the discipline are like, and might thereby provide us with some clues as to how we can settle the debate (for examples of such projects see Hohwy & Frith 2004; Gallagher et al. 2015; Tononi & Koch 2015; Chan & Latham 2019). This provides a useful framework for the naturalization of the philosophy of mind and phenomenology.

10. Final Thoughts

I have proposed HPA by using the strategy of SA, which is to search for an extension of current science. I have argued that it can survive the main problems facing DC and the dilemma of the right kind of similarity facing SA, that it can explain some patterns of behaviour and intuitions of scientists, and that it can offer us a useful framework to resolve some philosophical debates. Braddon-Mitchell and Jackson (2007, pp. 34–5), two proponents of SA, remark that 'The vagueness in [SA] can perhaps be left to advances in philosophy and indeed in science itself to sort out.' I consider my proposal to be a development of SA because they share the same strategy, which is to search for an extension of current science. It is in this light a response to Braddon-Mitchell and Jackson's invitation.

[19] For a detailed discussion of this problem and the implications it has for philosophy and science, see Gallagher 2018.

Moreover, it provides a framework that allows other advances in philosophy and science to reduce the relevant vagueness in an easier way.[20]

One possible worry is that HPA is largely based on socio-historical but not metaphysical considerations, for it is somewhat a socio-historical issue as to which concepts are found on the HPs and which are not. This is counterintuitive and unsatisfactory. For when we use the term '*metaphysical* naturalism', we are usually trying to express a metaphysical notion, not a merely socio-historical notion.[21] But this would be a mistake. MN can be understood as a minimalist or parsimonious metaphysical framework. Its exclusive commitment to entities or properties posited by HP-related theoretical concepts is a pursuit of ontological economy, which is itself not an HP-related theoretical concept but a metaphysical commitment. To be more precise, every reasonable metaphysical framework has metaphysical commitments to entities and properties posited by HP-related theoretical concepts; the naturalist is a minimalist or parsimonious metaphysician who always restricts her metaphysical commitments to such entities and properties and is unwilling to have any additional commitments. By contrast, the non-naturalist goes beyond the minimalist restriction and has additional metaphysical commitments such as <God> and <the fundamentally mental>.[22] Some might find this solution unsatisfactory, for it can only count MN and not the *notion of the metaphysically natural* as a metaphysical notion: while the exclusive commitment to HP-related theoretical concepts is a metaphysical doctrine, HP-related theoretical concepts are still defined solely in socio-historical terms. But this is not so. Once the above idea of MN is accepted, then being natural is not only about adequately relating to HP, but also about playing certain roles in the minimalist metaphysician's ideology: the natural is the only kind of thing the minimalist metaphysician would posit. This is, again, a metaphysical issue.

One last point. This proposal of HPA is still incomplete, for the idea of HP is historical and there is certainly a problem of historical accuracy. No doubt, there might be historical facts that are in tension with the proposal

[20] Of course, as I noted in Section 5, there is still vagueness in my approach. But I have argued, following Ravenscroft, that the existence of vagueness in a characterization of MN is not necessarily a failing and might even be a virtue. What is important here is whether or not unwanted and problematic vagueness—such as that which leads to the dilemma of the right kind of similarity—is reduced; and if my arguments in this chapter are on point, it is.

[21] For a related argument, see McPherson 2015, pp. 124–8.

[22] For a more detailed and sophisticated discussion of the minimality of MN, see Oppy 2018b, pp. 34–6.

outlined in this chapter, for the history of science is far less linear, systematic, and unified than it may seem: there is no absolute guarantee that, say, all concepts of spirits and souls are not on the HPs. Of course, even if there are a few such cases on the HPs, they might be taken as insignificant noises that can be ignored when we construct a simplified, general theoretical model of what the HPs are like.[23] For example, the possibility that, say, a few spiritual concepts are on a few HPs should not be necessarily irreconcilable with a simplified, general model of HPs that takes them as some scientific developments that have got rid of spiritual concepts. But, again, any such model is yet to be verified.

This chapter is concerned with metaphysics—or, more precisely, the metaphysical foundation of the discourses in several philosophical areas, such as the philosophy of mind, contemporary phenomenology, the philosophy of religion, and metaethics—and not historical research. Its commitments to history are not claimed to be true in terms of a historical study, but rather should be seen as a hypothesis that is waiting to be assessed on the basis of empirical studies in the history and philosophy of science. A possible solution I favour is to see the concept of HP as a folk scientific concept or even a popular myth. According to the solution, what HP are should be assessed in terms of intuitions about the history of science, or what popular science tells us about the history of science, rather than real historical research. This might release the user of the doctrine of MN from the burden of requiring professional historical knowledge. After all, while such intuitions are perhaps historically wrong or inaccurate, they are indeed what many philosophers (and scientists) believe, and what motivate many of their research directions. In this respect, the real history of science might even be less important. Taking this into account, the intuition proposal might be useful and interesting enough for some disciplines like the philosophy of mind, contemporary phenomenology, the philosophy of religion, metaethics, metaphysics, the philosophy of language, and even for the philosophy of science, in which hard facts of science and its history are a serious concern—for scientists' beliefs and motivations are hard facts as such. I am leaving the question of whether the intuition proposal or the real history proposal is correct as an open question and a topic for future research.

[23] For detailed discussions of theoretical modelling, see Nagel 1961; Laudan 1977; Musgrave 1981; Godfrey-Smith 2014.

Acknowledgements

For useful discussions or comments, the author would like to acknowledge Eran Asoulin, John Bigelow, David Braddon-Mitchell, Belinda Calderone, Ruey-Lin Chen, Michael Duncan, Frank Jackson, Justine Kingsbury, Uriah Kriegel, Andrew James Latham, Erick Llamas, James Norton, Graham Oppy, Alex Sandgren, Simon Varey, Wai-hung Wong, Lei Zhong, and the philosophy of mind group at the Australian National University. The author is also grateful to Caleb Liang and Yi-Cheng Lin for their helpful administrative support to this research project. Special thanks are due to Belinda Rickard. This project is supported by a grant from the Ministry of Science and Technology of Taiwan (grant number: 108-2410-H-002-241-MY3).

References

Armstrong, D. 1978, 'Naturalism, materialism, and first philosophy', *Philosophia*, vol. 8, no. 2–3, pp. 261–76.

Braddon-Mitchell, D. & Jackson, F. 2007, *The Philosophy of Mind and Cognition*, 2nd edn, Malden: Blackwell.

Chalmers, D. 1996, *The Conscious Mind: In Search of a Fundamental Theory*, New York: Oxford University Press.

Chan, L.C. & Latham, A.J. 2019, 'Four meta-methods for the study of qualia', *Erkenntnis*, vol. 84, no. 1, pp. 145–67.

Chen, R.L. 2017, 'Mechanisms, capacities, and nomological machines: integrating Cartwright's account of nomological machines and Machamer, Darden and Craver's account of mechanisms', in H.K. Chao, S.T. Chen & J. Reiss (eds), *Philosophy of Science in Practice: Nancy Cartwright and the Nature of Scientific Reasoning*, New York: Springer.

Copp, D. 2012, 'Normativity and reasons: five arguments from Parfit against normative naturalism', in S. Nuccetelli & G. Seay (eds), *Ethical Naturalism: Current Debates*, Cambridge: Cambridge University Press, pp. 24–57.

Cramer, J. 1988, 'An overview of the transactional interpretation of quantum mechanics', *International Journal of Theoretical Physics*, vol. 27, no. 2, pp. 227–36.

Craver, C. & Tabery, J. 2015, 'Mechanisms in science', in E. Zalta (ed.), *Stanford Encyclopedia of Philosophy*. Retrieved 8 November 2020, https://plato.stanford.edu/entries/science-mechanisms/

De Caro, M. & Macarthur, D. 2010, 'Introduction: science, naturalism and the problem of normativity', in M. De Caro & D. Macarthur (eds), *Naturalism and Normativity*, New York: Columbia University Press, pp. 1–22.

Devitt, M. 1998, 'Naturalism and the a priori', *Philosophical Studies*, vol. 92, no. 1–2, pp. 45–65.

Eccles, J. 1994, *How the Self Controls Its Brain*, Berlin: Springer-Verlag.

Gallagher, S. 2018, 'Rethinking nature: phenomenology and a non-reductionist cognitive science', *Australasian Philosophical Review*, vol. 2, no. 2, pp. 125–37.

Gallagher, S., Janz, B., Reinerman-Jones, L., Bockelman, P. & Trempler, J. 2015, *A Neurophenomenology of Awe and Wonder: Towards a Non-Reductionist Cognitive Science*, London: Palgrave Macmillan.

Godfrey-Smith, P. 2014, *Philosophy of Biology*, Princeton: Princeton University Press.

Hempel, C.G. 1969, 'Reduction: ontological and linguistic facets', in S. Morgenbesser, P. Suppes & M. White (eds), *Philosophy, Science, and Method: Essays in Honor of Ernest Nagel*, New York: St. Martin's Press, pp. 179–99.

Hohwy, J. & Frith, C. 2004, 'Can neuroscience explain consciousness?', *Journal of Consciousness Studies*, vol. 11, no. 7–8, pp. 180–98.

Jackson, F. 1998, *From Metaphysics to Ethics: A Defence of Conceptual Analysis*, Oxford: Oxford University Press.

Kim, J. 2003, 'The American origins of philosophical naturalism', *Journal of Philosophical Research*, vol. 28, pp. 83–98.

Kim, J. 2006, *Philosophy of Mind*, 2nd edn, Boulder: Westview Press.

Kuhn, T. 1996, *The Structure of Scientific Revolutions*, 3rd edn, Chicago: University of Chicago Press.

Laudan, L. 1977, *Progress and Its Problems: Towards a Theory of Scientific Growth*, London: Routledge & Kegan Paul.

Lewis, D. 1972, 'Psychophysical and theoretical identifications', *Australasian Journal of Philosophy*, vol. 50, no. 3, pp. 249–58.

Lewis, D. 1983, 'New work for a theory of universals', *Australasian Journal of Philosophy*, vol. 61, no. 4, pp. 343–77.

Lewis, D. 2009, 'Ramseyan humility', in D. Braddon-Mitchell & R. Nola (eds), *Conceptual Analysis and Philosophical Naturalism*, Cambridge, MA: MIT Press, pp. 203–22.

McDowell, J. 1994, *Mind and World*, Cambridge, MA: Harvard University Press.

Machamer, K., Darden, L. & Craver, C. 2000, 'Thinking about mechanisms', *Philosophy of Science*, vol. 67, no. 1, pp. 1–25.

McPherson, T. 2015, 'What is at stake in debates among normative realists?', *Noûs*, vol. 49, no. 1, pp. 123–46.

Melnyk, A. 2003, *A Physicalist Manifesto: Thoroughly Modern Materialism*, Cambridge: Cambridge University Press.

Montero, B. 2009, 'What is the physical?', in A. Beckermann & B.P. McLaughlin (eds), *The Oxford Handbook of Philosophy of Mind*, New York: Oxford University Press, pp. 173–88.

Montero, B. & Papineau, D. 2016, 'Naturalism and physicalism', in K. Clark (ed.), *The Blackwell Companion to Naturalism*, Malden: Wiley-Blackwell, pp. 182–95.

Musgrave, A. 1981, '"Unreal assumptions" in economic theories: the F-twist untwist', *Kyklos*, vol. 34, no. 3, pp. 377–87.

Nagel, E. 1961, *The Structure of Science: Problems in the Logic of Scientific Explanation*, London: Routledge.

Oppy, G. 2018a, *Naturalism and Religion*, Abingdon: Routledge.

Oppy, G. 2018b, Atheism and *A*gnosticism, Cambridge: Cambridge University Press.

Papineau, D. 2014, 'The poverty of conceptual analysis', in M.C. Haug (ed.), *Philosophical Methodology: The Armchair or the Laboratory?*, Abingdon: Routledge, pp. 166–94.

Pollard, C. 2014, 'Merleau-Ponty and embodied cognitive science', *Discipline Filosofiche*, vol. 24, no. 2, pp. 67–90.

Price, H. 1996, *Time's Arrow and Archimedes' Point: New Directions for the Physics of Time*, Oxford: Oxford University Press.

Quine, W.V.O. 1995, 'Naturalism; or, living within one's means', *Dialectica*, vol. 49, pp. 251–63.

Ravenscroft, I. 1997, 'Physical properties', *Southern Journal of Philosophy*, vol. 35, no. 3, pp. 419–43.

Stoljar, D. 2010, *Physicalism*, New York: Routledge.

Tononi, G. & Koch, C. 2015, 'Consciousness: here, there and everywhere?', *Philosophical Transactions of the Royal Society B Biological Sciences*, vol. 370, no. 1668. https://royalsocietypublishing.org/doi/10.1098/rstb.2014.0167.

Varela, F., Thompson, E. & Rosch, E. 1991, *The Embodied Mind*, Cambridge, MA: MIT Press.

Zhong, L. 2016, 'Physicalism, psychism, and phenomenalism', *The Journal of Philosophy*, vol. 113, pp. 572–90.