



The Richmond Journal of **Philosophy**

Issue Eight Winter 2004

In this  
***issue***...

**Dan O'Brien**  
on family life

**James Hill**  
on descartes

**Mark Sinclair**  
on history

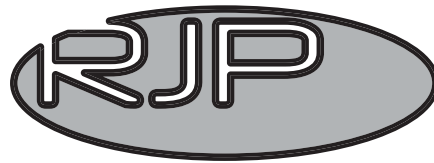
**Paul Sperring**  
on physicalism

**Peter Wyss**  
on emergence

**Sam Coleman**  
reply to Peter Wyss



Richmond upon Thames College



# **The Richmond Journal of Philosophy**

**Issue eight  
Winter 2004**

## **Editorial Board**

Stephen Grant

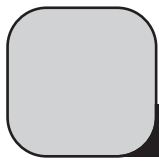
Paul Sheehy

Paul Sperring

Philosophy Department  
Richmond upon Thames College  
Egerton Road  
Twickenham  
Middlesex  
TW2 7SJ  
United Kingdom

email: [rjp@rutc.ac.uk](mailto:rjp@rutc.ac.uk)

[www.rutc.ac.uk/rjp](http://www.rutc.ac.uk/rjp)



## Contents

Editorial	p4
About the Editorial Board	p5
Family Life <i>Dan O'Brien</i>	p6
Descartes' Dreaming Argument and why we might be sceptical of it <i>James Hill</i>	p11
Nietzsche and the Problem of History <i>Mark Sinclair</i>	p18
Should we give up on Reductive Physicalism? <i>Paul Sperring</i>	p24
Two Aspects of Emergence <i>Peter Wyss</i>	p32
Reply to Peter Wyss <i>Sam Coleman</i>	p40
Notes on Contributors	p46
Notes for Contributors	p47
How to Subscribe	p49

# [Editorial]

Welcome to the eighth issue of the Richmond Journal of Philosophy.

We begin with Dan O'Brien's discussion on how we come to grasp certain concepts, and in particular on the influence of Wittgenstein. Next James Hill considers Descartes' dreaming argument and questions whether we should be sceptical of its claims. The third article by Mark Sinclair discusses Nietzsche and the sense in which history poses a philosophical problem. The remaining papers in this issue share a concern over our understanding of the mental. Paul Sperring asks whether we should abandon attempts to reductively analyse the mind in physical terms, while Peter Wyss examines the notion of emergence and the role it can play in the philosophy of mind. Our final paper is a critical response by Sam Coleman to Peter's defence of emergentism.

## Purpose of the Journal

The motivation for and ambition of the journal is to provide serious philosophy for students who are at an early stage in their philosophical studies. The style and content of the papers will be accessible to students who have yet to become hardened to the more technical and specialised journals of professional philosophy.

What do we mean by 'serious' philosophy? First, the content of the journal is not constrained by a remit to appeal to or reach the interested general public. Whilst the papers must speak to the needs of students who are relatively inexperienced in philosophy, they presuppose that their audience is actively engaged in philosophy. Second, the content is serious in its focus on the central areas of philosophy. One must beware of the dangers of trying to impose more precision on a subject than its nature will allow. Therefore, some degree of caution is called for in talking of the central areas of philosophy. Nonetheless, the big or traditional questions of metaphysics, epistemology, and ethics will provide the journal's centre of gravity. The third way in which the philosophy is serious is through the scope, variety and depth of analysis that can be achieved by the accumulation of papers over time. Moreover, each

paper is not simply an introduction to one of the main topics on A-level, IB or degree courses. Such papers will indeed have a role in the journal, but they will not be the only kind. Our contributors will be offering original papers based on their own research. The journal will be a forum for the kind of critical engagement and debate that characterise the practice of philosophy. The fourth way in which the philosophy is serious is in the contributors themselves. The vast bulk of the papers will be written by professional philosophers engaged in both research and teaching.

# About the [Editorial] Board

Stephen Grant is a full-time lecturer in philosophy at Richmond upon Thames College. He has also taught at King's College London where he is completing his doctorate on the emotions. His main interests are in the emotions, ethics and political philosophy. He has published on the ontological argument.

Dr Paul Sheehy teaches philosophy at Richmond upon Thames College and King's College London. His main areas of interest are in metaphysics, political and moral philosophy and the philosophy of the social sciences. His doctoral thesis was on the ontological and moral status of social groups, and he has published papers on social groups, voting and explanation and realism.

Paul Sperring is head of the philosophy department at Richmond upon Thames College and an A-level examiner in philosophy. He completed his undergraduate and masters studies at The University of Warwick, studying both analytic and continental philosophy. He is currently working towards his PhD at Birkbeck College. His research interests are metaphysics and the philosophy of mind.

# [Editorial]

## Dan O'Brien

# [Family]

## Life

In Plato's dialogues, Socrates pushes his companions hard to find the definitions of some of our ethical concepts such as justice, goodness and piety. Surely a concept must have such a definition if it is to have a determinate application or meaning. We can call this the Socratic assumption. I shall suggest, however, that this assumption is not always warranted, and I shall sketch Wittgenstein's alternative account of our grasp of certain concepts. The example that I shall give of a concept that does not have a definition may perhaps be surprising given its scientific nature; the concept is that of 'life', or of what it is to be 'living'. First, though, let us look a little closer at the content of the Socratic assumption.

### Philosophical Analysis

Socrates was engaged in philosophical analysis, and this is an activity that has to a greater or lesser extent engaged philosophers ever since. When pursuing such analysis we attempt to define the extension of our concepts, that is, we attempt to draw up rules that specify to what it is our concepts apply. One way of doing this is to lay down necessary and sufficient conditions for the application of a concept. For example, to be Premiership Champions it is necessary that one is a registered football club and that one has a squad of at least

eleven players. The sufficient conditions for being so are that such a team must have more points than anybody else. Note, though, that different properties can meet the requirements of the sufficient conditions. To be multilingual, say, it is sufficient to speak French, German and Greek, but other sets of languages will suffice.

Analyses of many concepts are easy to find. One may wonder just what a carburettor is, and in order to work this out we can analyse how we use the term 'carburettor' and thereby discover its definition. In doing so, we can work out that the term refers to the mechanism within an engine that mixes together air and petrol in order that the latter will efficiently combust. This, then, is what a carburettor is. Other analyses, however, are not so forthcoming. Socrates' concern with the nature of justice, say, is something

that continues to tax moral philosophers to this day.

A useful philosophical tool for pursuing such analyses is the thought experiment. One can consider how one would apply one's concepts in imaginary or counterfactual situations. And, the intuitions we have about such cases can be used to test our suggested analyses. In order to illustrate this let us consider a simple case. It may be suggested that 'water' refers to whatever liquid it happens to be that falls from the clouds and that fills the oceans. This, then, is an analysis of the concept 'water'. Let us, however, consider an alien world (or a distinct way in which our world could have developed). In such a world the rain and the oceans may consist of ammonia. Now, though, we encounter a problem. Our analysis suggests that 'water' should refer to this substance, but our intuition is that this is simply different stuff; it is not water. We must, then, re-think our suggested analysis; water cannot be defined as whatever liquid it is that happens to fall from the clouds and fills the oceans. If in a particular case the verdict of our analysis does not match our intuitions about the correct use of our concepts, then we must refine our analysis or reject it completely and start again from scratch. We shall see that this is the kind of procedure that we shall use below to assess a suggested analysis of the concept of 'life'.



## Life

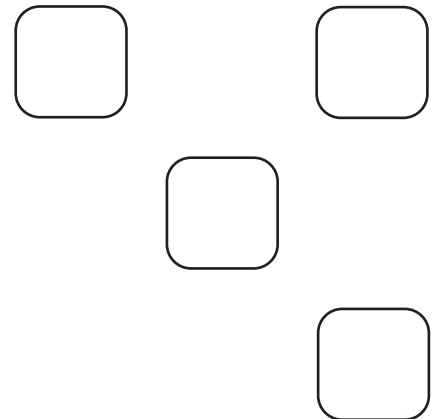
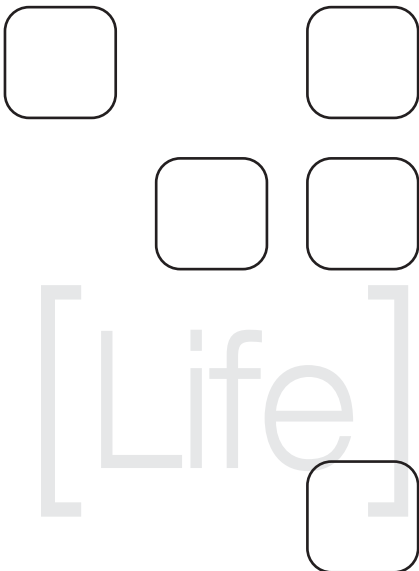
There are many clear-cut applications of 'life': sharks are alive and stones are not. But what is it that enables us to apply such a classification? Or, in accepting the Socratic assumption, what definition of 'life' are we using? It was once thought that living organisms contained a certain 'vital spirit', a substance that sparked sharks and not stones into life. Science, however, has found that such a substance does not exist, and accounts of life now focus on the functional or organisational level rather than on the possession of a particular type of substance. To be alive one does not require vital spirits; instead, one needs to perform certain actions or functions. And biologists have detailed just which functional characteristics are necessary and sufficient for life. Living creatures must eat, grow, excrete, move, respire, reproduce, and be sensitive to their environment. It is this definition of life that may be familiar from science classes at school and it is this definition that we shall go on to investigate.

First, we have to be a little careful about how we apply these seven criteria because straightaway we seem to find counterexamples, cases where our intuitions clash with our definition. Mules do not reproduce, my sweet peas do not move very far, and the cactus on my desk does not seem to eat anything, yet we should like to say that all three of these organisms are alive. There are, however, senses in which these organisms do reproduce, move and eat: the body cells of a mule are constantly replicating and being renewed; sweet peas turn their leaves toward the sun and grow towards the light; and, my cactus is taking in carbon dioxide from the environment and metabolising it into sugar which it uses as fuel. In order, then, for the above definition to be adopted we must not think of these criteria in anthropomorphic terms, and the biologist will need to tell us how we should conceive of such functions as reproduction and movement.

Let us next, however, consider some more exotic creatures. We shall conceive of certain organisms that may force us to reject the above definition. Let us imagine a creature that simply has an ongoing metabolism that produces coloured compounds, compounds that help camouflage this creature. We shall call this sea-dwelling creature a 'camouflaged metaboliser'. It is so efficient at this that its continued survival is assured. The metabolites of this creature are internally recycled and thus, there is no need for nutrition or excretion. Its cell structure is also so durable that reproduction is unnecessary, and since it need not eat or reproduce, it has no need for self-directed movement. Its coloured camouflage compounds, however, do degrade over time and therefore

ongoing metabolic recycling is required in order to maintain such compounds at the required level for survival. The energy necessary for such recycling is provided by the respiratory processes of the camouflaged metaboliser (the biological conception of respiration being that of a process of oxidation in which energy is transferred from certain energy rich metabolites to those that are more easily utilised by the organism). This is a creature, then, that only respire. And, my claim is that this is a creature that we would want to call alive even though it is not one that is picked out by the suggested definition above.

One response here would be to rethink our analysis. One could, perhaps, focus on the metabolism of creatures and claim that organisms are alive if they have respiration-driven metabolic processes of a certain complex kind. Creatures could then be seen as alive if they satisfied only certain important criteria of the original suggested seven. I, however, shall suggest an alternative response, one based on Wittgenstein's notion of family resemblance, and one that allows distinct living organisms to have no characteristics in common.



## Family Resemblance

Wittgenstein argues that there are no necessary and sufficient conditions for the application of at least some of our concepts; we should not assume that the various instantiations of a concept have anything in common. When we look at the use of some of our concepts we do not find such common features. His example is that of the concept 'game'.

Consider for example the proceedings that we call 'games'. I mean board-games, card-games, ball-games, Olympic games, and so on. What is common to them all? - Don't say: 'There *must* be something common, or they would not be called 'games' - but *look and see* whether there is anything common to all. (*Philosophical Investigations* §66)

And, if one looks, one does not find any common features.

Look for example at board-games, with their multifarious relationships. Now pass to card-games; here you find many correspondences with the first group, but many common features drop out, and others appear. When we pass next to ball-games, much that is common is retained, but much is lost. - Are they all 'amusing'? Compare chess with noughts and crosses. Or is there always winning and losing, or competition between players? Think of patience ...(*Philosophical Investigations* §66)

Wittgenstein goes on, and so can we: in looking at the various activities we call games we can see that there is no such thing as the essence of what it is to be a game. All we find is:

a complicated network of similarities overlapping and criss-crossing: sometimes overall similarities, sometimes similarities of detail. (*Philosophical Investigations* §66)

And:

I can think of no better expression to characterise these similarities than 'family resemblances'; for the various resemblances between members of a family: build, features, colour of the eyes, gait, temperament, etc. etc. overlap and criss-cross in the same way. - And I shall say: 'games' form a family. (*Philosophical Investigations* §67)

One can illustrate such a family as follows. We can see how it would be plausible to say that the five individuals below are of the same kind or family. Among them they have the properties A, B, C, D and E; and, any two individuals share three properties. There are, however, no properties that are shared by all.

X<sub>1</sub>: A B C D

X<sub>2</sub>: A B C E

X<sub>3</sub>: A B E D

X<sub>4</sub>: A E C D

X<sub>5</sub>: E B C D

Wittgenstein goes on to draw a further analogy, this time between the members of a family and the fibres of a rope, and this we shall see suggests a more radical construal of family resemblance.

And we extend our concept of [X] as in spinning a thread we twist fibre on fibre. And the strength of the fibre does not reside in the fact that some one fibre runs through its whole length, but in the overlapping of many fibres. (*Philosophical Investigations* §67)

As we saw above, X<sub>1</sub> and X<sub>2</sub> are of the same kind because they share three properties. Let us now, though, introduce X<sub>6</sub> (B C D E). X<sub>6</sub> also shares three properties with X<sub>1</sub> and thus should be seen as of the same kind. We can, therefore, imagine a progression of individuals as follows, each of which shares three properties with his nearest 'relations'.

X<sub>1</sub>: A B C D

X<sub>6</sub>: B C D E

X<sub>7</sub>: C D E F

X<sub>8</sub>: D E F G

X<sub>9</sub>: E F G H

We should, then, also like to say that these five individuals comprise a family, or that they are of the same kind. One should note, however, that we are now accepting that X<sub>8</sub> is of the same kind as X<sub>1</sub> even though they have no properties in common, just as two widely separated sections of the same rope may have no fibres in common. And, it is important to see that X<sub>8</sub> is only considered to be of the same kind as X<sub>1</sub> because of the existence of the rest of the family, a family that we have come to call by the same name. There may be another individual, X<sub>99</sub>, who possesses the properties R S T U. This individual is just as different from X<sub>1</sub> as X<sub>8</sub> is - both X<sub>8</sub> and X<sub>99</sub> have nothing in common with X<sub>1</sub> - yet only X<sub>8</sub> is of the same kind as X<sub>1</sub> because of the family of intermediates (X<sub>2</sub> to X<sub>7</sub>) through which they are related. Two activities, therefore, can both be games even though they have nothing in common, and, I shall argue, two entities can both be alive even though they have no shared properties. Understanding the notion of family resemblance should quell our 'craving for generality' (*The Blue Book*, p. 17).



## Why we might be sceptical of it

There are paradigm cases of life that possess all seven of the characteristics above. I have also argued that there could be living creatures that do not possess all seven, and that we would be prepared to call the camouflaged metaboliser alive even though it only respire. I shall now claim that in a world full of such metabolisers we can come to see other markedly similar systems as living, yet these creatures may have nothing in common with our camouflaged metabolisers.

In a world inhabited by camouflaged metabolisers we would also come to say that mock-metaboliser is alive. Mock-metaboliser has the same outwardly camouflaged appearance as our camouflaged metabolisers but it does not respire in order to maintain its metabolism. It does not have a metabolism. Its durable cell structure simply provides a safe haven for certain coloured compounds, compounds that would otherwise be broken down by the acidic nature of the alien seas. Such coloured compounds diffuse in and remain there shielding the structure from the

eyes of predators. As its load of such compounds increases so does its density, and thus it sinks lower into the seas where it so happens that its colouring no longer acts as camouflage. Down in the depths certain deep-dwellers start to nibble away at these now conspicuous compounds, thus decreasing its density and allowing this structure to float upwards into safety and to restock 'itself' with the coloured compounds it 'requires'. This structure can thus be seen as pursuing a strategy to prolong its existence in that it periodically rises and sinks in the ocean in a constant 'effort' to 'avoid' its prey. In a world of mules, sharks and sweet peas, such a structure may be seen as merely an interesting object of the seas. However, if such forms of alien life as our camouflaged metaboliser had previously helped clarify our conception of life then the mock-metaboliser could also be seen as living.

There may be other objects that we do not consider to be alive, yet they may be just as different from sharks as our mock-metabolisers are. It is the existence of the family member camouflaged metaboliser, however, that allows mock-metabolisers into the fold and not the others (just as  $X_8$  was allowed in but not  $X_{99}$ ). My claim, then, is that one can conceive of 'alive' as a family resemblance concept.

This entails that we will not be able to find a definition or a set of necessary and sufficient conditions for the application of the concept 'life'. (Note that the suggestion to pursue an analysis focused on the important process of respiration is now ruled out because mock-metaboliser does not respire.)



[Life]

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

## Biology

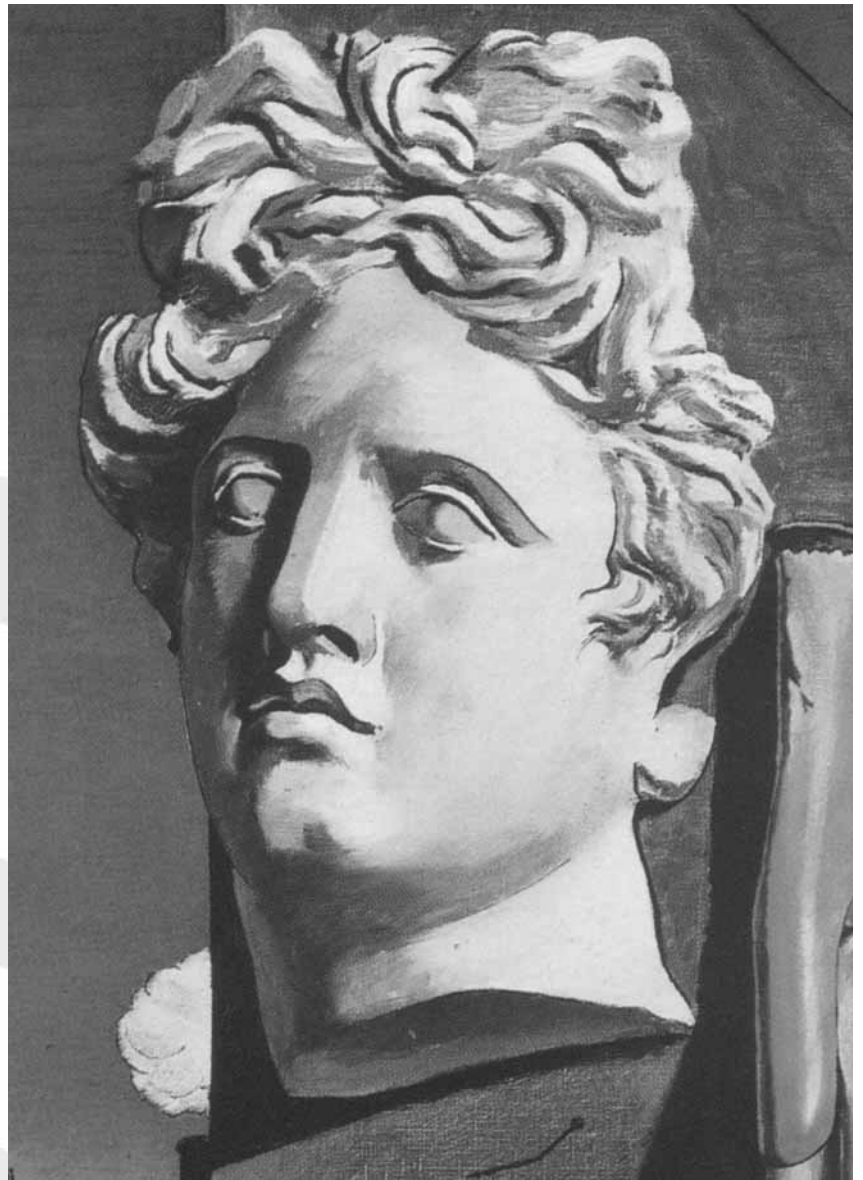
The biologist, then, cannot provide a definition that captures all instantiations of the concept 'life'. Nevertheless, the concept is perfectly meaningful; it is simply that the search for a definition is unnecessary. '[T]his does not [however] leave us with nothing to do; instead it invites us to trace out relationships, and this should be done with whatever degree of rigor the subject matter allows' (Wittgenstein, p. 138). Exploration can always throw up distinct manifestations of life and the biologist's task is then to catalogue the various properties that we see shared among our growing family of living organisms and trace the relations between them. She can also, of course, go on to investigate these properties and answer questions concerning how it is that certain family members move and what it is that enables some of them to respire. But importantly, the properties she investigates will not be those that are shared by all living creatures, but rather, they are those that ground the application of our concept in particular cases.

The main aim of this paper has been to illuminate the concept of family resemblance, and to show how in certain cases it can undercut the pursuit of philosophical analysis and the Socratic Assumption. In doing so, we have also come to reflect on the ontology of biology. Biological entities are those that are alive; that is the criterion that separates the subject matter of biology from that of the other physical sciences such as chemistry. We concluded that the biologist should not aim to give a unifying theory of what it is to be alive, but that she should instead simply provide a catalogue of living

creatures and note the various features that underlie her classification.

## References

- R. Fogelin, *Wittgenstein* (London: Routledge and Kegan Paul, 1976)
- L. Wittgenstein, *The Blue Book in The Blue and Brown Books* (Oxford: Blackwell, 1958)
- L. Wittgenstein, *Philosophical Investigations* (Oxford: Blackwell, 1958)



**James Hill**

# Descartes' dreaming [argument] and why we might be sceptical of it

The sceptical argument about dreaming that Descartes puts forward in the First Meditation is a landmark in philosophy. It raises a problem that seems to be both of the utmost simplicity and to be devastating in its impact on our claims to know things about the world around us. In a single paragraph Descartes describes how our confidence that we are awake may be shaken:

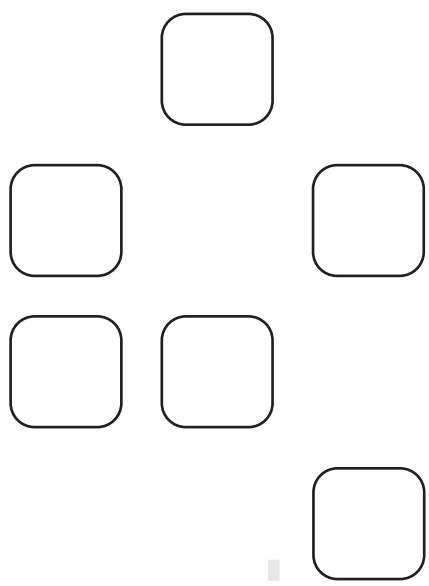
As if I were not a man who sleeps at night, and regularly has all the same experiences while asleep as madmen do when awake indeed sometimes even more improbable ones. How often, asleep at night, am I convinced of just such familiar events that I am here in my dressing-gown, sitting by the fire when in fact I am lying undressed in bed! Yet at the moment my eyes are certainly wide awake when I look at this piece of paper; I shake my head and it is not asleep; as I stretch out and feel my hand I do so deliberately, and I know what I am doing. All this would not happen with such distinctness to someone asleep. Indeed! As if I did not remember other occasions when I have been tricked by exactly similar thoughts while asleep!

As I think about this more carefully, I see plainly that there are never any sure signs by means of which being awake can be distinguished from being asleep. The result is that I begin to feel dazed, and this very feeling only reinforces the notion that I may be asleep.<sup>1</sup>

Our concern here is to make clear what exactly the argument in this paragraph presupposes and what it purports to show. I shall also suggest a way in which we might avoid the sceptical conclusions that seem to arise so irresistibly from the thought that one might, at any given moment, be dreaming.



To begin with we must see the dreaming argument in the context of the *Meditations* themselves. They are an unconventional philosophical text. Most strikingly, they are written in the first person, recording the reflections of a solitary thinker. But the *Meditations* are not autobiography: we know that Descartes himself did not literally go through the process described in arriving at his philosophical views. We are really dealing with a piece of fiction, and the narrator, the 'I' of the *Meditations*, is a fictional character. He introduces himself in the sketchiest of terms as a person of a reasonably mature age who is concerned with making progress in science. We also know that the meditator finds himself in retirement from the world and thus able to devote himself to a concentrated period of meditation. Beyond that, we know precious little. This paucity of detail is no accident. Descartes wants the 'I' of the *Meditations* to stand for *any* thinker setting out in a quest for certainty: the *Meditations* represents an unfolding viewpoint that we are all invited to share. For this to be possible we must be able to fully identify with the narrator and that means the fewer distinguishing features the better.



It is because the narrator is a kind of philosophical everyman, that Descartes has him dismiss the possibility that he is undergoing delusions similar to those in the insane when the thought suggests itself.

Unless perhaps I were to liken myself to madmen, whose brains are so damaged by the persistent vapours of melancholia that they firmly maintain they are kings when they are paupers, or say they are dressed in purple when they are naked, or that their heads are made of earthenware, or that they are pumpkins, or made of glass. But such people are insane, and I would be thought equally mad if I took anything from them as a model for myself.<sup>2</sup>

Few readers would be able to identify with a narrator who thought he might be insane. So, in the prosecution of his method of doubt, Descartes prefers to look to another consideration, the possibility that he is dreaming. Dreams serve Descartes' purposes well because they have the following double-aspect: they are on the one hand experiences of the ordinary, healthy mind, experiences with which no one could plausibly deny an acquaintance, and yet, on the other hand, they are comparable, in their extravagance and deceptiveness, to the delusions of insanity the narrator has just been considering.<sup>3</sup>

The key move that Descartes then makes is to highlight the lack of insight one has into one's condition when dreaming. It is this lack of insight, and Descartes' way of interpreting it, which forms the backbone of the dreaming argument. Descartes treats the lack of insight as the result of sense-perception and dreams having—or potentially having—the same content.

... I see plainly that there are never any sure signs by means of which being awake can be distinguished from being asleep. The result is that I begin to feel dazed, and this very feeling only reinforces the notion that I may be asleep.

The very situation in which the narrator finds himself right now, sitting by the fire in a dressing-gown with a piece of paper in his hand, might really be a scene in a dream he is having while fast asleep in bed. This in turn means that the meditator can no longer be sure of the existence of the things that he is perceiving, and thus he has a reason for a more general doubt about the overall veracity of his sense experience. The argument and conclusion are retrospectively summed up by the narrator in the Sixth Meditation thus:

... every thought I was having while awake I can also think of myself as sometimes having while asleep; and since I do not believe that what I seem to perceive in sleep comes from things located outside me, I did not see why I should be any more inclined to believe this of what I think I perceive while awake.

II

Let us return to the premise that there are 'no sure signs by means of which being awake can be distinguished from being asleep'. This premise has been repeatedly challenged in the long history of responses to Descartes' sceptical thought. I shall now describe three such challenges: those of Hobbes, Locke and Austin.

Thomas Hobbes thought that one distinguishing mark of dreams was the absence of a sense of the absurd. In dreaming, he claimed, we take in all

kinds of bizarre happenings without batting an eyelid. It does not occur to us that what we are experiencing is so crazy that it can only be an illusion. In waking life, on the other hand, our sense of the absurd is perfectly alive and so, looking back, we readily appreciate the absurdity of what we have dreamt. Hobbes can thus conclude that

... because waking I often observe the absurdity of Dreames, but never dream of the absurdities of my waking Thoughts; I am well satisfied, that being awake, I know I dreame not; though when I dreame, I think my selfe awake.<sup>4</sup>

Hobbes' view of dreams and waking involves an interesting asymmetry. Even if we do not know when we are dreaming, we may still know when we are awake that we are awake.

John Locke in his *Essay concerning Human Understanding* is less patient with the suggestion that he might be dreaming. But he does offer some reasons for his impatience, inviting his opponent 'to dream that I make him this Answer'. The central mark that distinguishes dreams from waking experience in Locke's view is the pleasure and pain experienced when awake but not found in dreams. His point is best made with acute pain:

I believe he [someone like Descartes] will allow a very manifest difference between dreaming of being in the Fire, and being actually in it.<sup>5</sup>

It is important to emphasise that Locke's example is one of *physical* pain, as nightmares involving *mental* pain, such as anxiety, shame or fear, are commonplace. Locke is saying that the actual physical sensation of pain, at least in its more extreme forms, is



The meditator cannot imagine himself to be deceived when reflecting on these truths which include the most straightforward statements in 'arithmetic, geometry and other subjects of this kind'. The intellect maintains its ability to discern transparent truths even when caught up in a dream and, as we shall see, this makes possible the metaphysical reflections of the following meditations.

### III

We have outlined two aspects of Descartes' concept of dream. Firstly its inexhaustible potential to mimic reality and secondly its failure to affect the clearest thoughts of the intellect. When we put these two things together we can characterise the concept of dreaming that Descartes presupposes in the *Meditations*. Most importantly, a dream for Descartes is a matter solely of the imagination. Because the imagination is treated as having the same characteristics and limits as sense-experience, the two are potentially indistinguishable.

If intellectual abilities can be contaminated by dreaming, that is presumably only because they involve imagery or memory, as when the complexity of the subject-matter demands that we faithfully remember a sequence of steps: the intellect when concentrating on one thing—a so-called 'simple nature'—cannot fail to be correct in its judgement.

Once we appreciate how Descartes treats dreams as concerns of the imagination alone we may also appreciate why the narrator, despite the initial shock he feels, brought on by his discovery, is ultimately so unfazed by the dreaming argument

itself. After all, the narrator produces his proof of why he is not dreaming only in the final paragraph of the book. And more than once before then the narrator tells us that the hypothesis is unthreatening to his philosophical progress. At the end of the Fifth Meditation, for example, when he has convinced himself by deductive proof that God exists and that everything else depends upon his existence, the narrator poses the question of whether he is not dreaming the proof and the conclusion:

Can one raise the objection I put to myself a while ago, that I may be dreaming, or that everything which I am now thinking has as little truth as what comes to the mind of one who is asleep? Yet even this does not change anything. For even though I might be dreaming, if there is anything which is evident to my intellect then it is wholly true.<sup>8</sup>

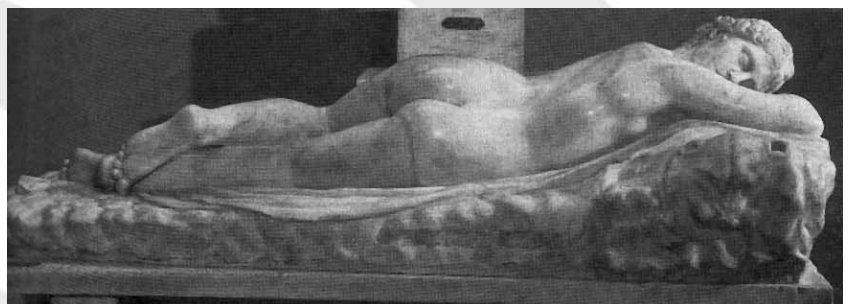
Thus the intellectual progress of the *Meditations* is assumed to be sound, even if the meditator is fast asleep. This is an important point. It means that dreams are not states of the whole subject. The subject can remain lucid if he or she ignores the deliverances of sense-perception/imagination. It also reveals Descartes' thinking about the lack of insight in a dream. The subject has no insight not because he or she is thoroughly confused or because the

ability to reflect is incapacitated, but because the data coming from the senses and the imagination are indistinguishable: the dream appears to be just like sense-perception.

### IV

Now that we have discerned the features of the concept of dreaming presupposed in the *Meditations* we can, I hope, begin to see the distance between this concept and our nightly experience of the dreamworld. It was precisely because dreaming is an experience that everyone is familiar with that Descartes employed it in the meditator's sceptical strategy. The hope was that there is a natural conception of dreaming in common life, the implications of which can be exploited to bring our most basic perceptions into doubt. But in the course of the argument we have found that the conception of dreaming appealed to is not quite uncontroversial. For the argument to go through as Descartes would like it to, dreaming must involve an inexhaustible capacity of the imagination to mimic experience of reality, on the one hand, but no power to upset the basic functions of the intellect in mathematics and logic, and in philosophical reflection, on the other hand.<sup>9</sup>

But to systematically challenge the conception of dreaming that Descartes presupposes is not at all easy. When



we seek a characterisation of what dreaming is we are thrown back on our own radically subjective experience. Dreams are supremely elusive. There are real problems, for example, with the reliability of memory. Memory of experiences in waking life can be confirmed by the memory of others and by the current state of physical things. I can be sure that I fed the cat, for example, because my wife remembers that I did so and because right now the cat is lying next to me purring, not begging for food. But can I be sure I had a dream in which I fed the cat? There is no way of checking whether my memory of doing so is not just an illusion. Might it not, for example, have appeared in my mind on waking?<sup>10</sup>

The problems with memory are compounded by the fact that dreams have a peculiar phenomenology which is often almost impossible to put into words. When reflecting on dreams we are always in danger of reconstructing them in terms of the everyday perception of objects and people.

Their strangeness and otherness is perhaps impossible to capture. What is garbled and chaotic in the dreamworld is translated into the clear and tidy terms of everyday experience. It is not even clear that the narrative structure we tell dreams in was really always there in the first place.

But the assumptions that Descartes makes about dreaming may at least be compared to what people can remember and relate of their actual dreams. Let us begin with the second assumption, the one that treats dreams as not affecting the reliability of the most transparent truths known to the intellect. This is a bold claim that was picked up on by Father Bourdin, author of the *Seventh Objections to the Meditations*, who gives us a very believable episode from a dream in which even the dreamer's ability to count is scrambled.

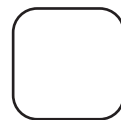
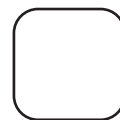
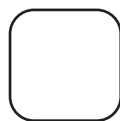
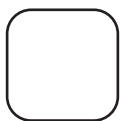
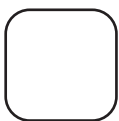
I know a man who once, when falling asleep, heard the clock strike four, and counted the strokes as 'one, one, one, one'. It then seemed to him that there was something absurd about this, and he shouted out: 'That clock must be going mad; it has struck one o'clock four times!' Is there really anything so absurd or irrational that it could not come into the mind of someone who is asleep or raving?<sup>11</sup>

Bourdin here, by appealing to his friend's dream, seems ready to treat dreams as states of the whole subject. Dreams are not restricted to the image-making capacities of the imagination. They have the potential to distort all our thinking and thus deprive us of any ability to follow the dictates of rationality. Bourdin compares dreams to 'raving'—presumably states of delirium—say in fever.

We might also compare them to various forms of intoxication which have a similar general reach, confusing our mental capacities right across the board, including the intellect (in Descartes' sense). A very drunk person, for example, might not be trusted in counting even the smallest numbers (especially if he is seeing double). Dreams, then, like raving and drunkenness, are best treated as states of the whole subject.

More important though for the dreaming argument is the first assumption we mentioned, the one that grants dreams the potential to replicate any experience that we might have when waking. Descartes seems to assume that if a perceptual situation is not actually logically impossible—if it respects the principle of non-contradiction—then it may be dreamt. Now clearly there is a lot that is weird and wonderful in the world of dreams. The versatility of the imagination here is hard to put any sort of limit on. What is less clear, though, is whether dreams ever replicate certain kinds of *everyday* experiences. Take Descartes own example:

How often, asleep at night, am I convinced of just such familiar events—that I am here in my dressing-gown, sitting by the fire—when in fact I am lying undressed in bed!



[Dreaming]

This exclamation is, on reflection, a peculiar one. Do we really dream that we are doing such familiar things as sitting by the fire writing? Do we really have such unextraordinary and, well, boring dreams? Many people who have examined dreaming itself, rather than the putative sceptical implications of dreaming, would say no. Here we will take one example, Havelock Ellis, a psychologist of the early twentieth century, who spent twenty years studying his own dreams, writing them down 'directly they have occurred, usually on awakening in the morning'. For Ellis 'the most elementary fact about dream vision is the perpetual and unceasing change which it is undergoing at every moment'. He used the term 'kaleidoscopic' for this quality of dreams. He believed that our freshest memories of dreams reveal their kaleidoscopic character, unlike the stylised stories that we subsequently construct when telling a dream.<sup>12</sup>

Would such kaleidoscopic experiences be able to replicate the situation that Descartes' narrator finds himself in? Hardly. The narrator is sitting by the fire in a dressing gown, holding a piece of paper in his hand. The situation is a static one, it involves perception of persisting physical objects, and those objects remain as they are even under careful examination. This experience is as far from being kaleidoscopic as any experience could be and therefore it contradicts what is, in Ellis' view, 'the most elementary fact about dream vision'. In addition there is concentrated reflection taking place and this requires a parallel steadiness at the conceptual level. To think about dreaming at all, for example, requires a continuing grasp of what a dream is.

If Ellis is fundamentally on the right lines in his empirical description of dreams, then dreams involve confused

thinking, which is constantly evolving and changing in curious ways. They are fluid and unstable. They do not present such everyday experiences as sitting calmly by the fire and surveying one's surroundings in a state of reflection. Therefore the contents of dreams are indeed different to waking states and there *is* a sure sign that can distinguish the waking state from dreaming.

It might now seem that we are now putting forward a new suggestion in the spirit of Hobbes, Locke and Austin. But there is a difference. In the view of those three philosophers, there is a difference in the content of dreams and waking experience. This different content can be sought and when found treated as a sure sign that one is asleep or awake. The present suggestion goes further. It says that not only are the contents of dreams and waking life different, but that the subject is unable to appreciate this difference in his dreaming state. The subject has no unconfused faculty that might detect the difference. In other words, dreams and waking life are distinguishable in content, but the subject in a dream is not in a state to distinguish them.

Why does the subject lack insight into his dreaming condition? Not because dreams and waking experience have the same content, as Descartes thought, but precisely because of the distinguishing feature of the dreamlike state, its kaleidoscopic character, precludes insight. If the contents of dreams are always evolving and changing there is no room for ordered rational reflection, including the reflection that would provide insight into the fact that one is at present dreaming.<sup>13</sup> This also explains why, *contra* Descartes, dreams are typically states of the *whole* subject, affecting and confusing our ability to use the

intellect in maths and logic and to reflect philosophically.

Let us return to Descartes' meditator. You will remember that he quickly became convinced that he could be dreaming because he treated the lack of insight in dreams as being the result of the lack of sure signs to distinguish dreams from reality. In other words the lack of insight was supposed to be the result of the contents of dreams and waking experience being (potentially at least) indistinguishable. We, on the other hand, are saying that the lack of insight is a result of the very *difference* in content between dreams and waking experience. Dreams are a flow of constantly shifting and transient impressions. Such a content cripples one's capacity for reflection on one's current state and therefore one experiences dreams without the knowledge that one is dreaming. We are therefore left agreeing with the asymmetry of insight pointed to by Hobbes'. The fact that I am not aware I am dreaming during a dream does not mean that now, in my reflective state, I may not be aware I am awake.

We have made a case for one way in which Descartes' sweeping argument might be resisted. But the strength and weakness of our argument lies in its empirical status. This is a strength when we compare it with Descartes' conception of dreaming that makes the empirically doubtful assumptions that dreams are able to replicate *any* waking experience and yet are not states of the whole subject. But it is also a weakness. Empirical characterisations are contingent and reversible. In the case of dreams this is especially a problem as data that are so hard to determine, when determined, may always rebel.

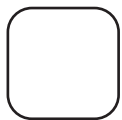
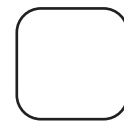
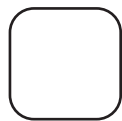
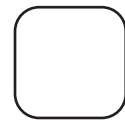
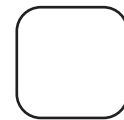
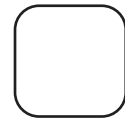
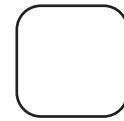
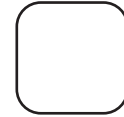
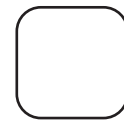




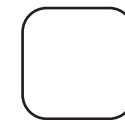
13 This is not to deny that there are dreams which involve such reflection and thus insight into one's dreaming state. These so-called 'lucid dreams' are cases where the dreamer is aware, quite rightly, that he is dreaming. As such they do not help the sceptic as the dreamer is cognisant of his or her state. An excellent introduction to the intriguing phenomenon of lucid dreams can be found in Celia Green and Charles McCreery's book *Lucid dreaming: The paradox of consciousness during sleep*, Routledge, 1994

- 1 CSM II 13 (*The Philosophical Writings of Descartes*, translated by Cottingham, Stoothoff and Murdoch, Cambridge University Press, 1984, vol. II, p. 13.)
- 2 CSM II 13.
- 3 Or as Sigmund Freud put it 'the dreams which we produce at night have, on the one hand, the greatest external similarity and internal kinship with the creations of insanity, and are, on the other hand, compatible with complete health in waking life.' *Two Short Accounts of Psycho-Analysis*, transl. A.S. Strachey, Penguin, 1962, see 'Five Lectures on Psycho-Analysis', Lecture 3.
- 4 *Leviathan*, ed. C.B. MacPherson, Penguin, 1968 (1651), Part I, chap. ii, p. 90.
- 5 Locke, *An Essay concerning Human Understanding*, ed. Peter Nidditch, Oxford: Clarendon, 1975, IV.ii.14.

- 6 Austin, *Sense and Sensibilia*, OUP, 1962, p. 49
- 7 CSM II 14.
- 8 CSM II 49.
- 9 This, in turn, means that Descartes' own distinction between the imagination and the pure intellect, advanced at the beginning of the Sixth Meditation, is presupposed here.
- 10 Thoughts like this lead Norman Malcolm, in his classic monograph *Dreaming*, Routledge, 1959, to take the radical step of treating all dream accounts as unverifiable and thus the claim that we have any kind of experiences in dreams as meaningless.
- 11 CSM II 306.
- 12 Havelock Ellis, *The World of Dreams*, Constable and Co, London, 1911, see chapter one.



[Dreaming]



**Mark Sinclair**

# [Nietzsche]

and the problem of history

It is something new in history that knowledge wants to be more than a mere means.<sup>1</sup>



The present essay examines Friedrich Nietzsche's articulation of the problem of history that is to be found in a short but nonetheless pivotal text entitled *On the Advantages and Disadvantages of History for Life*.<sup>2</sup> The text was first published as the second of four *Untimely Meditations* in 1874, and thus in what commentators have come to isolate as the 'early period' of Nietzsche's work. In the essay, I aim to show what sort of a philosophical problem the problem of history is for Nietzsche, and how his response to it offers us a key to understanding the development of his work from its 'early' to its 'middle' period and even his work as a whole. In addition, I aim to give some indications as to how Nietzsche's account of history as a problem has been taken up in the work of 20<sup>th</sup> century philosophers.

The 2<sup>nd</sup> *Untimely Meditation* announces the outbreak of a certain sickness or malady of the age: an

excess of historical study and historical education within the German, and, more broadly, European culture of the late 19<sup>th</sup> century. The massive growth of historical studies in the 19<sup>th</sup> century has been described metaphorically as a 'discovery of the continent of history', a discovery comparable to that of the New World. For Nietzsche, however, this discovery amounts to an excessive concern for the past, and this excess would lead to a sickness – as an excessive consumption of anything usually does – because a limited range of historical knowledge is necessary to the health, happiness and creative powers of a people. Certainly, some historical knowledge is advantageous and necessary to life, but too much of it would be disadvantageous, having a harmful effect on the quality of our life itself. Although Nietzsche describes several different symptoms of this sickness in the course of his text, he initially makes this latter point by comparing the life of a culture or people to that of an individual: the vitality and vigour of a culture requires a limitation of the range of its historical knowledge, just as the moments of happiness or the moments of decision and action in our personal lives require us to live fully in the present, limiting our recollection of the past; by, in other words, forgetting the past. Without some such forgetting – which is always prior to the possibility of actively remembering

something and reflecting on it – we would, of course, become self-conscious to the point of distraction and alienation. If there is a historical malady in modern European culture, then, it would seem that the cure to the sickness lies in finding the right balance, for the sake of our lives, between an excess of historical knowledge or remembering and an excess of forgetting; between living in the manner of someone unable even to lift a finger because, remembering everything, he sees only becoming in things, only the transitory nature of events, and living in the ignorance – however blissful it may be to forget everything as soon as it happens – of a cow or goldfish. The ability to find this balance is what Nietzsche terms the 'plastic power'<sup>3</sup> of an individual or people.

On this basis the 2<sup>nd</sup> *Untimely Meditation* may appear to present some interesting psychological analysis and cultural criticism, but one might wonder how it could be claimed to possess a pivotal importance within Nietzsche's philosophical thinking. The philosophical stakes of this diagnosis of a historical malady begin to become clearer, however, with the recognition that the sickness arises 'through the demand that history be a science'.<sup>4</sup> The 'discovery of the 'continent of history' in the 19<sup>th</sup> century is accompanied by, and, in fact, occasioned by, the apparently reasonable idea that the

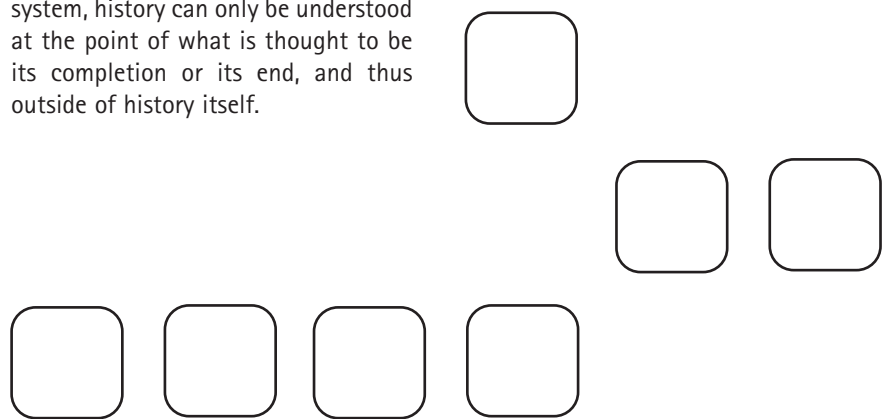
historian should learn from the modern natural sciences, adopting the position of the neutral and indifferent observer who does not import her prejudices into the object of the study. In her occupation with what can be known objectively in things, with knowledge that is valid for everyone and at all times, the modern scientist attempts to distance herself from her own subjectivity, her own particularity and concerns as an individual human being. This calculating objectivity is, then, what is required from the historian with the demand that history be a science. Henceforth historical knowledge is supposed to be an objective account of past events and epochs, an account that consequently exists for its own sake, rather than for the sake of life or of anything else; and it is the pursuit of historical knowledge for its own sake, which is to say the pursuit of historical knowledge as good in itself, that is precisely, as Nietzsche shows, what leads to an excess of historical knowledge:

Now life is no longer the sole ruler and master of knowledge of the past: rather all boundary markers are overthrown and everything which once was rushes in upon man. All perspectives have shifted as far back as the origins of change, back into infinity. A boundless spectacle such as history, the science of universal becoming, now displays what no generation has ever seen; of course, she displays it with the dangerous boldness of her motto: *fiat veritas pereat vita*.<sup>5</sup>

The demand that history be a science promotes the value of objective truth over and above any actual concern for our lives, and thus it can be characterised by the dictum: *let there be truth and may life perish*.

It is the conception of history as an objective science, then, that is the origin of the 19<sup>th</sup> century excess of historical knowledge. Yet what is the philosophical basis or the basic philosophical presupposition of this demand for objectivity? It is that the course of history as the object of historical studies is an object arrayed before an a-historical gaze; that, in other words, the human being insofar as it thinks is independent of history, outside of time; or that the human being is, in Nietzsche's words, 'an eternal non-subjectivity'.<sup>6</sup> In one sense, this determination of the human being as at least partially residing outside of history, and thus time, derives from Plato's determination of the soul in relation to the true, timeless world of the ideas, and it is more or less a constant, in some guise or other, in philosophy before Nietzsche. Even in the work of the 19<sup>th</sup> century philosopher G.W.F. Hegel, which constitutes an attempt to incorporate history and the history of philosophy in a philosophical system, history can only be understood at the point of what is thought to be its completion or its end, and thus outside of history itself.

In another sense, however, this determination of the human being is peculiar to modernity and modern philosophy, for it is here that thought aims to access truth, eternal truth, in deliberate and methodical abstraction from received wisdom and history. If every age almost inevitably believes itself to be wiser, cleverer than the last, then this tendency is nevertheless radicalised in modernity, which claims to be able to enlighten itself in putting paid to a long history of error. In the text that is generally held to constitute the beginning of modern metaphysics, namely Descartes' *Meditations on First Philosophy*,<sup>7</sup> we indeed witness the attempt of thinking to ground itself in and from itself independently of the history of philosophy; according to this method, the human being is apprehended as a self-grounding, timeless thinking thing for which everything else, save perhaps God, is arrayed as an object before it.



[History]

Nietzsche's diagnosis of the historical malady as deriving from the demand that history be a science, then, implies a critique of the philosophical tradition and, more specifically, a critique of the basic, Cartesian position of modern metaphysics; a position according to which the human being is abstracted from its own historical life. It follows from this that if Nietzsche is going to offer a different account of historical study than that holding it to be an objective science, this account will involve a transformation of the basic position of modern metaphysics. With what, then, does the *2nd Untimely Meditation* propose to replace the idea of historical study as an objective science? And what is the philosophical or metaphysical basis of this transformed account of the study of history?

Responding to the first question leads to a response to the second. There are, however, two levels to the response to the first question. In general, the task of the historian is compared to that of the artist, and here we meet one aspect of the early period of Nietzsche's work, a period which he himself characterised as an 'artist's metaphysics'. Historical study is shown to be always a question of interpretation and thus creation. The facts of history only have meaning within a framework of interpretation, a framework that the historian imposes on them from her own historical situation, and this is what Nietzsche means when he writes that the 'fact is always stupid'.<sup>8</sup> The very idea of objectivity, which Nietzsche shows to be an impossible 'ideal', only serves to conceal the particular prejudices and presuppositions of a historian and her age. Against the demand for such objectivity, then, a demand that assumes 'that whoever is

*quite unconcerned* about a past event has a calling to describe it',<sup>9</sup> Nietzsche argues that 'the past always speaks as an oracle; only as master builders of the future who know the present will you understand it'.<sup>10</sup> The meaning of what an oracle says is always, of course, a question of interpretation and, for Nietzsche, only those with a creative concern for our lives in the present and the future are fit to interpret it. In stressing thus the irreducibly interpretative and hence creative nature of historical study he introduces a problematic that has occupied historians themselves to this day, and which will be developed in the 20th century hermeneutic philosophy of, amongst others, Hans-Georg Gadamer and Paul Ricoeur.<sup>11</sup>

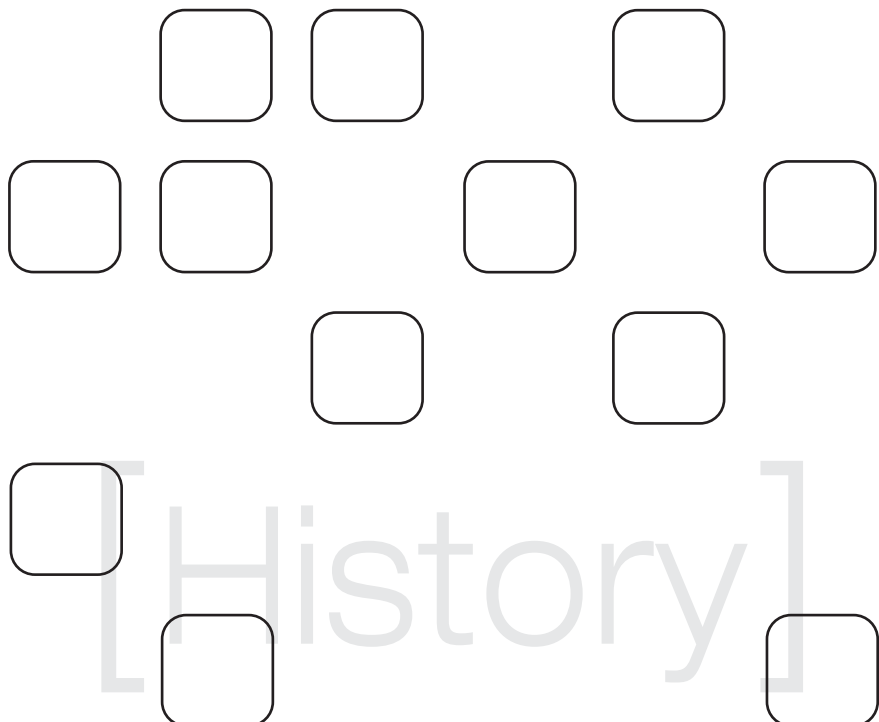
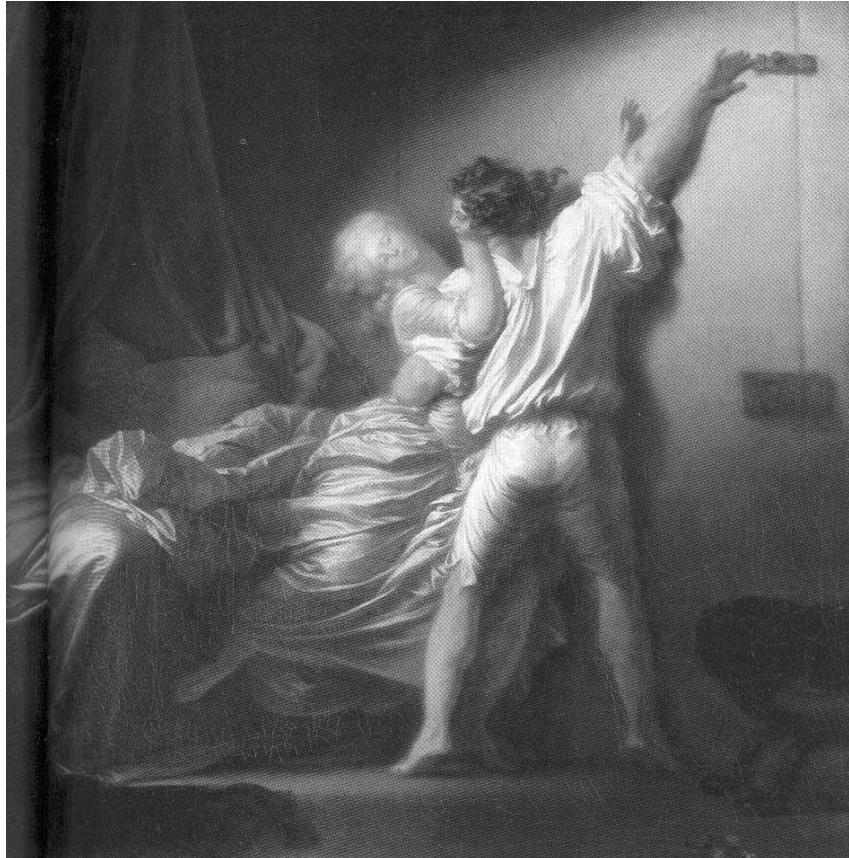
More specifically, the *2nd Untimely Meditation* articulates three seemingly distinct kinds or modes of historical inquiry, each of which can be advantageous to life, but also disadvantageous if pursued to excess. Briefly, these are: the *monumental*, which celebrates the great deeds of the past, essentially in the fashion of a 'great-man' theory of history, for the sake of great deeds in the future, but which can descend into all sorts of fanaticism and distortion of history; the *antiquarian*, which happily stresses tradition and the roots of the present in the past, but which can become a stultifying force; and the *critical*, which analyses, dissects and criticises the past, thus distancing it from the present. History carried out as a science is predominantly critical history, the excess of which, as we have seen, leads to the historical malady with which Nietzsche is concerned. We are led to imagine, then, that in some sense, a sense that Nietzsche does not at all clarify, history as a creative task would comprise these three modes, which

each seem to have a different relation to the past from the particular perspective of one of the three different 'moments' or 'aspects' of time: the *monumental* study of history privileges the future insofar as it is concerned with the *possibility* of great deeds, possibility being that which is not yet present or actual; the *antiquarian* privileges the past insofar as it seeks to show the roots of the present in the past; and the *critical* privileges the present, insofar as it seeks to distance the past from the present.

It is by asking in what sense time and its 'moments' could be intended here that we arrive at a response to the question of the transformed metaphysical basis of Nietzsche's thinking. This basis concerns the nature of time and the way that the human being exists as a historical being. First of all, the argument that the study of history is a creative act, and thus a function of the future, does not oppose the conception of an a-historical, eternal subject with the idea that the human being is simply *in time*, and that is to say in the present moment. Such would be the metaphysical position underlying historicism or historical relativism, which merely negates the possibility of objective historical knowledge with the simple claim that all knowledge is relative to the particular and present historical situation of the knower. Nietzsche does more than oppose historical relativism to historical objectivism; as he argues, it is the demand that history be a science that ultimately leads to relativism, since from the idea that previous periods of history can be examined objectively there is only a very short, perhaps inevitable, step to be made to the recognition that our age is itself just another one of these periods in history,

one with its own prejudices and historical positions. If the study of history is always an act of creative interpretation, then, it is not the case that the human being is either *inside* or *outside* time. It would be much better to say that the human being *is* time, *that the human being is historical in its essence*. Yet this idea necessitates a transformation of the common and quite traditional linear conception of the nature of time as a mere succession of present moments or 'nows' – nows that are no longer present, a now that is present, and nows that are not yet present – that receives its first sustained philosophical articulation in Aristotle's *Physics*.<sup>12</sup> For the claim that historical inheritance is always a task, that the past is only accessible by means of our opening onto the future, amounts to, and in fact presupposes, the idea that the past *is* what it *is* only by means of the future. But if the past is what is only by virtue of the future, then we can no longer understand the past to be simply sequentially prior, as a 'now' which is no longer, to the future as a 'now' which is not yet. In order to begin to get to grips with this difficult thought, we are in fact required to transform our understanding of the verb 'to be', if any meaning has previously been granted to the verb at all, since on this account it can no longer mean simply to be *present*, for the past *is*, and it is in a sense other than that of a 'now' which is no longer.

The *2nd Untimely Meditation* compels us to recognise – in the words of the philosopher Georg Simmel, a reader of Nietzsche, who recognised the 'logical obstacles' of this way of speaking – that 'life is really past and future'.<sup>13</sup> In his brief text Nietzsche himself does not explicitly elaborate on this sense of life and time; the concept of life



that features in the title remains undeveloped and ambiguous within the body of the text itself. Yet it is developed in the most deliberate fashion within Martin Heidegger's master-work of 1927, *Being and Time*, in which he remarks that Nietzsche's text 'allows us to suppose that he understood more than he has made known to us'.<sup>14</sup> Heidegger distinguishes between the study of history and what he terms historicity, which is human life or human existence as a being-historical. The former studies the latter, but the latter is the ever pre-objective movement of history or time itself. And yet history or time here is not to be understood according to what Heidegger terms the 'vulgar' conception of time that has been dominant in philosophy since Aristotle; it is here to be understood in its more original guise as what is termed temporality, according to which the past, present and future are not mere successive 'nows' but are rather – as one might say, although the expression is hardly adequate to the nature of the problem, since it supposes that temporality is *in* time – *at one and the same time*.

If, however, Heidegger's *Being and Time* does present us with an elaboration of the particular instantiation of Nietzsche's early 'artist's metaphysics' in the *2nd Untimely Meditation*, then this apparently happy philosophical partnership almost immediately falls asunder with the beginnings of Nietzsche's move beyond this 'artist's metaphysics' itself. These beginnings are indicated towards the end of his short text. For there is one crucial problem that Nietzsche will not allow himself to avoid here, a problem to which the turn to art is, in the end, held to be an inadequate response: in a sense his own thinking suffers from

the historical malady that it itself diagnoses; the critical and alienating distance from life that Nietzsche diagnoses in modern historical consciousness and modern metaphysics is ultimately repeated in his own thinking insofar as it is itself a criticism of the age. The problem relates to the very idea of an *untimely* meditation – to pose the problem in an interrogative form: how is it possible to take a critical distance from the present age, without staking a claim to an a-historical truth or an a-historical essence of the human being? It is as a result of *this* aspect of the problem of history that Nietzsche issues a threefold imperative:

The origin of historical education [...] *must* itself in turn be historically understood, history *must* itself dissolve the problem of history, knowledge *must* turn its sting against itself – this threefold *must* is the imperative of the spirit of the 'new age' if it really does contain something new, mighty, original and a promise of life.<sup>15</sup>

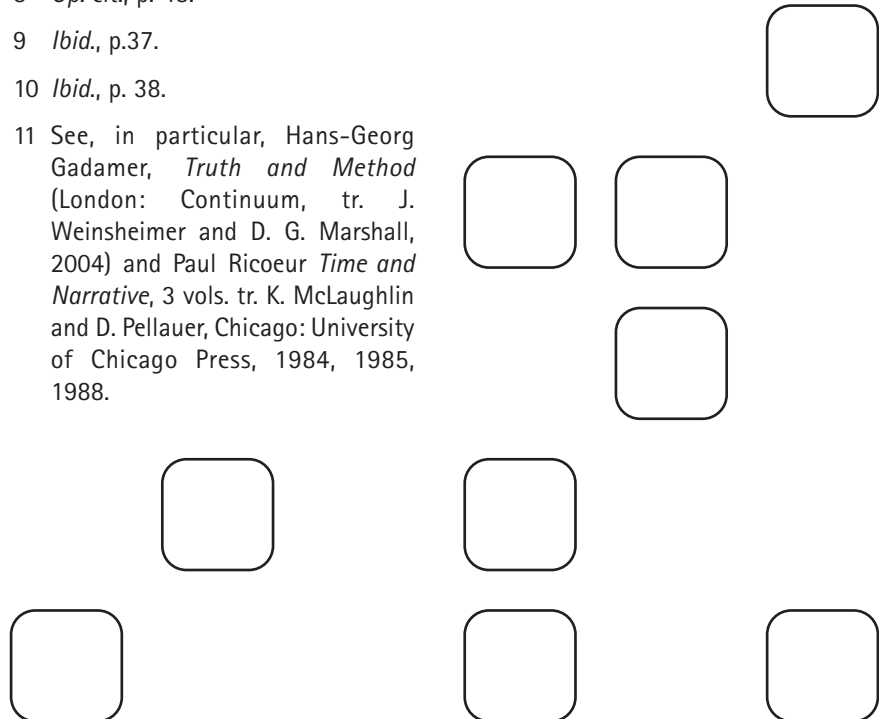
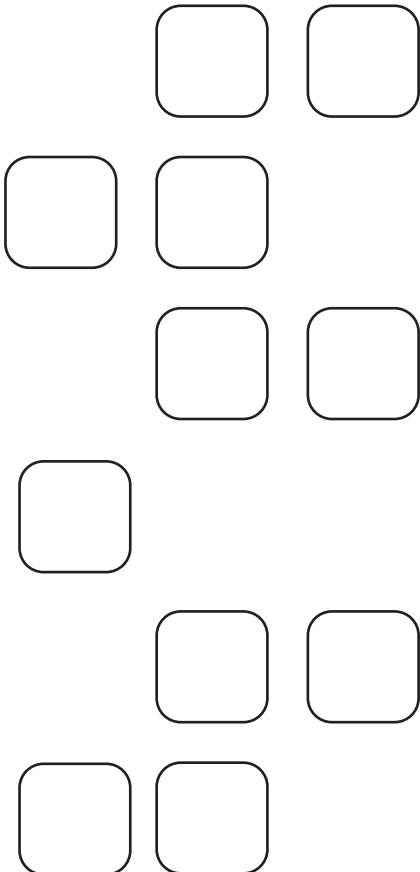
With this threefold imperative Nietzsche announces what commentators have called the 'middle period' of his work, which commences with the text *Human all too Human*, and ends with *The Gay Science*.<sup>16</sup> According to these three 'musts', the problem that modern historical education and the modern mode of knowledge represent is no longer to be addressed by means of an appeal to art and the activity of the artist. It is rather to be countered internally, as it were, by means of knowledge and historical education itself; the veritable antidote to the modern historical malady is now to be found within the cause of this malady itself. Such an attempt to 'turn the sting of knowledge against itself' is precisely what Nietzsche attempts in the texts

of the middle-period, texts which attempt to overturn the basic, Cartesian position of modern metaphysics by means of an extension, a more radical practice, of objective, scientific knowledge.

Many commentators have held that the transition in Nietzsche's work from its early to middle periods represents a transformation in his estimation of modern science; Nietzsche the artist would become Nietzsche the scientist. It is, however, necessary to recognise that this shift represents nothing like a change of opinion, a change in his estimation of both modern science and the metaphysical position on which it is based, but rather an attempt to 'turn the sting of knowledge against itself' in order to achieve a transformation of the way that we understand ourselves and the world, and, consequently, a transformation of the way that we exist. Perhaps the reading of the *2nd Untimely Meditation* that I have sketched in this essay would allow us to recognise that Nietzsche's thought as a whole is to be understood less as a set of changing positive philosophical doctrines than as a series of attempts to turn modern metaphysics against itself, for the sake of a 'promise of life', i.e. for the sake of a transformed historical human existence.



- 1 Friedrich Nietzsche, *The Gay Science*, tr. W. Kaufmann, New York: Vintage, 1974, § 123, p. 180.
- 2 Friedrich Nietzsche, *On the Advantages and Disadvantages of History for Life*, Hackett: Indianapolis, tr. P. Preuss, 1980.
- 3 *Ibid.*, p. 21.
- 4 *Ibid.*, p. 23.
- 5 *Ibid.*, p.23.
- 6 *Ibid.* p. 31.
- 7 René Descartes, *Meditations on First Philosophy*, tr. & ed. J. Cottingham, Cambridge University Press, 1996.
- 8 *Op. cit.*, p. 48.
- 9 *Ibid.*, p.37.
- 10 *Ibid.*, p. 38.
- 11 See, in particular, Hans-Georg Gadamer, *Truth and Method* (London: Continuum, tr. J. Weinsheimer and D. G. Marshall, 2004) and Paul Ricoeur *Time and Narrative*, 3 vols. tr. K. McLaughlin and D. Pellauer, Chicago: University of Chicago Press, 1984, 1985, 1988.
- 12 Cf. Book IV of Aristotle, *Physics*, tr. P. Wickstead & F. Cornford, London: Loeb, 1957.
- 13 Cited by Hans-Georg Gadamer in *Truth and Method*, *op. cit.*, p. 264.
- 14 Martin Heidegger, *Being and Time*, tr. J. Macquarrie & E. Robinson, Oxford: Blackwell, 1995, p. 448.
- 15 *Op. cit.*, p. 45.
- 16 Friedrich Nietzsche, *Human all too Human*, tr. R. Hollingdale, Cambridge University Press, 1996.



[History]

**Paul Sperring**

# Should We Give Up on [Reductive] Physicalism?

Supposing you were a physicalist in the late 1950s, early 1960s, and supposing you were Australian too, it is highly likely you would have thought that mental properties could be reduced to physical properties. Now, suppose you are a contemporary philosopher of mind and suppose further that you are also of a physicalist stripe. Will you be inclined to think that mental properties are reducible to physical properties? It's by no means certain. These days physicalists fall into two, broadly conceived, camps: (i) the reductionist physicalists who think that minds (or mental properties, or states or events<sup>2</sup>) can be reduced to brains (or something smaller) and; (ii) the non-reductive physicalists who think that minds are not straightforwardly reducible to some lower level set of physical properties.

In truth if one were to carefully classify all the physicalist positions in contemporary philosophy of mind we would need distinctions of a much finer grain than this story suggests. For the purposes of this paper, however, those philosophers who have thought that mental properties can be reduced to lower level properties will be lumped together (and called 'reductionists') and those philosophers who, although embracing physicalism, have thought that mental properties in principle defy reduction to something lower down will also be

lumped together (and called 'non-reductive physicalists'). As far as name calling goes these days 'reductionist', in some hands, takes a bit of beating as a philosophical insult.<sup>3</sup> The problems with reductionism have been well documented of course, and I will focus on one particular attack. But I also want to say something about what might be worrying about the non-reductive physicalist's position.<sup>4</sup>

## The Type-Identity Thesis

One of the most famous (if short-lived) reductionist physicalist views is known as the Type-Identity thesis. Various philosophers in the 1950s argued that mental properties are identical to physical properties – thoughts, sensations, feelings and so on (mental 'types'<sup>5</sup>) *are just* brain processes (physical 'types'). So when persons are said to have some mental property, say the property of being in pain, then analysis will reveal such a property to be a property of their brains, and nothing more.

Perhaps you are struck by the *prima facie* plausibility of such a position. Certainly, these days, when we think of minds we automatically assume a role for the brain, knowing what we know about the close relationship between the two. So when we come to examine how the mind and the

brain are related we often think about it in terms of correlation or cause. By correlation we just mean that wherever we have some mental property it is always accompanied by some physical property. Now, mere correlation explains nothing of course. We want to know *why* every mental property is accompanied by a physical property. So perhaps we think about the relationship in terms of causation. All mental properties, we might say, are *caused* by physical properties of the brain. However, if we stop here then we still face a problem. How, we ask, do physical properties give rise to mysterious non-physical properties, and how, if those mental properties are non-physical do they exert any influence over the properties of the brain (this is the well-known interaction problem that Princess Elisabeth of Bohemia brought to Descartes' attention in the Seventeenth century, although the debate was couched in terms of substances rather than properties)?

One way out of this problem is to say that the reason that mental properties are always accompanied by brain properties is because mental properties *just are* properties of the brain. There is also no residue to explain causally – i.e. how mental properties *arise* as a result of brain goings on – since we only have one type of phenomenon, at bottom. If A and B are identical then A can't be the





cause of B, and there is no mystery how B can cause things if there is no mystery about how A causes things (and if A is a brain property then ultimately physics will tell us how that does what it does<sup>6</sup>).

Some of the advantages of accepting the type-identity thesis ought to be pretty clear. Firstly, the problem Descartes wrestled with, how minds and bodies could be in causal communion, seems to melt away. There is no violation of well established physical laws on this view (for instance the principle of the conservation of momentum) since we have well-behaved physical particles all the way down (well maybe most of the way down – but I'll leave quantum particles to one side).

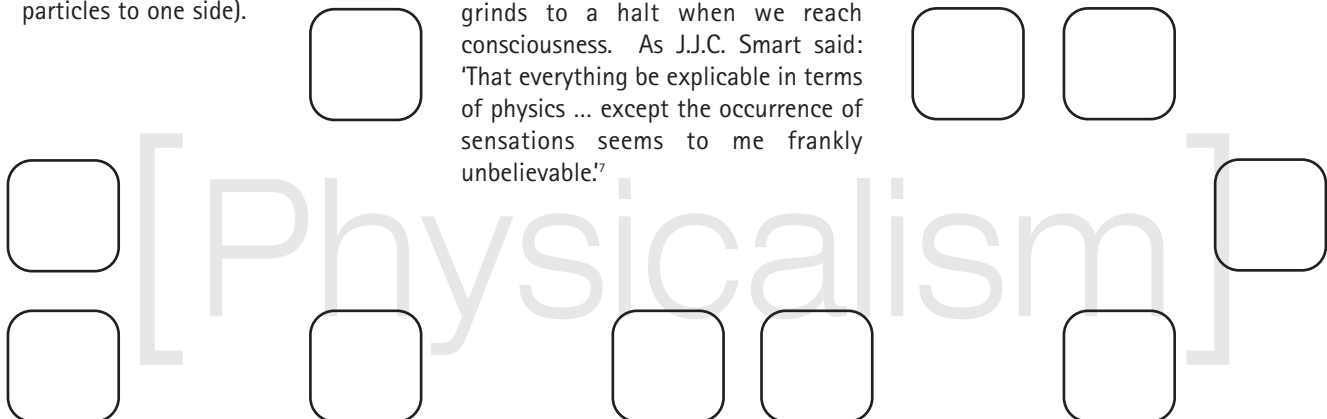
Secondly, we get a much simpler view of the world – one type of phenomenon subject to one set of laws. There is no need to account for consciousness by introducing some new set of fundamental laws into our story of everything (anyone sympathetic to Occam and his razor will find something to cheer about here).

Further, as a general claim in support of the reductive account of mind, science has been incredibly successful in explaining all sorts of phenomena reductively. Things of a complex nature which used to be considered as real 'in themselves' turned out to be made of up things of a simpler nature, but in a way that held them to be just the same thing as the simpler phenomena. By way of a well worn example, it was discovered that 'temperature' is in fact 'mean molecular kinetic energy' – not, of course, that the mean molecular kinetic energy of your bath water *causes* (still less is just correlated with) its temperature. Given this success it might then strike you as odd that the process of explaining things which has been applied all the way up to biological mechanisms (think, for instance about 'life' and how it can be explained reductively in terms of complex bio-chemical processes, which themselves can be explained in terms of simpler phenomena) suddenly grinds to a halt when we reach consciousness. As J.J.C. Smart said: 'That everything be explicable in terms of physics ... except the occurrence of sensations seems to me frankly unbelievable.'<sup>7</sup>

## The Multiple Realizability Argument

'Well, you had better believe it', said a number of philosophers. One was Hilary Putnam in a paper called 'The Nature of Mental States', which purported to show that the type-identity theorist's project was hopeless. His argument went something like this:

If the type-identity thesis is to make good the claim that some mental property type M is identical to some physical property type P then it must provide an account of how it is that for some token<sup>8</sup> m of M, m can be realized not merely by tokens of type P but also, under different circumstances, by tokens of types P<sub>1</sub>, P<sub>2</sub> ... P<sub>n</sub> (where these are tokens of distinct types of physical properties). According to Putnam, that is, the defender of type-identity must have it that for any organism S in M there must be some P had by S *necessarily*. Since some S could have M in virtue of having P<sub>1</sub> but lack P entirely then, simply, M can't be identical to P. Conversely, if the having of P does not guarantee the having of M then, once more, M and P don't appear to be identical.



Let's go a little more slowly. Suppose we take 'being in pain' as our specified type of mental property *M*, and then specify that the neural substrate (or brain property) type, *P*, with which it is supposedly identical, is some C-fibre stimulation. Now, 'do octopi have pains?' asks Putnam. He supposes so. 'Do octopi have C-fibre firings?' He supposes not. And if his suppositions are right, and they seem to be, then we have a clear counterexample to the type-identity ( $M=P$ ) thesis. Identity claims being what they are this just won't do. If *M* is supposed to be identical to *P* then wherever there is a *P* there ought to be an *M*, and wherever there is an *M* there ought to be a *P*. If Cicero=Tully, then when Cicero is in the bath so is Tully, if Tully is combing his hair, then so is Cicero. In short, whatever is true of one of the 'pair' in an identity relation must be true of the other. In the case of the octopus an *M* is present but no *P* (and, for the converse, if there were a zombie<sup>9</sup> who had *P* then there would be, by stipulation, no *M*). So, either there is something wrong with Putnam's argument or we ought not to claim that  $M=P$  with respect to pain. The same will be true for any mental state that can be realized by distinct types of physical properties.

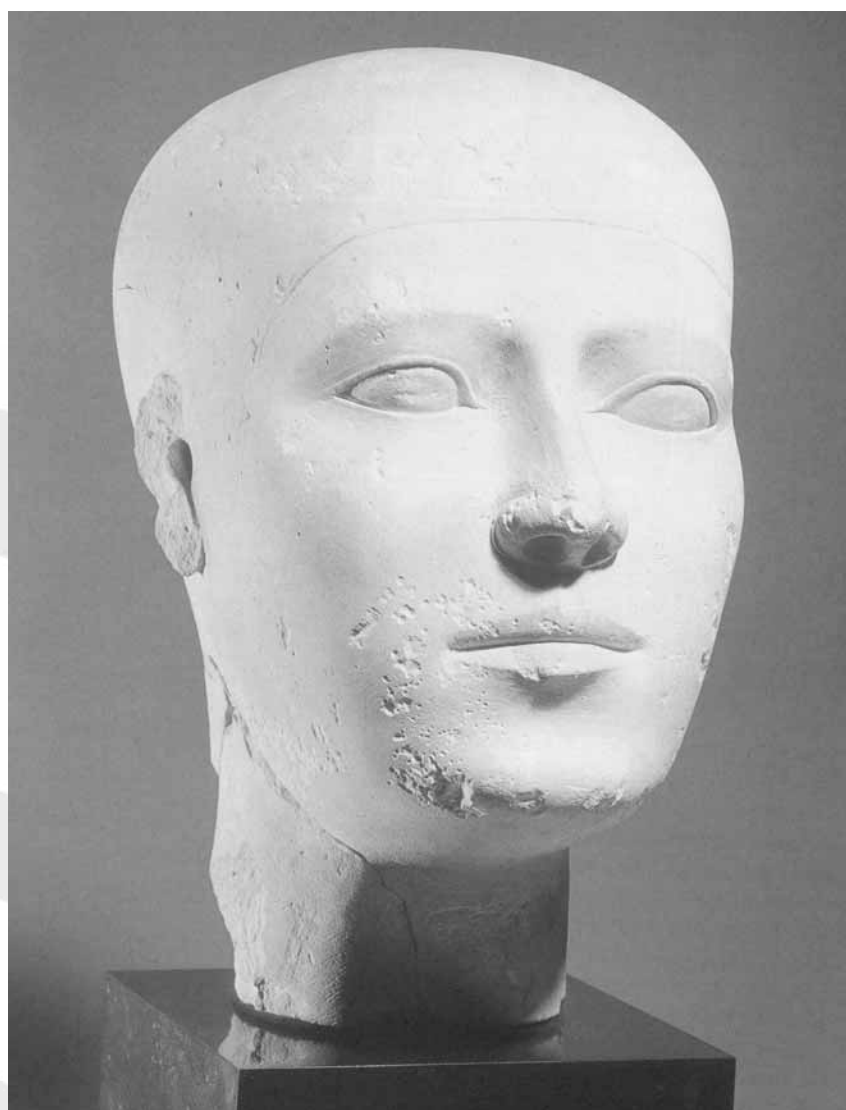
For a number of commentators Putnam's argument pretty much decided things against the type-identity thesis. Better to say, the conclusion ran, that mental states are identifiable with *functional* states which then allows for their multiple realizability in all sorts of physical organisms (perhaps even in non-physical ones, like angels, as well).

### Why is Multiple Realizability a Problem for the Reductionist?

Jaegwon Kim<sup>10</sup> offers an analysis of why, exactly, 'the Multiple Realization Thesis' (or MR) is standardly taken to be problematic for the identity thesis.

Among the reasons given for the death of type materialism is, firstly, that if mental state types can be realized in heterogeneous ways then they cannot be *defined* physically. The idea here

seems to be that any definition *D* of some mental state type *M* will be at best disjunctive since tokens of *M* are realizable by tokens of *P* or *P*<sub>1</sub> or *P*<sub>2</sub>... or *P*<sub>*n*</sub>. What this means is that the definition of 'is in pain' would include not just one neat physical property, but a rather wild, open-ended, collection of physical properties – 'has brain property of a human variety or has brain property of an octopus variety or ...' and so on. But a disjunction of such properties isn't itself a physical property (in fact, it might be said, it is a proposition and



not a property at all, and no physicalist is going to claim that mental states are identical to propositions, since these are abstract entities).

Secondly, and similarly, there can be no *reduction* of a single psychological property to some basic physical property because the former property needn't involve any relation at all to the latter (certainly in some possible world there could be just such and such a psychological state without there being the physical state that, in this world, it is supposed to be reducible to). If the putative reduction base isn't the same across all possible worlds then it's hard (well, impossible) to see how it (the absent physical base) can just be the same as the very real mental phenomenon in question.

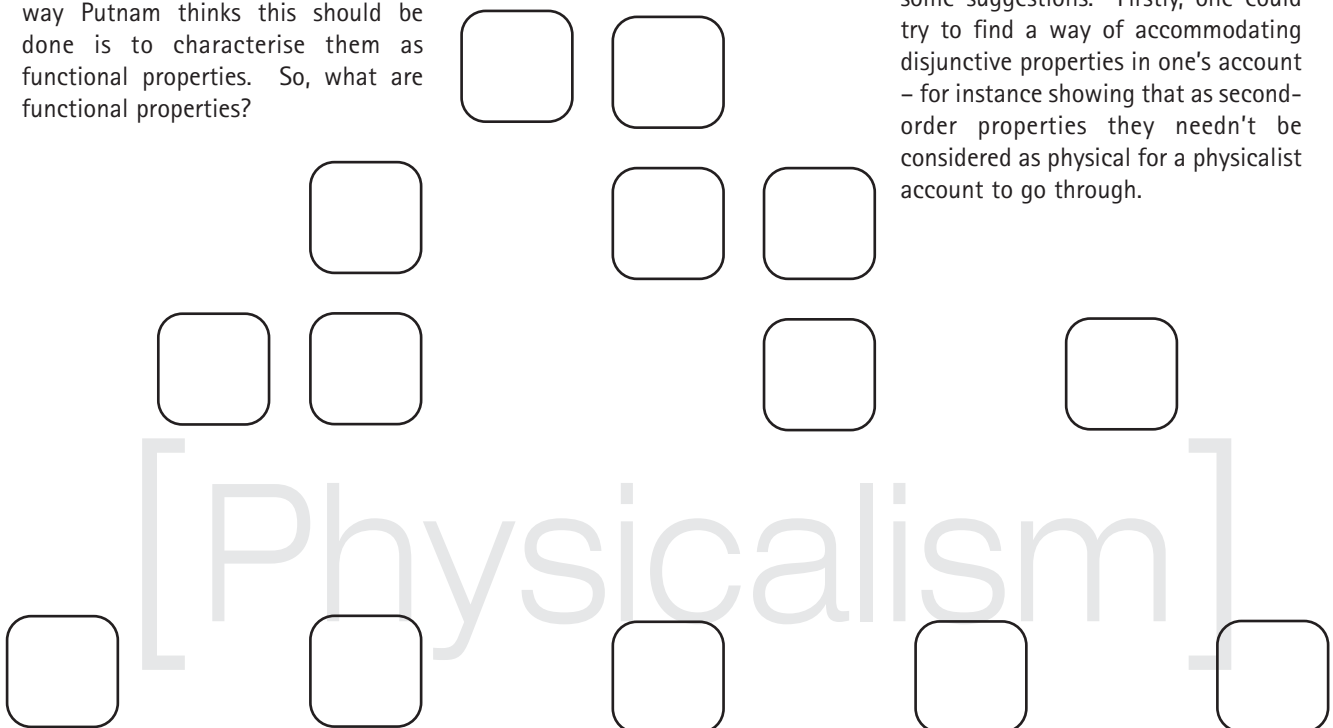
What, then, are the consequences of MR? Well, as suggested above, we would need to characterise mental properties differently in order that they may be predicable of a range of distinct (physical) phenomena. The way Putnam thinks this should be done is to characterise them as functional properties. So, what are functional properties?

A functional property is to be understood dispositionally, or in terms of the causal role that it plays. It is easy to see how this works with something like our pain case. You and the octopus have just been badly bitten by something and we would say, ordinarily, that both of you are in pain. What this means is not that the two of you share some physical state type or property, but rather that you are both in the same functional state or possess the same functional property. Pains typically arise from bodily damage, are typically related to a system's other internal states, and typically manifest themselves in some sort of behaviour (usually behaviour related to fleeing the cause of damage, or attending to the damage, perhaps involving seeking some means of repairing it, and so on). How is it then that these very different organisms, with their wildly distinct physiology and neurology, can both be said to be in the same state, or both sharing some *one* property?

Well, says the functionalist, functional properties are 'higher level' properties. They come with lower level properties at their base, to be sure, and it seems sensible to say that the base will be a physical base, but that they are not simply identified with the base. Functional properties are had by different things insofar as these things have the appropriate dispositions given a range of inputs and internal workings, and which in turn generate a series of outputs. These functional, or higher level, properties are *realized* by a range of lower level properties. The realizers may be octopi neurological properties, or human neurological properties, or perhaps they might even be non-neurological properties (properties had by machines or Martians).<sup>11</sup>

### What Now?

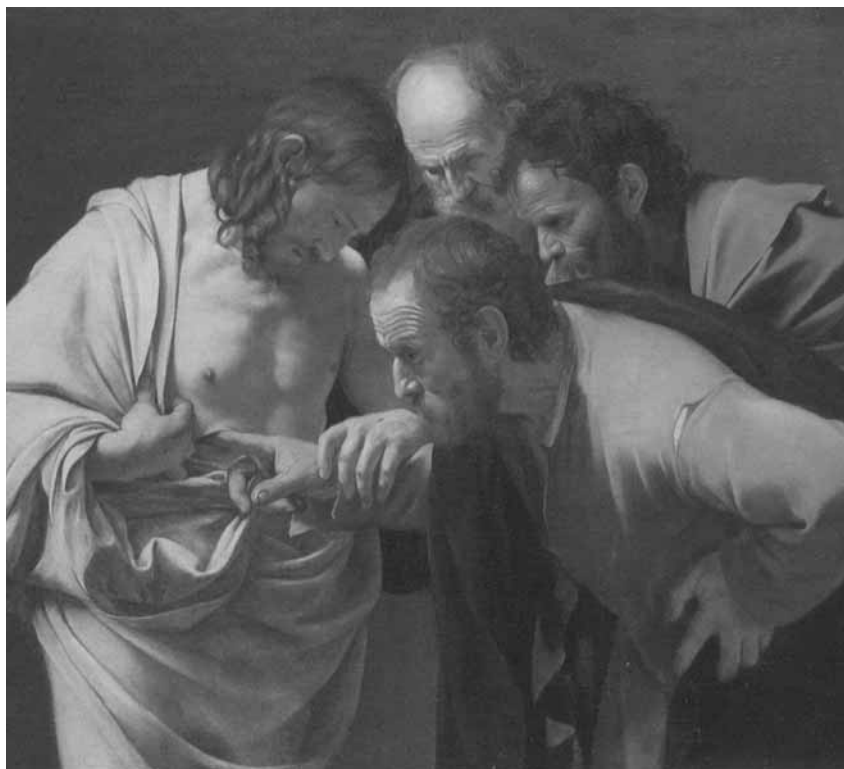
How might one respond to MR if one were inclined to hold on to something like reductive physicalism? Here are some suggestions. Firstly, one could try to find a way of accommodating disjunctive properties in one's account – for instance showing that as second-order properties they needn't be considered as physical for a physicalist account to go through.



Secondly, one might show that mental states, globally conceived, are not individuable as kinds. On this view there would be no such thing as 'pain as such'. While this approach might appear to have an eliminativist ring to it, it doesn't actually do away with mental states such as pains,<sup>12</sup> but rather dispenses with Pain as a state applicable across a multitude of distinct physical organisms (and, perhaps, some non-physical ones too). So different pains (in Martians, in octopi, in humans) get individually classed as kinds but Pain, supposed to include all of the instances in one kind, turns out not to be a kind. So, the pain that one gets from treading on a dead bee with one's bare foot (sharp, hot, centripetal waves of it, as I discovered to my cost this past summer) still really exists and, further, there is no need for 'wild' disjunctive properties to account for it.<sup>13</sup>

On both suggested saving models mental concepts are taken to be diffuse considered in one sense, and simple considered in another. On the first account Pain (big P) turns out to be rather coarse-grained and picks out a whole cluster of things (supportably as a disjunction of properties, or insupportably as a hodgepodge of disparate properties, not really classifiable under the one concept). On the second account, pains, relativized to a subject or collection of similarly constituted subjects are perfectly acceptable as properties and neatly co-variant with the brain properties (or whatever) that they get identified with.

Some problems with the second sort of strategy spring immediately to mind. If the model is going to be acceptable then it better account for all mental states in this way (or explain why pain is a special case, and deal with the other sorts of mental states



differently). So, one might ask, are properties like 'being hungry', or having some desire, or some belief or other, relativizable to species? If a Martian believes (while visiting Earth and watching the Old Trafford Test Match on TV.) that 'it is raining in Manchester' is this a case of Martian belief, distinct from human belief, or is the same belief (the same mental property) being had by two distinct sorts of organisms?

If we do relativize all mental states to species or collection of subjects (and perhaps to one and the same subject across different times and in different circumstances) aren't we in danger of multiplying perfectly intelligible states endlessly, and in a rather ad hoc fashion, simply as a means of escaping the MR claim? Well, perhaps. What seems true, at least, is that psychology, considered as a science of mentality wherever it is found, no longer looks

very unified on this account. We might even struggle to neatly carve it up into sub-branches, such as 'human psychology', or 'bat psychology', or 'Martian psychology', since human pains can be realized by distinct physical types of brain properties in different persons (and in the same person at different times).

But at this point someone might ask why go to all this trouble to avoid MR. It seems perfectly natural to think of distinct sorts of physical things having one and the same property so why not just accept that higher level mental properties cannot be reduced to their lower level bases and that's that. Well, I am going to finish with a suggestion why things aren't so easy for the non-reductive view, despite the apparent reasonableness of MR.

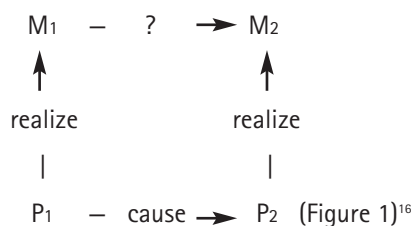
## Non-Reduction and Mental Causation

I suggested earlier that one of the reasons for finding reductive physicalism attractive was that the intimacy of the identity relation provided us with the resources to solve the problem of mental causation. The thought was that if mental properties just are physical properties, and we have no problem understanding how physical properties are causally efficacious,<sup>14</sup> then mental causation comes for free.

Now, the non-reductive physicalist wants to give an account of mental causation too, keeping the mental real (i.e. not wanting to eliminate it) and making it entirely dependent on, but irreducible to, the physical. The question then arises, how do mental properties make a difference?

Suppose that we accept the non-reductive physicalist's claim that for every mental property there is some physical property that realizes it. Suppose also that we are happy enough with the view that for every physical effect there are physically sufficient causes for that effect. And finally, suppose we are also accepting of our non-reductive physicalist's belief that physical effects are not systematically overdetermined – that is, she takes it that, most of the time, for some particular effect, *e*, that *e* was brought about by some one cause, *c*, and that *e* would not have occurred were it not for the occurrence of *c*.<sup>15</sup>

Taking all of these things together we have enough to raise a worry about the causal efficacy of the mental. Here's how:

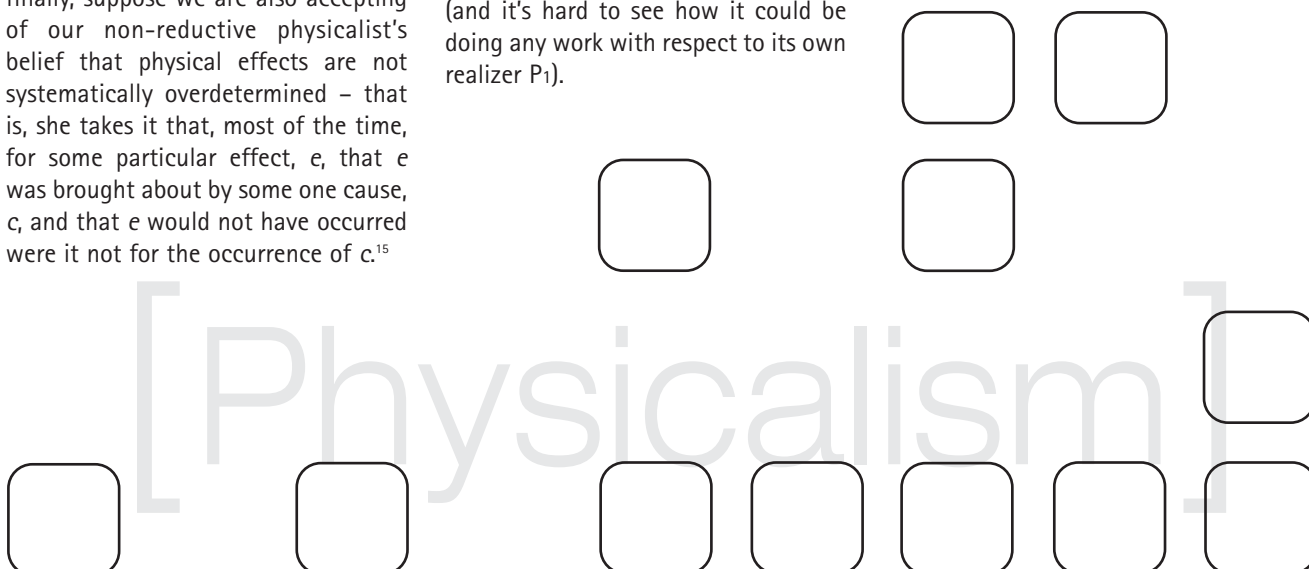


In figure 1 *P*<sub>1</sub> caused *P*<sub>2</sub>, where the *P*<sub>s</sub> in question are physical properties, and this seems to be reasonably uncontroversial (at least we'll assume so here). Now each *M*, or mental property, in this story is dependent on its physical realizer, so the question is, what work is *M*<sub>1</sub> doing? If we take *M*<sub>1</sub> to be the desire to do something or other, and *M*<sub>2</sub> the intention to bring it about that that something or other be done, we can ask how *M*<sub>1</sub> is relevant to the bringing about of *M*<sub>2</sub>. If *P*<sub>1</sub> realizes *M*<sub>1</sub>, and *P*<sub>2</sub> realizes *M*<sub>2</sub>, and *P*<sub>1</sub> causes *P*<sub>2</sub> (and hence *M*<sub>2</sub>) to come about, then is there anything at all left for *M*<sub>1</sub> to do? It doesn't seem to be relevant to the bringing about of either *M*<sub>2</sub> or *P*<sub>2</sub> (and it's hard to see how it could be doing any work with respect to its own realizer *P*<sub>1</sub>).

We have assumed, remember, that all physical effects are covered wholly by physical causes, and that physical effects are not overdetermined, for the most part, so is there any room for the mental to enter into the causal story. If not, then we have the unfortunate consequence that the mental is *epiphenomenal*, that is to say, it is causally inert.

Why should we worry about epiphenomenalism? Well, according to some philosophers if epiphenomenalism were true then it would be wrong to say that, for instance, we speak or act at all.<sup>17</sup> Or, even more dramatically, as Jerry Fodor says, if it were true that the mental had no causal powers, then 'practically everything I believe about anything is false and it's the end of the world.'<sup>18</sup>

The argument I have sketched out (roughly) here is sometimes called the 'exclusion' argument, since it excludes mentality from the causal story. The mental appears to be indispensable to an account of the right causal story for why we do the things that we do, however. It is certainly hard to see how we could be responsible for the things we do if the causes operate at a level below our awareness.



If we want to be non-reductive physicalists then, and also keep the mental operative in the causal nexus, then we should have to find something wrong with the exclusion argument. There may well be something wrong with it, although I am not going to pursue that discussion here. I simply wanted to point out that rejecting the reductive picture of mind presents its own set of thorny philosophical problems. As Kim says, there are no 'metaphysical free lunches' in this debate, so while MR seemed to push us away from reductionism, the exclusion threat might, in the end, lead us back there.

- 1 The mind-brain identity thesis was popularized particularly by Australian philosophers such as U.T. Place and J.J.C. Smart.
- 2 I intend to talk primarily about mental *properties*.
- 3 '[The] word 'reductionism' seems by now to have acquired a negative, faintly disreputable flavor – at least in the philosophy of mind. Being a reductionist is a bit like being a logical positivist or member of the Old Left – an aura of doctrinaire naïveté hangs over him.' Kim (1989: 32)
- 4 I take no credit for these arguments, they can be found in the works of J. Kim and others.
- 5 'Types' are to be distinguished from 'tokens'. An itch or a pain considered in general is a mental *type* – 'the itch that I am feeling right now in my left ear' is a *token* of the type.
- 6 Of course, metaphysics might have something to say about that.
- 7 Smart (1959) 'Sensations and Brain Processes' quoted in Kim *Philosophy of Mind* (1998: 53)
- 8 Type/token talk is, crudely, just general/particular talk. For example, a type of property, such as 'being red', is had by three red snooker balls. We have in this case one property type, and three tokens of that type.
- 9 A zombie is a theoretical entity that is supposed to be a molecule for molecule replica of you minus all of your mental properties (well, specifically the conscious ones, such as your qualia). If zombies are possible beings then, the argument runs, mental properties cannot be physical properties. I don't intend to say anything more about zombies here.
- 10 'Multiple Realization and the Metaphysics of Reduction'
- 11 I am not here saying anything about the different sorts of functionalist positions (as with physicalism there are many different versions of functionalism), but rather am lumping together any position that sees mental states as (a) functional states and (b) multiply realizable.
- 12 Not that eliminativists ever *really* deny that there are such phenomena – although they are often caricatured as taking us to be insensate automata.
- 13 Jaegwon Kim (1992) discusses strategies of these sorts, although my account of them shouldn't be thought to be a detailed and accurate rendering of his position – I simply offer here a boiled down version of the general reductionist program that Kim offers there and elsewhere.
- 14 Not everyone accepts that physical causation is metaphysically unproblematic – but there appears to be a long tradition of philosophers happy to accept that if only we could make sense of mental items in terms of physical items we would thereby make sense of their causal status. At the very least the ontologically simpler view on the cards with reductive physicalism means that we avoid the troubles with causation across completely distinct phenomena – which was thought to be what did for Descartes.
- 15 This wouldn't be the case if the effect were overdetermined, since if the cause had been missing the effect would still have come about (because the other, overdetermining, cause would have done the work). When the tightrope walker falls but is attached by means of a wire to the circus tent's roof, and has a safety net below her, we have a case of the overdetermination of the effect of her *safely* falling.
- 16 I take this from John Heil who raises worries about the non-reductive physicalist view in his book *From An Ontological Point of View*
- 17 Norman Malcolm 'The Conceivability of Mechanism'
- 18 Fodor 'Making Mind Matter More' p. 156

Kim, J 'Multiple Realization and the Metaphysics of Reduction' *Philosophy and Phenomenological Research* 52: 1-26 1992

Putnam, H 'The Nature of Mental States' *Mind, Language and Reality: Philosophical Papers Volume 2* Cambridge: Cambridge University Press 1975

**References**



Fodor, J 'Making Mind Matter More', from *A Theory of Content and Other Essays* (London: MIT Press, 1992)

Heil, J *From An Ontological Point of View* (Oxford: Oxford University Press, 2003)




Kim, J 'Multiple Realization and the Metaphysics of Reduction', from *Philosophy and Phenomenological Research* (1992)

*Philosophy of Mind* (Boulder: Westview Press, 1996)



*Mind in a Physical World* (Cambridge, Mass.: MIT Press, 1998)

Malcolm, N 'The Conceivability of Mechanism', from *The Philosophical Review* (1968)



Putnam, H 'The Nature of Mental States', from *Mind, Language and Reality: Philosophical Papers Volume 2* (Cambridge: Cambridge University Press, 1975)




Smart, J.J.C. 'Sensations and Brain Processes', from *The Philosophical Review* (1959)






[Physicalism]

**Peter Wyss**

Two Aspects of

# [Emergence]

It was on a dreary night of November, that I beheld the accomplishment of my toils. With an anxiety that almost amounted to agony, I collected the instruments of life around me, that I might infuse a spark of being into the lifeless thing that lay at my feet. It was already one in the morning; the rain pattered dismally against the planes, and my candle was nearly burnt out, when, by the glimmer of the half-extinguished light, I saw the dull yellow eye of the creature open; it breathed hard, and a convulsive motion agitated its limbs. (Shelley, 1818, p. 57)

## I. Introduction

This is the dramatic moment when Frankenstein assembles his pitiable monster: the creature cannot but open its 'dull yellow eye', once a masterly arranged configuration of material components is in place. Of course, rather than giving a detailed account of how to instigate the emergence of life from matter, Shelley unfolds the moral consequences of Frankenstein's nightly activities. The emergence of life and consciousness remains elusive and enigmatic to this day.

Many philosophers react in one of two ways to emergence. Some shrug it off as an impertinent metaphysical extravagance, while others excitedly

welcome the idea as a trendsetting *dernier cri* without caring too much about the details. I shall here resist either extremes and examine more soberly why emergentism—*viz.* the idea that there is emergence—is still around after its first philosophical appearance over hundred years ago. My concern is the *pull* the idea exerts in recent philosophy. Given the fact that emergentism stubbornly perseveres, there seems to be a market for its message. Hence my two questions are, why does emergentism *reappear*, and, what is the message of emergentism?

I will explore the following answers: emergence is a versatile idea whose resilience is fuelled by discontent. Its attraction lies in a combination of ontological proposals and a measured attitude of moderation—the two aspects of emergence. As a pointedly middle way, emergentism has the power, so to speak, to stimulate re-thinking of both reductive and non-reductive physicalism, and hence to mould the current discourse in the metaphysics of mind.

## II. The Legacy of Classical Emergentism

The resurgent interest in emergence during the past decade has produced not only a staggering amount of published work (more than seventy

publications since 1995), but also a confusing variety of emergentisms, all tailor-made to specific purposes. Now that the dust is settling after the renaissance of emergence, there are attempts to chart and classify the various ideas, and a proper re-evaluation has only just begun (see, e.g., Gillett, 2002; O'Connor & Wong, 2002; Silberstein, 2002; Stephan, 1999b; Van Gulick, 2001). Although it is common today to present a coherent picture of classical emergentism (doubtlessly coating the current brands in authority and dignity), this conceals the fact that emergentism in its heyday was as diverse as it is today. Now the diversity is truly enormous: it ranges from the cognitive emergence of new ways of thinking to the emergence of individuals with free will that survive death. Rather than providing a survey of the current conceptions, let me revisit the stalemate between vitalism and mechanism of which emergentism was the product, followed by a brief look at the current context of revival. I will then extract two central ontological posits and the characteristic attitude of emergence. First, here are simplified versions of vitalism and mechanism.

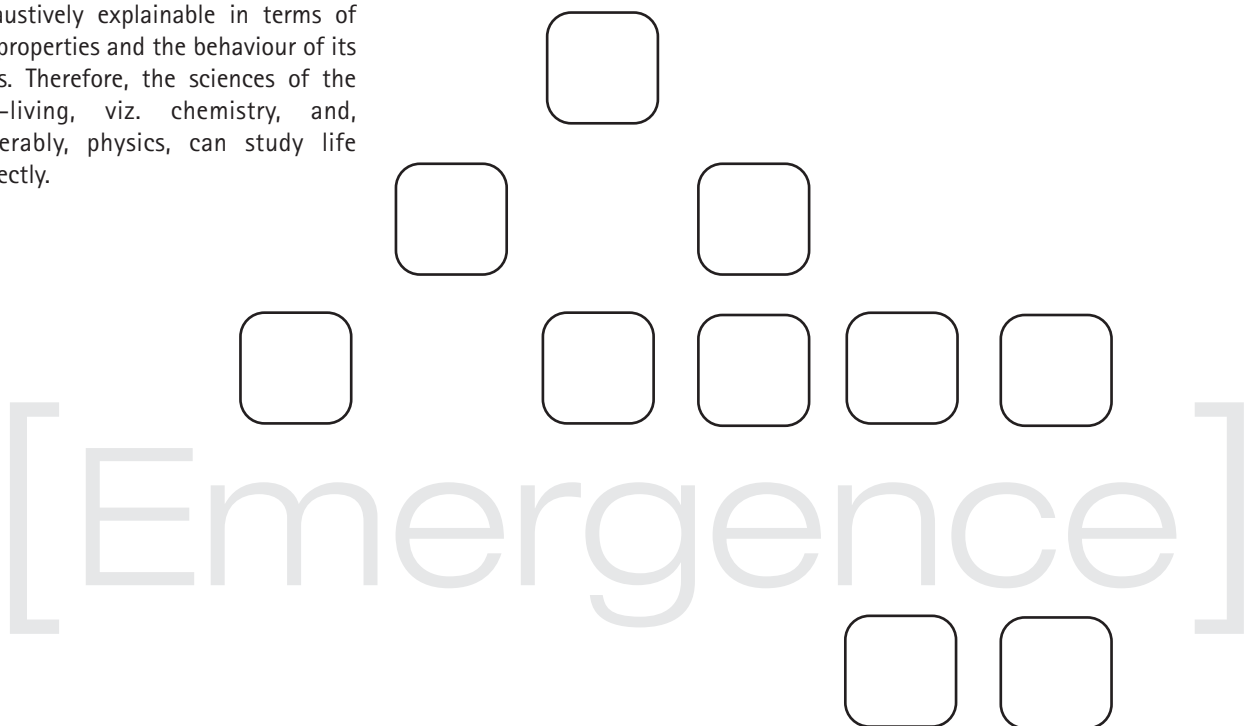
*Vitalism* is the theory that the presence of a special component is decisive for turning mere chunks of matter into organisms. This peculiar non-spatial ingredient is a



'quasi-intelligent directive force' (Lloyd Morgan, 1899, p. 184), which has 'the power to interfere by way of regulation or control with the physicochemical processes of the body [and] can suspend the second law of thermodynamics' (Hoernle, 1918, p. 630). Organisms are hence compounds of some suitably arranged material structure and a further non-material part. (Where did Frankenstein get *that* from?) Since the material parts do not determine the properties of an organism as a whole, vital properties are not explainable in terms of the parts' properties. If vitalism is true, science cannot explain life and mentality. *Mechanism*, in contrast, claims that organisms are not essentially different from stones or clocks. Life is simply due to a peculiar dynamical arrangement of matter. Organisms are but the sum of their parts working together like a machine, and the laws governing the whole are a consequence of the laws holding for the parts. Likewise, every (collective) property of the organism as a whole is exhaustively explainable in terms of the properties and the behaviour of its parts. Therefore, the sciences of the non-living, viz. chemistry, and, preferably, physics, can study life perfectly.

When in this tension emergentism arose, the idea was to find some middle ground between these extremes. Emergentists voiced concern and dissatisfaction, yet spotted valuable insights in both theories. On the one hand, with mechanists, classical emergentists thought there is no extra thing inherent in organisms; on the other, in line with vitalists, they thought life and mind add new and distinctive features that could not be captured by a purely mechanical approach. Two central ideas emanate from classical emergentism: first, life and mind are essentially distinctive from physical matter, yet, secondly, life and mind are dependent on physical matter, or some more fundamental stuff. In short, while life and mind are 'grounded' in matter, they 'go beyond' it. The different brands of current emergentisms mentioned above are mixtures of these two ideas.

The contemporary revival of emergence does not relate to the problem of life, but to the problem of consciousness (or 'the' mental in general). More specifically, emergence makes its re-appearance in the context of the debate about reductionism in the philosophy of mind. Kim is on the right track when he writes: 'The fading away of reductionism and the enthronement of nonreductive physicalism as the new orthodoxy simply *amount* to the resurgence of emergentism' (1999, p. 5, his emphasis). As in the vitalism/mechanism debate, emergence is propelled by *discontent*—in particular by the disquieting suspicion that physicalism perpetuates, or perhaps even generates, the mind-body problem (see, e.g., Kim, 2001). Emergence makes a comeback because of the rising impression that something is wrong with physicalism.





Let me characterise physicalism crudely as follows—well aware that defining it is a philosophical problem in itself. Ontologically, physicalism assumes that everything, including ourselves, is *constituted* by components postulated by physics, which excludes any special (vitalist) ingredients. As this boils down to a roughly materialist position, it is in line with emergentism. However, physicalism assumes also that every (instance of a) property is exhaustively *explainable* in physical terms, together with the basic laws of physics. The idea is that all facts hold in virtue of physical facts alone (see, e.g., Horgan, 1994; Loewer, 2001). This is not in line with emergentism anymore.

Furthermore, physicalism is either reductive or non-reductive. Resembling mechanism in its rigor, *reductive* physicalism regards the mental as reducible to the physical, often in the sense that mental properties are identical with, or 'nothing but', physical properties. (In a weaker sense, reduction means that

the mental is explainable in terms of the physical.) *Non-reductive* physicalism denies the identity of mental and physical properties. Instead, it is suggested that mental properties supervene on physical properties. Here is a rough and ready characterisation of supervenience. For two families of properties *P* and *Q*, *P* supervenes on *Q*, if two things that are indiscernible with respect to *Q* are indiscernible with respect to *P* (cf. Kim, 1998, pp. 9–10). Alternatively, an instance of *P* supervenes on an instance of *Q*, if there can be no change in *P* without corresponding change in *Q* (though there may be *Q*-change without *P*-change). For a while non-reductive physicalists thought that mind-body supervenience could establish a position tight enough to go through as physicalism while being loose enough as *not* to entail reduction. But doubts have been raised recently whether supervenience can in fact substantiate these hopes (see, e.g., Heil, 1998; Kim, 2003a). It is striking that interest in emergence resumed when enthusiasm for supervenience started to subside roughly ten years ago.

### III. The Core Idea of Emergence: a Blend of Ontology and Attitude

I will now turn to what I see as the central features of classical emergentism (see also Crane, 2001; Kim, 2003b). For ease of exposition, I assume that *properties* emerge (rather than substantive souls).

The first feature is the *distinctiveness* of emergent properties. The idea is that an emergent property *P* individuates the thing that has it as something new—a piece of matter that instantiates vitality is no longer a

*mere* piece of matter but *also* an organism (Alexander, 1920, ii, pp. 45f.). As it were, *P* confers a new identity on the thing that has it. Combined with the idea that nothing is real unless it has causal powers—that is, having causal powers is the mark of reality (Alexander, 1920, ii, p. 8)—, *P* is distinctive in the sense that it makes a causal difference to the thing that has *P*, that is, *P* endows it with new causal powers. (Otherwise, we would not even know that something has *P*.) Hence epiphenomenalism is rejected. Distinctiveness also means that the causal powers of emergent properties are irreducible and fundamental. The thought is this: if the causal powers of emergent properties were reducible to (or identical with) the causal powers of their base properties, they would be neither new nor distinct. As a consequence, the causality of emergent properties is inexplicable in terms of, and theoretically unpredictable from, those of their base properties.

The second feature is the *dependence* of emergent properties. The idea that life and mind depend on some basis makes emergentism a naturalist position. Dependence has a temporal (diachronic) and a non-temporal (synchronic) sense. Lloyd Morgan writes: '[I]f by Vitalism we give expression to the fact that living matter has certain distinctive properties, it may be freely accepted; but [...] if by it we imply that these properties neither are nor can be *the outcome of evolution*, it should be politely rejected [...]' (1899, p. 196, my emphases). The message is clear: we should renounce miraculous extra-bits in our cosmology, if only because we have no idea of how matter, life, and mind relate (see Lloyd Morgan, 1923, pp. 12f.). For emergentism, all there is has originated *from within*, or, to use

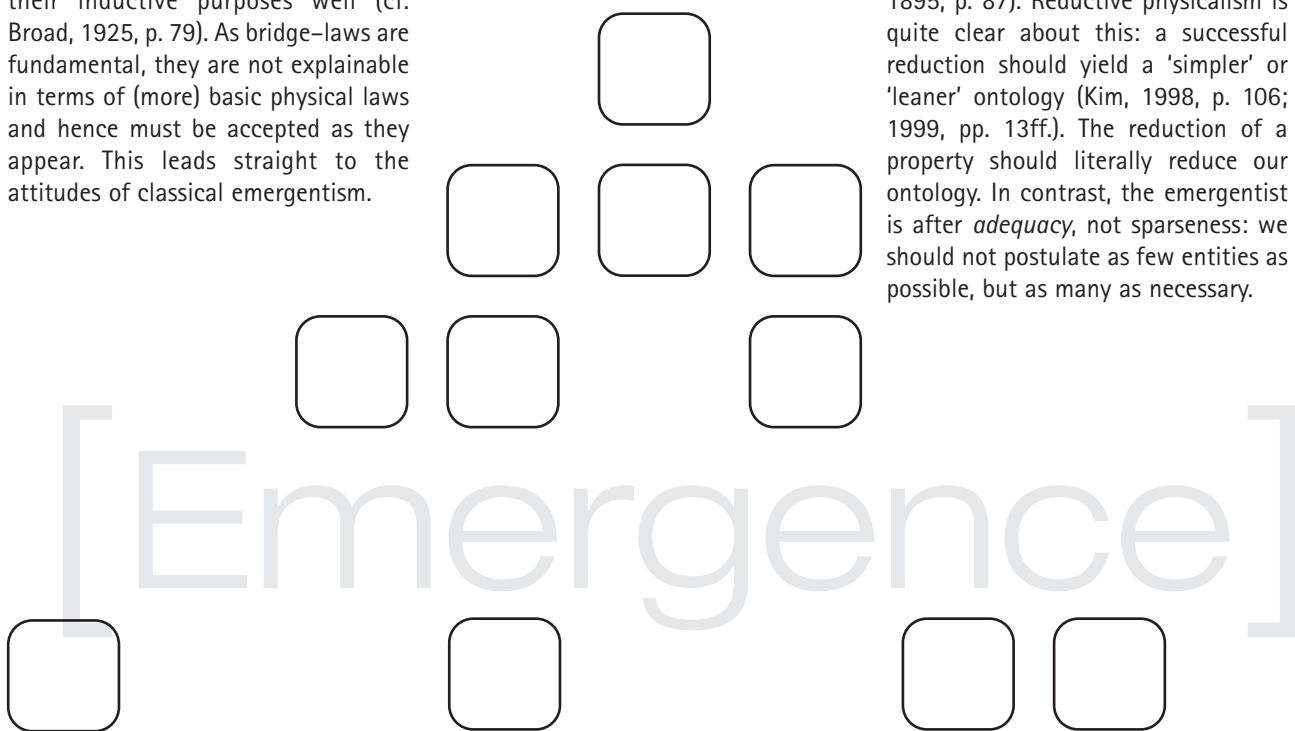
Alexander's phrase, 'blossomed out' of, nature (1920, ii, p. 52). The idea that everything is ultimately made of some one stuff is compatible with mechanism and physicalism. However, if all change is the result of a regrouping of components that have a few intrinsic properties such as mass or charge and are governed by a couple of basic laws, then it is hard to see how anything new could ever appear (cf. Broad, 1925, p. 76; Lloyd Morgan, 1923, p. 2).

Dependence also conveys the thought that an emergent property *P* cannot be (instantiated) unless the thing that has *P* also has certain other properties that serve as *P*'s 'emergence base'. So *P* is existentially dependent on, and supervenes on, its base properties. Supervenience captures the idea that there is a nomological link between *P* and its base—that there are so-called 'bridge-laws' that relate emergent properties with their base properties. In spite of the fact that they are primitive and basic, such laws serve their inductive purposes well (cf. Broad, 1925, p. 79). As bridge-laws are fundamental, they are not explainable in terms of (more) basic physical laws and hence must be accepted as they appear. This leads straight to the attitudes of classical emergentism.

First of all, there is epistemic modesty and humility. Alexander writes: 'The existence of emergent qualities [...] is something to be noted, as some would say, under the compulsion of brute empirical fact, or, as I should prefer to say in less harsh terms, to be accepted with the "natural piety" of the investigator. It admits no explanation' (1920, ii, pp. 46f.). There are questions, he thinks, 'we have no right to ask' (1922, p. 620), as there are facts we must accept in the mood of 'reverent temper' (*op. cit.*, p. 621). Natural piety, in other words, is 'the habit of knowing when to stop in asking questions of nature' (*op. cit.*, p. 609). By embracing a stance of 'deliberate innocence', or 'strenuous naïveté', we prevent ourselves from asking silly questions or filling the explanatory void with some purpose-built entity (Alexander, quoted by Muirhead, 1939). When faced with explanatory and purported ontological gaps, emergentism implores us to stay cool and relax.

Second, there is a resistance to excessively optimistic scientism and explanatory chauvinism. The successes of mechanistic explanation and the reductive strategy do not carry over to the cases of life and consciousness. Aiming at explaining the complex in terms of the simple certainly is a valuable heuristic. But its downside is a tendency to over-simplify facts. As Broad sharply remarks of mechanism, it 'of course reaches its wildest absurdities in the attempts [...] to treat mental phenomena mechanistically' (1925, p. 77). Besides, as participants in the vitalism/mechanism debate had access to the *same* biological knowledge, their radically different outlooks must have been driven not by facts, but by certain explanatory pre-conceptions. Emergentism tells us not to let our metaphysics be compelled by our favourite explanatory strategies.

Third, there is a certain dislike for obsessive ontological tidiness or sparseness (see, e.g., Lloyd Morgan, 1895, p. 87). Reductive physicalism is quite clear about this: a successful reduction should yield a 'simpler' or 'leaner' ontology (Kim, 1998, p. 106; 1999, pp. 13ff.). The reduction of a property should literally reduce our ontology. In contrast, the emergentist is after *adequacy*, not sparseness: we should not postulate as few entities as possible, but as many as necessary.



There is also wariness about pretentious theories which tend to equalise or belittle the variety of being, rather than saving the differences. For this reason, it is perhaps right to attribute to emergentism an inclination for metaphysical pluralism.

The last feature is tactical. As reductionism has *not* shown that all properties are, or must be, reducible, emergence is at least not an absurdity (cf. Broad, 1933). Let the opponents of emergence show that there are *no*, or cannot be, emergent properties. In a similar vein, modern physicalists beseech us to accept the following dilemma: *either* you save mental causation at the price of reductionism, *or* you save the distinctiveness of the mental at the price of its causal efficacy (see Kim, 1998). Emergentism is the strategic refusal to play along with such either-or-isms.

#### IV. Emergence at Work: a Twofold Problem for Supervenience

I will now illustrate how emergentism, as a blend of ontological theory and attitude, stimulates the re-thinking and re-shaping of the debate about non-reductive physicalism. In particular, it brings to the fore a twofold problem with supervenience: supervenience alone cannot separate non-reductive physicalism from emergentism, and supervenience eludes an explanation in physical terms.

I take it that a primary motivation in favour of physicalism is a worry about mental causation: if we want mental properties to make a causal difference to the thing that has them, and in some sense bring about *physical*

effects, we must 'attach' them to physical properties. In the background, there is the thesis of *causal completeness* (or causal closure), which says that every physical effect has a (sufficient) physical cause. In other words, physical effects do not have (non-redundant) non-physical causes. Reductive physicalism strongly attaches mental properties to physical properties by identifying them—where we thought there were two properties, in fact there is one. So reductionists think that mental properties have causal powers because they *are* physical properties, and thus get a handle on mental causation. However, they now need to explain (away) the apparent difference between mental and physical properties, for reducing the mental makes it disappear as something unique and special. As Kim aptly puts it: 'In what sense, then, have we saved *mental* causation?' (1996, p. 237, his emphasis). This is especially worrying if we assume that phenomenal or qualitative aspects are essential to the mental.

In contrast, non-reductive physicalists hold that mental properties *supervene* on physical properties without being reducible or identifiable with them. They think that mind-body supervenience secures a minimal physicalism, because supervenience incorporates mind-body dependence and hence the primacy of the physical. Here is the first problem: emergentism also accepts mind-body supervenience and thus seems to *coincide* with non-reductive physicalism (or minimal physicalism). Both emergentists and non-reductive physicalists are confident they can keep mental properties irreducibly distinct from, yet somehow grounded in, physical properties. Together they resist being wheedled into accepting that *As* really are 'nothing but' *Bs*; or that *As* are

much better understood in terms of *Bs*; or that the *Bs* explain the *As*; or that there are no *As* as we know them. Because of this, some philosophers think there are no, or only marginal, differences between emergentism and non-reductive physicalism (see, e.g., Crane, 2001; Kim, 1993; Pereboom, 2002). Supervenience is too weak a criterion to separate emergentism from non-reductive (minimal) physicalism.

The second problem with supervenience is this: why does mind-body supervenience hold? How does physicalism explain the covariation of mental and physical properties? As mentioned, emergentism accepts the nomic link between emergent properties and their base-properties as fundamental, and hence mind-body supervenience as unexplainable, which makes (strong) emergence unattractive to physicalism. The reason is a further physicalist thesis (mentioned above), *viz.* the thesis of *explanatory completeness*, which says that every physical fact has an explanation in exclusively physical terms (see, e.g., Kim, 2001; Papineau, 2001). That is, the physical realm is explanatorily self-sufficient. In itself, this is a fairly innocuous claim, warranted by the explanatory success of physics. The aspiration of physicalism, however, is to account also for *non-physical* facts; mind-body supervenience is a case in point, as this is not a purely physical fact. As long as there is no robust physical explanation of why and how mind-body supervenience holds, physicalists should worry about explanatory completeness (see Horgan, 1993). If this is right, not only is *unexplained* supervenience too weak to delineate minimal physicalism, it is in fact a threat to physicalism, as it *undermines* the idea

that an explanation in purely physical terms gives the *full* story.

If physicalism takes bridge-laws to be a consequence of physical laws, and hence in principle deducible and *explainable* by those laws, then mind-body supervenience must be explainable in terms of basic physical laws. If we fail to explain the covariation of *P* and *Q*, this leaves supervenience an unexplainable brute fact (Kim, 1998, p. 96). Of course, reductive physicalists go on to say that the *reduction* of *P* to *Q* easily explains why *P*-instances covary with *Q*-instances: *P* supervenes on *Q* because *P* and *Q* are one and the same (Kim, 1999, pp. 13ff.). A physicalist who flinches from reduction, however, is in the same boat as the emergentist: bridge-laws *describe*, but do *not explain*, the covariation of properties. In a nutshell, the worry is this: if supervenience eludes a physical explanation, yet explanatory completeness is a defining feature of physicalism, this entails that non-reductive physicalism is not physicalism anymore—but, again, coincides with emergentism.

Since unexplained supervenience is accepted as a 'brute fact', and, in contrast to physicalists, they do not subscribe to the thesis of explanatory completeness, emergentists have peace of mind. It is not *absurd* for emergentism to deny explanatory completeness, because even physicalists admit that it is a contingent thesis distilled from years of inductive science (cf. Papineau, 2001). The vigour with which they cling to completeness, however, suggests that physicalists conflate the *denial* of completeness with the *acceptance* of supernaturalism, that is, to go along with unexplainable ontological dwellers. While emergentism denies the explanatory completeness of physics, it defends the idea that novelty arises *from within* nature. After all, its quintessential idea is that the emergence of novelty is a diachronic process building on material, so to speak, which already is in nature (or the physical realm) (see Lowe, 2000). If my inkling is right, these two facts are recognised thanks to emergence: physicalism cannot be *founded* on supervenience which is congruous with the denial of explanatory completeness, and non-reductive physicalism needs something other than supervenience to *remain* physicalism.

## U. So, Why Does Emergence Reappear?

There is a surprising twist for reductive physicalism. The standard Nagel (1961) model of reduction based on bridge-laws does *not* require them to be explainable, and is therefore compatible with emergentism (see, e.g., Kim, 1999, p. 12). But if reductive physicalism based on bridge-laws is compatible with emergentism, and these positions *oppose* each other, then standard reduction cannot be the hallmark of reductive physicalism. In other words, reductive physicalism must operate on a stricter model of reduction. Ironically, an alternative model of reduction offered by Kim (1998, ch. 4) fails to get a grip on phenomenal properties, or the intrinsic character of conscious experience (i.e. *qualia*). In this situation it is awfully painless to call emergence to the rescue (Kim, 1998, p. 103; see also Stephan, 1999a, p. 195). The strategy is obvious enough: move from failing to reduce a property *P* (on any model of reduction) to declaring *P* emergent. What generates this pressure to take *qualia*-emergentism seriously, I suggest, is the insistence on an explanation of mind-body supervenience, and the recognition that this can only be done by identifying phenomenal properties with physical properties—which even physicalists feel is counterintuitive (cf. Kim, 1998, p. 117). But if emergentism about *qualia* is considered an option, the situation for physicalism must be despondent indeed.

Emergence

In a revealing passage, Kim says: 'But if a whole system of phenomena that are prima facie not among basic physical phenomena resists physical explanation, and especially if we don't even know where or how to begin, it would be time to reexamine one's physicalist commitments' (1998, p. 96, my emphasis). This is precisely the impact of emergentism: shaping awareness for the difficulties that plague an overarching physicalist worldview. The idea of emergence works as a corrective, and it plays the role, as it were, of a jester: not by direct proposals, but by bringing others to reflect on their own positions. As doubt is a fertile ground for emergence, the idea works better on physicalists (like Kim) who self-confessedly struggle with the heritage of neo-positivism that still 'constrains' their thinking (see Kim, 1998, p. 2).

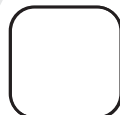
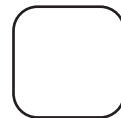
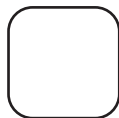
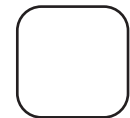
To sum up. I suggest that part of the *appeal* of emergence is its expression of an attitude of metaphysical humility, which is fostered and motivated by discontent or disquiet about extreme positions such as physicalism. In turn, some of the re-thinking in the current metaphysics of mind is due to the resilient reappearance of emergence. The 'metaphysical discomfort' (see Kim, 1999, p. 28) that some physicalists feel towards emergence does not reflect troublesome aspects of emergence, but reveals doubts about physicalism itself.

## References

- Alexander, S. (1920). *Space, Time, and Deity. The Gifford Lectures at Glasgow 1916-1918 (2 Vols.)*. London: Macmillan.
- Alexander, S. (1922). Natural Piety. *The Hibbert Journal*, 20, 609-621.
- Broad, C. D. (1925). *The Mind and its Place in Nature*. London: Kegan Paul.
- Broad, C. D. (1933). *Examination of McTaggart's Philosophy, Vol. I, 'The Nature' of a Continuant'* (pp. 264-278). Cambridge: Cambridge University Press.
- Crane, T. (2001). The Significance of Emergence. In C. Gillett & B. Loewer (Eds.), *Physicalism and Its Discontents* (pp. 207-224). Cambridge: Cambridge University Press.



- Gillett, C. (2002). The Varieties of Emergence: Their Purposes, Obligations, and Importance. *Grazer Philosophische Studien*, 65, 95-121.
- Heil, J. (1998). Supervenience Deconstructed. *European Journal of Philosophy*, 6, 146-155.
- Hoernle, R. F. A. (1918). Mechanism and Vitalism. *The Philosophical Review*, 27, 628-645.
- Horgan, T. (1993). From Supervenience to Superdupervenience: Meeting the Demands of a Material World. *Mind*, 102, 555-586.
- Horgan, T. (1994). Physicalism. In S. Guttenplan (Ed.), *A Companion to the Philosophy of Mind* (pp. 471-479). Oxford: Blackwell.
- Kim, J. (1993). The Non-Reductivist's Troubles with Mental Causation. In J. Heil & A. Mele (Eds.), *Mental Causation* (pp. 189-210). Oxford: Oxford University Press.
- Kim, J. (1996). *Philosophy of Mind*. Boulder, CO: Westview Press.
- Kim, J. (1998). *Mind in a Physical World*. Cambridge, MA: MIT Press.
- Kim, J. (1999). Making Sense of Emergence. *Philosophical Studies*, 95, 3-36.
- Kim, J. (2001). Mental Causation and Consciousness: The Two Mind-Body Problems for the Physicalist. In C. Gillett & B. Loewer (Eds.), *Physicalism and Its Discontents* (pp. 271-283). Cambridge: Cambridge University Press.
- Kim, J. (2003a). Supervenience, Emergence, Realisation, Reduction. In M. J. Loux & D. W. Zimmerman (Eds.), *The Oxford Handbook of Metaphysics* (pp. 556-584). New York: Oxford University Press.
- Kim, J. (2003b). *Being Realistic about Emergence*. Unpublished Manuscript.
- Lloyd Morgan, C. (1895). Naturalism. *The Monist*, 6, 76-90.
- Lloyd Morgan, C. (1899). Vitalism. *The Monist*, 9, 179-196.
- Lloyd Morgan, C. (1923). *Emergent Evolution. The Gifford Lectures at St. Andrews 1922*. London: Williams and Norgate.
- Loewer, B. (2001). From Physics to Physicalism. In C. Gillett & B. Loewer (Eds.), *Physicalism and Its Discontents* (pp. 37-56). Cambridge: Cambridge University Press.
- Lowe, E. J. (2000). Causal Closure Principles and Emergentism. *Philosophy*, 75, 571-585.
- Muirhead, J. H. (1939). Samuel Alexander. *Philosophy*, 14, 3-14.
- Nagel, E. (1961). *The Structure of Science. Problems in the Logic of Scientific Explanation*. London: Routledge & Kegan Paul.
- O'Connor, T., & Wong, H. Y. (2002). *The Metaphysics of Emergence*. Unpublished Manuscript.
- Papineau, D. (2001). The Rise of Physicalism. In C. Gillett & B. Loewer (Eds.), *Physicalism and Its Discontents* (pp. 3-36). Cambridge: Cambridge University Press.
- Pereboom, D. (2002). Robust Nonreductive Materialism. *Journal of Philosophy*, 99, 499-531.
- Shelley, M. (1818). *Frankenstein, or the Modern Prometheus*. Oxford: Oxford University Press (Reprint 1992).
- Silberstein, M. (2002). Reduction, Emergence, and Explanation. In P. Machamer & M. Silberstein (Eds.), *The Blackwell Guide to the Philosophy of Science* (pp. 80-107). Oxford: Blackwell.
- Stephan, A. (1999a). *Emergenz. Von der Unvorhersagbarkeit zur Selbstorganisation*. Dresden: Dresden University Press.
- Stephan, A. (1999b). Varieties of Emergentism. *Evolution and Cognition*, 5, 49-59.
- Van Gulick, R. (2001). Reduction, Emergence, and Other Recent Options on the Mind/Body Problem. *Journal of Consciousness Studies*, 8, 1-34.



[Emergence]

## Sam Coleman

# [ Two Aspects of Emergence: A Response ]

In the previous paper Peter identifies emergence as an attempt to find a middle road between reductive physicalism on the one hand, and dualism on the other. This attempt, he says, is fuelled by an attitude of discontent; a discontent directed primarily, I guess, at the ruthless desire to cram all existents into a reductive physicalist ontology. But the discontented emergentist also finds cause to complain about the dualist alternative.

The dualist, it is felt, is a bit metaphysically extravagant, positing two radically different kinds of property (I follow Peter in sticking to property-talk) where arguably just one kind will do.

The problems all stem, of course, from the fact that reducing the mental to the physical seems to many to be impossible. Philosophers of all kinds of stripe are moved by the thought that no matter how comprehensive our knowledge of, say, the physical workings of the visual system, nothing will follow about the character of the experiences being undergone by the owner of the system. And this fact has created a drive, as Peter notes, to find an ontology that can be seen to be adequate; rich enough to contain mentality, and not merely as pared down and austere as possible.

On this description of the motivations that have prompted a move away from strictly reductive physicalism, I agree

with Peter wholeheartedly. And it certainly seems true that, historically, in the debate between vitalists and mechanists which paralleled that between dualists and physicalists today, it was emergentism that tried to steer a middle course between the extremes of unnecessary economy and overblown metaphysics. I guess my worry and confusion is over where the motivations, the discontent, should lead us today. I'm not sure that there now exists a gap in the market that can be uniquely occupied by emergence. It is this thought that I'm going to probe a little.

We currently have a very popular position that tries to take a middle line between reductive physicalism and dualism. This is non-reductive physicalism (NRP). A non-reductive physicalist is free to hold, with Peter's emergentist, that the mental supervenes on the physical, whilst holding that there is no possible reductive explanation of the mental in terms of the physical. And, also just like emergentism as Peter characterised it, according to NRP what is going to keep the mental supervenient on the physical, in the absence of any explanatory relation between the two, are some fundamental, that is brute, psychophysical laws. Finally, again to the emergentist's liking, we can hold on this picture that mental properties have distinctive causal powers. They have causal powers by virtue of being

physical properties, and explanations featuring such properties will be indispensable: it won't be possible to give equivalent explanations in terms of lower order properties because of the failure of reduction.

This NRP looks a lot like emergentism as Peter describes it. It stems from the same discontent with the extreme positions, and it defends the commitments important to the emergentist. Here there is only one kind of thing in nature; nature is 'all joined up', we might say, in that there are only physical properties at the high level just as at the low level. Moreover there is a failure of reduction, but this is not seen to imply epiphenomenalism about the mental. The mental remains distinctively causal.

So my first question for Peter is: does he see any daylight between this form of NRP and emergentism as he understands it?

I guess it's possible that NRP might not be metaphysically 'punchy' enough for Peter. One question that has always come up for the non-reductive physicalist is how to account for the fact that the mental is physical. We just can't see any kind of entailment between them, or any explanation of the mental in physical terms. Frank Jackson puts this poser in a typically eye-catching way:

It is implausible that there are facts about very simple organisms

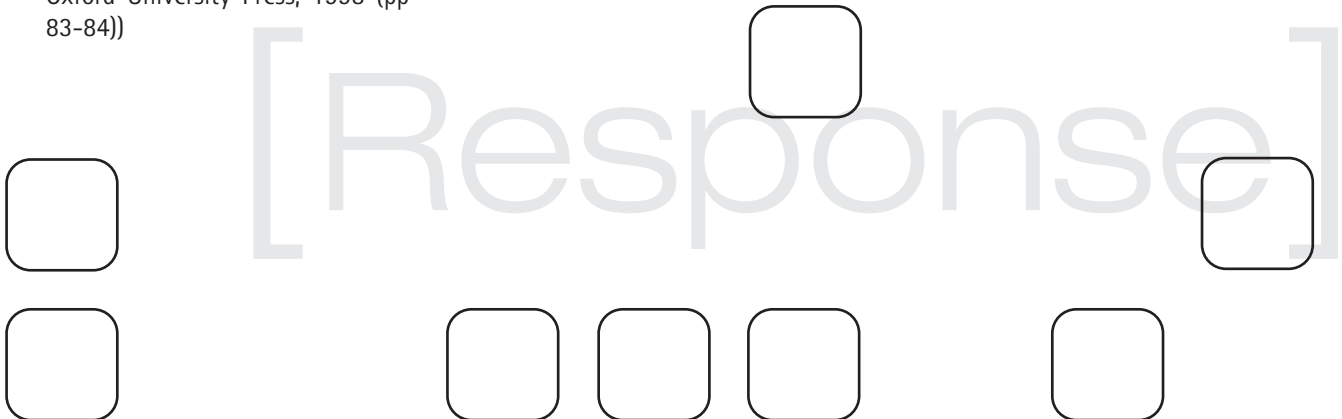
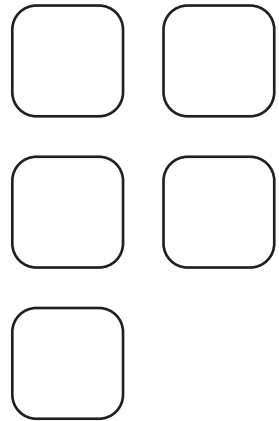


that cannot be deduced a priori from enough physical information about their physical nature and how they interact with their environments, physically described. The physical story about amoebae and their interactions with the environment is the whole story about amoebae...But according to materialism we differ from amoebae essentially only in complexity of ingredients and their arrangement. It is hard to see how that kind of difference could generate important facts about us that in principle defy our powers of deduction...Think of the charts in biology classrooms showing the evolutionary progression from single-celled creatures on the far left to the higher apes and humans on the far right: where in that progression can the physicalist plausibly claim that failure of a priori deducibility of important facts about us emerges? Or, if it comes to that, where in the development of each and every one of us from a zygote could the materialist plausibly locate the place in which there emerge important facts about us that cannot be deduced from the physical story about us? (*From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford: Oxford University Press, 1998 (pp 83-84))

The thought is that the defender of NRP owes us a story as to why it should be opaque that the mental is entailed by the physical, if it is indeed so. There have been appeals to a posteriori necessity, to the idea that we couldn't even see a priori that, say, water was H<sub>2</sub>O, as in fact it is. This tactic has been largely dropped in recent writing, with non-reductive physicalists acknowledging that more needs to be said, since mentality seems to be in a uniquely intractable position as regards any entailment between it and the physical.

Recent accounts have tried to offer an *epistemological* solution to the opacity. Many holders of NRP now claim that there is something distinctive about the concepts we use when talking about mentality, in particular consciousness. For example one thought might be that, due to evolution, mental concepts and physical concepts have developed to play radically different cognitive roles, with the result that there are no conceptual connections between them at all. This might mean that, even when faced with a true identity, imagine c-fibre firings = pain is one, we wouldn't be able to see that the mental and physical concepts flanking the identity co-referred. We would never see that the presence of c-fibre firings implied that of pain and vice-versa.

This sort of move is now very popular, but as I say, as an explanation of the lack of entailment between physical and mental, I'm not sure it is metaphysical enough. The move suggests that in *nature* there *is* an entailment between physical and mental: mental properties really are systematically and logically the upshot of lower-level physical stuff and its properties. It's just that, because of the nature of our concepts, we are barred from ever seeing this. Indeed the position suggests that there could be beings, with concepts different from our, who could see the physical to mental entailment. This is as much as to say that we are contingently cognitively barred from being reductivists, but that reductivism is really the position we ideally ought to aim to hold.



The gap between physical and mental allowed by such a theory is just an epistemological one, perhaps even merely a contingently human one. This does not seem to do justice to Peter's thoughts when he says that mental properties are to be kept 'irreducibly distinct from physical properties', that 'it is...right to attribute to emergentism an inclination for metaphysical pluralism'. Perhaps an even more telling way to bring out how the defender of NRP using the epistemological move may not be what Peter wants is to think about causation. Peter says: 'the causal powers of emergent properties are irreducible and fundamental. The thought is this: were the causal powers of emergent properties reducible to...the causal powers of their base properties, they would be neither new nor distinct'. But this epistemological brand of NRP precisely does hold that mental properties' causal powers are reducible to those of their physical base. They are reducible in the metaphysical sense that the physical base and its properties entail the whole story about what mental properties are instantiated. This entailment remains opaque to us, and we can make no predictions from physical data alone. Causal explanations employing mental concepts remain indispensable, but it is still not true that mental properties have radically distinct and *fundamental* causal powers, metaphysically speaking.

So I don't think that this NRP makes mental properties quite *distinct enough* for Peter, thus arguably it is not quite the same as the emergentism he described in his paper. The question that presses now, then, is just what kind of picture would make mental properties distinct enough for Peter to class it as emergentism?



I find this territory tricky and confusing. We need a picture where emergent mental properties are not just epistemologically, but *metaphysically* novel when compared with their subvenient physical bases, whatever this means. Such a picture would look, epistemologically, just like the NRP considered just now. That is, we would be unable to go beyond some brute, fundamental, psychophysical laws in expressing the relationship between mental and physical. But this time, rather than these laws merely arising from a cognitive impairment of ours, or some peculiarity of our concepts, they would actually reflect the way mental and physical interacted in reality. If this is emergentism, it is a strange view. It is plausible that, setting mentality aside, the macro properties of the world are entailed by the nature of the lower level physical. We have good evidence that this is the case with successful reductions like that of

heat to molecular kinetic energy. The emergentist as portrayed would have, in the face of future successful reductions, to be willing to endorse the existence of a metaphysical gap solely between the physical and the mental. They may be willing to hold this, those who have given up on the emergence of life holding onto mentality as the last bastion of irreducible reality. But it is reasonable to ask how the emergentist can account for the existence of this unique metaphysical gap. A dilemma lurks here which threatens to push the emergentist either into the arms of the epistemological NRP, or towards full blown dualism. In any case it is a move away from holding a stable and distinctive position.

The epistemological NRP has a neat explanation of the uniqueness of mental/physical relations. Really there is nothing unique about mental physical relations; the mental is



The real issue for me at this point is what the emergentist gains by distancing herself from *dualism*. I can see why she wants to avoid epistemological NRP, because it doesn't make mental properties robustly their own kind of distinctive thing. The difference between the dualist and the emergentist as we've set things up is that the dualist says the mental is metaphysically distinct and non-physical, while the emergentist says the mental is metaphysically distinct but still physical, or at least, still natural.

What is gained by insisting on hanging on to the claim that the mental is of a piece with the physical or the natural? The initial motivation for resisting dualism was that we needed no intruders into nature. But the dualist can just say that the mental is perfectly *natural*, just that nature turns out to contain two kinds of property, mental and physical. So the disagreement seems to turn on the retention or not of the word 'physical'. Does the emergentist want to hang on to the thought that the mental is of a piece with the physical? If no, then I suggest she happily acquiesces in dualism. If yes, then two questions face her. 1. What is the benefit, and what is the content of saying that the mental is one with the physical? And 2. How can she elucidate the oneness of the mental with the physical, alongside the claimed metaphysical distinctiveness of the mental? Perhaps there are answers to these questions, but it does look like a hard road.



# Notes On [Contributors]

## Sam Coleman

is pursuing a doctorate at Birkbeck College London. His main interests are in metaphysics and the philosophy of mind, and the main focus of his research is in the area of consciousness.

## Mark Sinclair

teaches philosophy at The Manchester Metropolitan University and The University of York. His research is concerned with phenomenology, Greek philosophy and the history of aesthetics.

## Peter Wyss

is a graduate student at Birkbeck College London. His main areas of interest are metaphysics and the philosophy of mind. He is completing a doctoral thesis on emergence.

## James Hill

lectures at Charles University, Prague. He studied at Oxford, Geneva and King's College, London, and has published a range of articles on the empiricists. At the moment he is preparing a book on John Locke's critique of mechanism and working on Descartes' concept of thinking. He is also collaborating on the translation of several English-language philosophers into Czech.

## Dan O'Brien

is lecturer in philosophy at the University of Birmingham. His main research interests lie in epistemology and the philosophy of mind. He is currently working on two books: *An Introduction to Epistemology*, Polity Press, Cambridge (forthcoming 2005), and *Readers Guide to Hume's Enquiry Concerning Human Understanding*, Continuum, London (with A. Bailey) (forthcoming 2006). He has been published in *Philosophy*, *Philosophy in Review*, and *The Internet Encyclopaedia of Philosophy*.

## Paul Sperring

is joint editor of the RJP.

# [Contributors] Notes For

## Content

We welcome articles on any area in philosophy. Papers may be broad or narrow in their focus (for instance a discussion of the mind/body problem, or an analysis of Hume's treatment of causation in the *Enquiry*). We would particularly encourage contributions which reflect original research on the following philosophical themes: epistemology, metaphysics, philosophy of religion, ethics, philosophy of mind, philosophy of science, political philosophy, religious ethics; and texts, such as: *The Republic*, *The Nicomachean Ethics*, *The Meditations*, *An Enquiry Concerning Human Understanding*, *Beyond Good and Evil*, *On Liberty*, *Existentialism and Humanism*, *The Problems of Philosophy*, *Language Truth and Logic*.

The articles should be around 3000-4000 words.

## Style

The language used in the articles should be as non-technical as possible whilst preserving the richness of the arguments. Where technical terms are unavoidable they should be explained and examples offered.

Notes should be presented as endnotes. Textual references should be presented in the following format: Barry Stroud, *Hume* (London: Routledge, 1977), 77-91.

## Presentation

Articles should be written in *Word* (any version).

## Contributions

Articles for this journal are currently written by a panel of philosophers from a variety of universities in Britain, Australia and the United States, whose work is edited by the journal's editorial board. We invite unsolicited contributions from philosophers working in any field. The contributions should be submitted via email attachment to [rjp@rutc.ac.uk](mailto:rjp@rutc.ac.uk)

## Copyright

The RJP retains the option of reprinting published articles in later RJP publications. Authors may republish articles with the journal's permission provided that they acknowledge that those articles were first printed in the RJP. Papers should only be submitted if the author is willing and able to be bound by the conditions set out in this paragraph.

# [Contributors]

# Richmond upon Thames College

is a large further education college located in  
Twickenham, West London  
offering 16-19 students one of the widest  
curriculum choices in the country.

Last year, we came top of all  
London further education colleges  
in the Times league tables  
and we are proud of our reputation  
for achieving excellent results year after year.

We are well known nationally for  
our high quality staff, excellent student support systems  
and the inclusive education we offer to all our students.

If you would like to find out more about us -  
please contact our Course Information Unit on

**020 8607 8305 / 8314**

or visit our website on

**[www.rutc.ac.uk](http://www.rutc.ac.uk)**



**Richmond upon Thames College**



# [Subscribing]

to the **RJP**

The RJP comes out in Autumn, Spring and Summer. To subscribe you need to select the appropriate price from the table and complete the mailing information at the bottom of the form. All prices include post and packaging. The bottom half of the form should then be detached and sent with a cheque made payable to Richmond upon Thames College to the address below:

RJP Subscriptions  
 Philosophy Department  
 Richmond upon Thames College  
 Egerton Road  
 Twickenham  
 London TW2 7SJ  
 United Kingdom

Please allow one week for delivery in the UK, and two weeks for the rest of the world.

 **Annual Subscription : Current Rates**

Institutional Subscriber in the UK	£33.00
Individual Subscriber in the UK	£18.00
Institutional Subscriber from the rest of the EU	63.00 Euros
Individual Subscriber from the rest of the EU	40.50 Euros
Institutional Subscriber outside the EU	\$67.50 US
Individual Subscriber outside the EU	\$45.00 US

Resubscription YES / NO From issue .....

Your name: ..... Your organisation's name (if appropriate) .....

Address .....

..... Post code/Zip code.....

Telephone number (including full international dialling code) .....

Email address .....

I enclose a cheque for the amount of ..... to purchase an annual subscription

Signed ..... Date .....



[Philosophy]



The Richmond Journal of **Philosophy**

family life **Dan O'Brien** 5

descartes **James Hill** 11

nietzsche **Mark Sinclair** 18

reductive physicalism **Paul Sperring** 24

two aspects of emergence **Peter Wyss** 32

reply to Peter Wyss **Sam Coleman** 40

Richmond Journal of Philosophy  
Philosophy Department  
Richmond upon Thames College  
Egerton Road  
Twickenham  
Middlesex  
TW2 7SJ  
United Kingdom

Tel: 020 8607 8000  
Fax: 020 8744 9738  
Email: [rjp@rutc.ac.uk](mailto:rjp@rutc.ac.uk)  
[www.rutc.ac.uk/rjp](http://www.rutc.ac.uk/rjp)



Richmond upon Thames College

ISSN: 1477-6480