

Cosmic Hermeneutics vs. Emergence: The Challenge of the Explanatory Gap

Tim Crane

ABSTRACT

This paper is a defence of Terence Horgan's claim that any genuinely physicalist position must distinguish itself from (what has been traditionally known as) emergentism. I argue that physicalism is necessarily reductive in character -- it must either give a reductive account of apparently non-physical entities, or a reductive explanation of why there are non-physical entities. I argue that many recent 'non-reductive' physicalists do not do this, and that because of this they cannot adequately distinguish their view from emergentism. The conclusion is that this is the real challenge posed by Joseph Levine's 'explanatory gap' argument: if physicalists cannot close the explanatory gap in Levine's preferred way, they must find some other way to do it. Otherwise their view is indistinguishable from emergentism.

1. The explanatory gap

Joseph Levine is generally credited with the invention of the term 'explanatory gap' to describe our ignorance about the relationship between consciousness and the physical structures which sustain it.¹ Levine's account of the problem of the explanatory gap in his book, *Purple Haze* (2001), may be summarised in terms of three theses, which I will describe and name as follows:

(SP) *Supervenience physicalism*: Every minimal physical duplicate of the actual world is a duplicate in every respect.

(DE) *Deductive Explanation*: the explanation of consciousness must consist in a deduction of the truths about consciousness from the physical truths.

(EG) *Explanatory gap*: we lack an adequate deductive explanation of all the truths about consciousness in physical terms.

SP, DE and EG are not inconsistent, since consciousness could supervene on the physical without its being explicable. Consciousness might be wholly determined by the physical but nonetheless be inexplicable. This is what 'mysterians' believe. Nonetheless, Levine thinks (and many agree with him) that there is a challenge here which physicalism has to meet. If physicalism is to be an adequate account of the world, it must not postulate too many 'brute' or inexplicable correlations and identities. Accepting a mere brute correlation is accepting a mystery. In this he is echoing Thomas Nagel, who said famously that someone who asserts that consciousness is a process in the brain would be in the same epistemological position as an ancient Greek who asserted that matter is energy: they would have said something true, but they would not have understood how it could be true (Nagel 1974).

That there is an explanatory gap (in the sense of EG) is surely undeniable. The question is what its significance is. Some physicalists think that it is of little significance, because it is only a matter of time before we arrive at an explanation. Others think it is of little significance because physicalism does not require such an explanation of consciousness. However, others think that with some additional

¹ See Levine 1983, 1997, 2001. For a discussion of some interesting historical precursors, see Tennant forthcoming.

assumptions, the explanatory gap shows that physicalism is false. In this paper I will argue that the real significance of the explanatory gap lies elsewhere. In my view, the explanatory gap creates a challenge for the proper formulation of *non-reductive* versions of physicalism. Non-reductive physicalists have attempted to distinguish their theories from reductive versions of physicalism on the one hand, and ‘emergentism’ on the other. But their standard response to the explanatory gap fails to distinguish their non-reductive physicalism from emergentism. Since whatever emergentism is, it is not physicalism, this is a problem for the proper formulation of physicalism. What the explanatory gap shows is what is properly required from an adequate non-reductive physicalism.

In section 2 I will sketch some physicalist responses to the challenge of the explanatory gap; in section 3 I will give an account of emergent properties; in section 4 I will argue that physicalism is necessarily reductive in character; and in section 5 I will show how the explanatory gap challenges so-called ‘non-reductive’ versions of physicalism to distinguish themselves from emergentism.

2. Physicalism and cosmic hermeneutics

Like Nagel before him, Levine did not argue that the explanatory gap shows that physicalism is false. As just noted, EG, DE and SP are mutually consistent. Nagel’s view was that physicalism is true, but that we cannot fully understand it. Levine, similarly, thinks that physicalism is explanatorily inadequate until EG is shown to be false. The explanatory gap is precisely that: an *explanatory* gap. There is no *metaphysical* gap, on the Levine-Nagel view. But it is only if there is a metaphysical gap that physicalism is false.

Frank Jackson (1998) rejects this account of the situation. Jackson thinks that a serious version of physicalism needs not just SP, but an explanation of SP itself. SP makes a claim about the necessary connection between the physical truths and all the other truths. Jackson argues first that this necessity cannot be ‘brute’; and secondly that it must be explicable as conceptual necessity, that is, as the necessity deriving from relationships among concepts. The now well-known picture that Jackson arrives at is that all truths stated in non-physical vocabulary must in principle be derivable from the physical truths plus conceptual analyses of the non-physical concepts. Jackson’s thesis therefore implies what Terence Horgan (1983) calls ‘cosmic hermeneutics’: that full knowledge of the microphysical supervenience base would allow us in principle to ‘read off’ all the truths about the world from the truths about the supervenience base (see also Byrne 1999). Understood in Jackson’s way, cosmic hermeneutics is the following doctrine:

(CH) *Cosmic hermeneutics*: there can be an *a priori* deduction of all the truths about the world from the microphysical truths about it plus the conceptual truths about non-physical concepts.

So stated, CH is a claim about possibility: it says merely that there *can* be such an *a priori* deduction of all the truths, not that anyone has actually done it. Jackson argues that if physicalism is true, then CH is true. Physicalism implies cosmic hermeneutics.

An over-simplified example can illustrate. Suppose that we achieve a complete analysis of the concept of water, in terms of concepts like *wet*, *transparent*, *odourless* etc. And let’s suppose that we have an analysis of these concepts in terms of the causal roles of the things to which they apply: we have a causal analysis of the

concept of wetness, for example. Let's abbreviate our analysis of the concept of water to *the watery stuff*. Equipped then with a complete physical knowledge of the world – knowledge of all the particular matters of physical fact, and the physical laws – we will recognise the truth of the proposition:

(a) H₂O covers most of the earth

Given our grasp of the concept of *the watery stuff*, and our knowledge of physics and chemistry, we can also recognise the truth of the proposition:

(b) H₂O = the watery stuff

And now given our conceptual analysis of *water*, we can then deduce, without any need for any further empirical investigation:

(c) Water covers most of the earth

Given a conceptual analysis of mental concepts – even concepts of consciousness – we could extend this kind of *a priori* 'reading off' to the relationship between the mental and the physical.²

Since CH implies the *a priori* deduction of all truths from the physical truths and the conceptual analyses about the other truths, it implies DE. For it will imply a deduction of all the truths about consciousness from all the physical truths, including the laws of nature, and such a deduction would be of the form that a Deductive Nomological explanation of Hempel's form requires (Hempel 1966). But DE does not imply CH, since it does not say that the deduction of all the truths about consciousness must be an *a priori* deduction from the supervenience base: i.e. a deduction which requires no more empirical information than the information about the supervenience base.

So CH is stronger than the conjunction of SP and DE. It is for this reason that Jackson's view (that physicalism implies CH) makes him more vulnerable to attack from non-physicalists. For if Jackson is right about what physicalism involves, a successful attack on CH would undermine physicalism. For example, a non-physicalist could construct this familiar metaphysical extension of the explanatory gap argument:

- (i) Physicalism implies that cosmic hermeneutics is possible;
- (ii) Cosmic hermeneutics is impossible;
- (iii) Therefore physicalism is false.

The defence of premise (ii) could come from an argument that there is no *a priori* incoherence in supposing the supervenience base of consciousness to be present in the absence of consciousness – no matter how good our conceptual analysis of consciousness. Given this, then there can be no *a priori* deduction of the sort envisaged by CH, and since physicalism implies CH, physicalism is false.

Support for (ii) could be bolstered by a zombie argument (Chalmers 1996): if zombies are conceivable, then they are possible, and if zombies are possible, then physicalism is false. However, if a non-physicalist took this route, then they would not need the detour via CH. It is a nice question whether CH or the move from conceivability to possibility is a more controversial one; but not a question we have to settle here.

Levine, in contrast to Chalmers, accepts that the zombie argument shows that zombies are conceivable, but denies that this implies their possibility. What the zombie argument shows is the presence of an explanatory gap, not a metaphysical gap:

² The origin of these ideas in David Lewis's analytical functionalism should be obvious: see Lewis 1972. For similar ideas, see Lewis 1995 and Pettit 1993.

No matter how rich the information-processing or the neurophysiological story gets, it still seems quite coherent to imagine that all that should be going on without there being anything that it's like to undergo the states in question. Yet if the physical or functional story really explained the qualitative character, it would not be so clearly imaginable that the qualia should be missing. (Levine 1997: 549)

It is clear here that Levine thinks that it is perfectly conceivable that all the physical might exist without the qualia; but he will not go as Chalmers in saying that this is a real possibility. So he remains a physicalist, despite the explanatory gap.

Some physicalists, however, disagree with Jackson *and* with Levine. All these physicalists will deny that physicalism requires CH (see Byrne 1999) but some also deny that it requires DE (see, e.g. Block & Stalnaker 1999). That is, they deny that physicalism needs to explain consciousness in terms of a link to its physical basis which must be articulated by a deductive argument. I call these philosophers 'non-reductive physicalists'. They accept SP, but they do not claim that SP must have an explanation. The necessity postulated in SP is a brute necessity, it is what Chalmers calls a 'strong metaphysical necessity'.

I think these physicalists are right to reject CH and DE. But the problem now is what makes their view a genuinely physicalist one, in any interesting sense. Physicalism can not simply be the denial of Cartesian dualism – it cannot simply be the assertion that there are no purely mental particulars or souls – since then there would be little debate about the truth of physicalism (Chalmers would be a physicalist on this definition). At the heart of physicalism, it seems to me, is a commitment to either the ontological or the explanatory priority of the physical ('physical' in the sense of 'physics' or 'physical science'). Non-reductive physicalism can certainly assert the ontological priority of the physical in the sense of SP. But, as we will now see, the doctrine known as *emergentism* can do the same. Yet ever since the advent of physicalism as a serious metaphysical view, physicalists have sought to distinguish their view from emergentism. If non-reductive physicalism cannot be distinguished from emergentism then it barely deserves the name of physicalism at all. In what follows I shall develop this line of thought.

3. Emergence

In order to see the problem for non-reductive physicalism, we first need to distinguish emergence from reduction. The term 'emergence' could be used for a number of different kinds of phenomena. The philosophically interesting use of the term is to express the view that certain properties of things are fundamentally different from others: certain properties are 'emergent' properties and others are not. To distinguish this idea from the idea that the mind developed or emerged over time, I shall call this 'synchronic' emergence. Synchronic emergence is the view that it is true of a system or entity *at a time* that some of its properties are emergent and others not, regardless of how it evolved or whatever its history (i.e. the distinction between emergents and non-emergents would remain even if the world were created in an instant, with no evolution or historical development of mind).

Synchronic emergence in this sense is centrally a feature of *properties*. This is how the term was introduced by some philosophers of the 19th century to describe

certain features of macroscopic objects.³ The rough idea is that these features of objects are genuinely ‘novel’ in the sense that they are not purely ‘consequences’ of microscopic parts, and yet they are not ‘added from outside’ in the way that is claimed by (for example) a Cartesian conception of mental properties, or a vitalist conception of biological properties. (Vitalism is the view that the explanation of biological life cannot be given in chemical and physical terms alone, but requires the postulation of vital forces.) So emergentism – the idea that some properties are emergent – is intended to steer a middle path between reductionism and forms of dualism like Cartesianism and vitalism.

The roughness of this description of emergent properties is indicated by the metaphors and scare quotes. How can we make these ideas more precise? Let us start with the idea of novelty. Certainly there are properties of macroscopic wholes which are not identical with properties of their parts, yet which are in an obvious sense ‘nothing over and above’ the properties of their parts. An object’s weight is an example. An object may weigh ten kilos, and yet in an obvious sense this weight is nothing over and above the weights of (say) its ten parts, each of which weighs a kilo. The weight of the object is a simple function of the weights of its parts; the object’s weight is only ‘novel’ from what Ernest Nagel (1963) once called ‘the additive point of view’. Properties like weight are what the 19th and early 20th century British emergentists called ‘resultants’. Emergents were contrasted with resultants, in that they were not *a priori* deducible from the properties of a thing’s parts, by ‘adding’ the properties of the parts.

Yet although this is one of the traditional ways of introducing the term ‘emergent’, this criterion will not distinguish for us the precise sense in which emergent properties are novel. For there are many properties of macroscopic things which are plausibly reducible to properties of the parts without being reducible by this simple ‘additive’ method: the temperature of some types of substance is an example. And this is true even if the temperature of something is a numerically distinct property from any of its microproperties. Distinctness is not sufficient for novelty.

We have more success if we turn to the idea of ‘adding something from the outside’. Though vague, the idea that emergent properties are properties which are not ‘added from the outside’ suggests the following, somewhat more precise thought: emergent properties of macroscopic objects are *dependent* on the properties of their parts, in such a way that there is no variation in the object’s macroproperties without variation in its parts’ microproperties. In other words, the macroproperties *supervene* on the microproperties, in one of the many senses of that term employed in recent philosophy. Indeed, Jaegwon Kim has claimed that this is all emergence really means:

According to emergentism, higher-level properties, notably consciousness and other mental properties, emerge when, and only when, an appropriate set of lower-level ‘basal conditions’ are present and this means that the occurrence of the higher properties is determined by, and dependent on, the instantiation of appropriate lower-level properties and relations. In spite of this, emergent properties were held to be ‘genuinely novel’ characteristics irreducible to the lower-level processes from which they emerge. Clearly, then, the concept of emergence combines the three components of supervenience, namely, property

³ For the historical background, and a brilliant discussion of the merits of the view, see McLaughlin 1992. For a sympathetic discussion of the views McLaughlin describes, see Crane 2001.

co-variance, dependence and non-reducibility. In fact, emergentism can be regarded as the first systematic formulation of non-reductive physicalism. (Kim 1995: 576-7)

Kim here claims that emergence is supervenience, and that emergentism is a type of physicalism. I think Kim is right to connect the ideas of dependence and co-variance with emergence, but wrong to think that emergence simply *is* supervenience. And he is also wrong to say that emergentism is a form of physicalism.

Supervenience is not sufficient for emergence because the supervenience of A on B is compatible with the reduction of A to B. Indeed, the idea of supervenience in general is compatible with almost any position in the metaphysics of the mind, so it can hardly be used to express such positions.⁴ Certainly if mental properties *are* emergent properties, then it ought to *follow* that they are supervenient – since this is what is meant by ‘not being added from the outside’. But this only means supervenience is *necessary* for emergence, not that it is sufficient. What more is needed to pick out the idea of an emergent property?

Sometimes it is said that emergent properties are those properties of a thing whose instantiation cannot be predicted from knowledge of the thing’s parts (Broad 1929: 67-68). We might be concerned here that prediction is an epistemic notion, relating to our ways of knowing things, and it is unwise to mark a distinction in kinds of properties (emergent vs. reducible) in epistemic terms. And there are other difficulties with the idea of predictability too, since whether we can actually predict something depends on having a vocabulary in which to describe it, and the existence of such a vocabulary cannot determine whether a property is emergent or not (see Crane 2001: §2).

Nonetheless, properly understood, the idea of predictability contains the key to emergence. For prediction is linked to the idea of explanation. When one has an explanation of X in terms of Y, then often there is some prospect of predicting X from the presence of Y. If we have no further explanation of Y, then we can say that Y is *explanatorily basic*.

Now, physicalists are those who think that physical science has some kind of priority in our account of the world. One way to understand this priority is merely in terms of the idea that everything has physical properties, or that everything is subject to the laws of physics. As noted above, this anodyne (though plausible) view is surely not strong enough to count as physicalism. We need to add to this the idea that the physical sciences are *explanatorily* more basic as well. Physicalism must also contain the idea that explanations of our world must come to an end with physical principles and the appeal to purely physical entities. Explanations of natural phenomena (of whatever form they take) must bottom out in terms of explanations in the physical sciences. And this is *why* they are also *metaphysically* basic: because, according to physicalism, physics provides the basis of the true metaphysics.

In a particularly insightful discussion of this matter, Terence Horgan says:

A physicalist position should surely assert, contrary to emergentism ... that any metaphysically basic facts or laws—any unexplained explainers, so to speak—are facts or laws within physics itself (Horgan 1993: 560).

This, plus the supervenience thesis, is the key to emergentism. A doctrine which holds

⁴ Here I agree with Kim’s later thoughts in his 1998: chapter 1.

that mental and other higher-level phenomena supervene on the physical, but that the supervenience in question has no explanation from within physics, is not physicalism by Horgan's lights. It is, rather, a doctrine which fully deserves the name *emergentism*. This interpretation of the situation makes it intelligible why physicalists have traditionally wanted to distinguish themselves from emergentists.

In the next section, we will see where this leaves non-reductive physicalists.

4. Physicalism and reductionism

Let me repeat again that if physicalism is to be interesting doctrine, something worth asserting and also worth denying, then it must be more than the claim that all objects have physical properties. The denial of this is the claim that not all objects have physical properties; or in other words, there are objects with no physical properties. Ignoring numbers for the time being, such objects are the traditional mental substances of Cartesian metaphysics. Such mental substances receive few supporters these days, and I will pass them over in silence here.

The claim that all objects have physical properties (the denial of Cartesian dualism) I shall call *the generality of physics*, and I shall take it to be partly constitutive of our (contemporary) idea of the physical that physics has this generality (see Crane 2001a: chapter 2). We believe that the laws of physics apply unrestrictedly across the universe; there are no regions where these laws fail or break down, and there are no kinds of entities which are immune from the effects of gravity and so on. But for the laws to have this generality, then all the objects to which they apply must have the kinds of properties which these laws concern: physical properties. Everything in spacetime has (or has parts which have) these properties: for example, mass, temperature, electric charge and so on.

The generality of physics is consistent with there being many kinds of objects: some objects, for example, are mental, some biological or organic, and some are social. These objects pose no problem for the generality of physics, so long as these objects or their parts have uncontroversially physical properties. A mental object is just an object with a mental property, a social object one with a social property, and so on. So, in particular, the generality of physics is compatible with the existence of emergent properties, defined above as: the supervenient properties of a thing not identical to any properties of its parts, and where the supervenience has no explanation in physical terms.

It is for this reason that I insist that anything worth calling the name 'physicalism' must be a reductive or reductionist view about the mental, *either* in the sense that it requires that the mental be metaphysically 'grounded in', 'realised by' or identified with physical phenomena, *or* in the sense that it requires that the mental be *explained* in physical terms.

This is the distinction which in previous work (Crane 2001a: chapter 2) I described as the distinction between 'ontological reduction' and 'explanatory reduction'. The distinction can be summarised as follows:

(OR) *Ontological reduction*: All entities (objects, properties, relations, facts etc.) belong to a subclass of the class of physical entities.

(ER) *Explanatory reduction*: All truths (particular truths, or general theoretical truths or laws) can be explained in principle in terms of broadly physical truths.

A classic example of OR in the philosophy of mind is Davidson's (1970) anomalous

monism, which identifies the class of mental events as a sub-class of physical events. Another classic example is D.M. Armstrong's (1968) type identity theory, which identifies the class of mental properties as a sub-class of the physical properties (by means of an identification of both as the states apt for bringing about a kind of behaviour).

A classic example of ER is the Armstrong/Lewis reduction of mental concepts to functional role concepts. On this view, mental concepts imply certain generalisations about the typical causes and effects of the properties or states to which those concepts refer. If this is right, then the fact that someone is in a certain mental state can be explained in terms of their being in a state which is apt for bringing about certain states of affairs, given certain other input from other mental states or from the environment (see Lewis 1972). Since the proposal is that these causal truths can be characterised in non-mental terms, we have a reduction of the mental to the non-mental.

It is clear from these examples that the two kinds of reduction are independent of one another. Armstrong and Lewis hold both. But some physicalists hold OR without holding ER: Davidson (1970) is an obvious example. And some hold ER without holding OR: a recent example is Melnyck (2004), but see also Smith (1992). Those who hold ER without holding OR typically say that they have no concern to identify mental properties with physical properties – perhaps because they are persuaded by Hilary Putnam's 'multiple realisation' argument. But nonetheless, they think that the relationship between the mental and the physical needs to be explained: it cannot be an inexplicable mystery.

5. The challenge of the explanatory gap

The upshot of the previous section is that any doctrine worthy of the name of physicalism holds either OR, or ER or both. If this is right, then the problem becomes apparent for those non-reductive physicalists mentioned in section 2 above, who respond to the challenge of the explanatory gap by saying that CH or DE is false (Byrne 1999 and Block and Stalnaker 1999). These philosophers reject Jackson's requirement of cosmic hermeneutics, and they also reject Levine's demand for a deductive-nomological explanation of consciousness. Yet unless they give some other explanatory account of the relationship between the physical and the mental, then this form of physicalism collapses into emergentism.

To put the matter another way, both non-reductive physicalism and many traditional emergentists (e.g. Broad 1929) hold the following two doctrines:

(SP) Any minimal physical duplicate of the actual world is a duplicate in every respect.⁵

(Not-OR) It is not the case that all entities (objects, properties, relations, facts etc.) belong to a subclass of the class of physical entities.

Nor-OR implies that there are some non-physical entities: for example, non-physical properties. By 'non-physical properties', I simply mean properties which do not figure in physical theories. Non-reductive physicalists hold not-OR because of their denial

⁵ I am assuming here that emergentism holds the supervenience in question to be necessary; this is clear in Broad 1929 and the argument can be found in McLaughlin 1992 and Horgan 1993.

of the identity theory. But I argued above that physicalists must either hold OR, or ER or both. So the difference between physicalism and emergentism must come down to their attitude to ER:

(ER) All truths (particular truths, or general theoretical truths or laws) can be explained in principle in terms of broadly physical truths.

The emergentist, of course, must deny ER. But I hope it is obvious now why a non-*ontologically* reductionist physicalist must accept it. For to deny ER is to hold that there is no explanatory reduction of the mental in physical terms. This amounts to accepting that the connection between the physical and the mental is a brute, inexplicable necessity. But this was supposed to be the characteristic thesis of emergentism. So if non-ontologically reductionist physicalists do not accept ER, then their physicalism collapses into emergentism.

Jackson's position can now be seen as a challenge to other forms of physicalism: either give your own solution to the explanatory gap, or become an emergentist. Recall that Jackson argued that (a) the necessity involved in SP cannot be 'brute'; and (b) that it must be explicable as conceptual necessity, in terms of an *a priori* analysis of mental concepts. Emergentism, as we have seen, denies (a) and (b). Physicalists may reject (b), along with their rejection of cosmic hermeneutics (CH). But they cannot reject ER, on pain of abandoning what is essential to physicalism. What physicalists need to do, then, is to give another kind of explanatory account of the relationship between consciousness and the physical world. The existence of statistical models of explanation or teleological forms of explanation, for example, shows that explanation need not fit the deductive-nomological form. Not all explanation is deductive-nomological. I think this is what Brian Loar is getting at when he writes that 'it is a mistake to think that, if physicalism is true, consciousness as we conceive it at first hand needs explaining in the way that liquidity as we ordinarily conceive it gets explained' (Loar 1997: 609). Note that an identity theory itself might be such an explanation, so long as asserting an identity between A and B is a way of explaining why A is B. This is, however, somewhat controversial.

In an insightful passage at the end of his book on pain, Nikola Grahek opposes Levine's claim that there is an explanatory gap:

The fact that we can conceive such states of affairs [as pain without injury] will not signal – contrary to Levine – the presence of an explanatory gap; it will not reveal the unintelligibility of the connection between pain and injury ... That we can conceive of injury without pain as well as pain without injury is rather to be explained by the simple fact that they are, to use traditional terminology, distinct existences. ... But this does not mean at all that pain and injury are arbitrarily stacked together; that their relationship is mysterious or unintelligible because we do not see at all why they should be tightly connected. It only speaks against the identification of pain with injury and merely shows that pain cannot be *a priori* analysed in terms of its causal role ... The general lesson to be learned from these consideration is that the *a priori* analysis of phenomenal concepts in functional terms is not a prerequisite for adequate or intelligible psychophysical explanations. (Grahek 2001: 150)

It seems to me that what Grahek says about the relationship between pain and injury can be applied also the case of pain and its neurophysiological basis. Non-reductive

physicalists will want to join Grahek in rejecting the need for an *a priori* analysis of phenomenal concepts in functional terms. But they cannot rest here if they want to avoid emergentism: they should also join Grahek in looking for empirical explanations of why pain and its neural basis are not ‘arbitrarily stacked together’.

So the explanatory gap does present a real challenge to physicalism; but not exactly in the way that Levine thought. Levine thinks that the explanatory gap shows physicalism to be incomplete. Without a solution to the explanatory gap, physicalism might be true, but it would be unintelligible. The argument of this paper has been, rather, that without closing the explanatory gap, there would be no way a non-ontologically reductive physicalist can remain a physicalist. Although neither deductive explanation (DE) nor cosmic hermeneutics (CH) should be required for a physicalist account of the mental, physicalists cannot simply deny the need to give *some* kind of explanatory reduction of the mental. Rather, they must give some other explanatory reduction if physicalism is not to collapse into emergentism. They need to give an account which is consistent with and explains the necessity of (SP). So far no such account has been forthcoming. But the tenability of non-reductive physicalism ultimately depends on giving such an account, and with it a solution of the problem of the explanatory gap.⁶

*Department of Philosophy
UCL
Gower Street
London WC1E 6BT
&
Institute of Philosophy
University of London
Malet Street
London WC1E 7HU*

REFERENCES

- Armstrong, D.M. (1968) *A Materialist Theory of the Mind* London: Routledge and Kegan Paul.
- Block, Ned and Stalnaker, Robert (1999) ‘Conceptual Analysis, Dualism, and the Explanatory Gap’ *Philosophical Review* 108:1-46.
- Broad, C.D. (1929) *The Mind and its Place in Nature* London: Routledge and Kegan Paul.
- Byrne, Alex (1999) ‘Cosmic Hermeneutics’ *Philosophical Perspectives* 13:347-83.
- Chalmers, David (1996) *The Conscious Mind: In Search of a Fundamental Theory*, Oxford: Oxford University Press.
- Crane, Tim (2001) ‘The Significance of Emergence’ in *Physicalism and its Discontents* C. Gillett and B. Loewer (eds.) Cambridge: Cambridge University Press, 207-224.
- Crane, Tim (2001a) *Elements of Mind* Oxford: Oxford University Press.
- Davidson, Donald (1970) ‘Mental Events’ in *Experience and Theory* L. Foster & J. Swanson (eds.), London: Duckworth, 79-101.
- Grahek, Nikola (2001), *Feeling Pain and Being in Pain* Oldenburg: Hanse Institute

⁶ An earlier version of this paper was read at the ‘Explanatory Gap’ conference in memory of Nikola Grahek, held in Belgrade in April 2005. I am grateful to Katalin Farkas, Mike Martin, Massimiliano Vignolo and Tim Williamson for discussion. I dedicate the paper to the memory of Nikola Grahek.

for Advanced Study.

Hempel, C.G. (1966) *Philosophy of Natural Science* Englewood Cliffs, NJ: Prentice Hall.

Horgan, Terence (1984) 'Supervenience and Cosmic Hermeneutics', *Southern Journal of Philosophy* 22 (Spindel Conference Supplement on Supervenience) 19-38.

Horgan, Terence (1993) 'From Supervenience to Superdupervenience: Meeting the Demands of a Material World' *Mind* 102: 555-586.

Jackson, Frank (1998) *From Metaphysics to Ethics* Oxford: Oxford University Press.

Kim, Jaegwon (1995) 'Supervenience' in *A Companion to the Philosophy of Mind* S. Guttenplan (ed.) Oxford: Blackwell, 575-587.

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

Levine, Joseph (1983) 'Materialism and qualia: the explanatory gap.' *Pacific Philosophical Quarterly*, 64: 354-361.

Levine, Joseph (1997) 'On leaving out what it's like' in *The Nature of Consciousness* eds. N. Block, O. Flanagan and G. Güzeldere, Cambridge, Mass.: MIT Press, 543-556.

Levine, Joseph (2001) *Purple Haze* Oxford and New York: Oxford University Press.

Lewis, David (1972) 'Psychophysical and Theoretical Identifications' *Australasian Journal of Philosophy*, 50: 249-58.

Lewis, David (1995) 'Reduction of Mind' *A Companion to the Philosophy of Mind* S. Guttenplan ed. Oxford: Blackwell, 412-431.

Loar, Brian (1997) 'Phenomenal States' in *The Nature of Consciousness* eds. N. Block, O. Flanagan and G. Güzeldere, 597-616.

McLaughlin, Brian (1992), 'The Rise and fall of British emergentism' in A. Beckermann *et al* (eds.) *Emergence or Reduction?* Berlin: De Gruyter, 49-93.

Melnyk, Andrew (2004), *A Physicalist Manifesto* Cambridge: Cambridge University Press.

Nagel, E (1963) 'Wholes, Sums and Organic Unities' in *Parts and Wholes* D. Lerner (ed.) New York: Free Press, 135-155.

Nagel, Thomas (1974) 'What is it Like to be a Bat?' *Philosophical Review* 83: 435-450.

Pettit, Philip (1993) 'A Definition of Physicalism' *Analysis* 53: 213-23.

Smith, Peter (1992) 'Modest Reductions and the Unity of Science', in *Reduction, Explanation and Realism* D. Charles and K. Lennon (eds.) Oxford: Oxford University Press, 19-44.

Tennant, Neil (forthcoming) 'Mind, Mathematics and the *Ignorabimusstreit*' *British Journal for the History of Philosophy*