

Ontological Novelty, Emergence, and the Mind-Body Problem

Katalin Balog
Yale University

(in (Günter Abel, ed.) *Kreativität*, Hamburg: Meiner Verlag, 2005)

There are two views, or types of views concerning fundamental ontology that are of particular interest in the current discussion of the Mind-Body problem, which I will call “Physicalism” and “Emergent Property Dualism.” According to Physicalism, the world’s fundamental ontology is physical and the best account of that ontology is provided by fundamental physics. According to contemporary physics, this ontology consists of particles, strings and fields of various types that occupy space-time (or bear spatio-temporal relations to one another) and possess a limited number of quantitative properties (mass, charge, electromagnetic potential and so on). Physics also claims that there are only a few fundamental dynamical and perhaps non-dynamical laws that govern the structure of space-time and evolution of its occupants. Physicalism asserts that everything else, whatever other entities and properties there are, is composed out of and realized by configurations of this fundamental physical ontology. Physicalism also asserts that all macroscopic or special science laws, causal relations, probabilities are ultimately derived from the laws of fundamental physics and the arrangement of fundamental physical entities.¹

According to “Emergent Property Dualism”, the fundamental ontology includes, in addition to the ontology of physics, fundamental mental or proto mental properties and may also include fundamental laws that link mental

¹ Lewis 1983, Jackson 1993, Papineau 1993, Loewer 2001, Melnyk 2003, e.g., are physicalists of this sort. The first precise formulation of physicalism comes from Lewis 1983. Subsequent discussions are variations of the same theme. Many philosophers, among them non-physicalists, accept Lewis’s definition as capturing the intuitive notion of physicalism (see, e.g., Chalmers 1996 p 41-42). The term “physicalism” is sometimes used to label another, weaker doctrine, i.e., the view that all *entities* are physical or physically realized. Davidson 1980 seems to have such a view.

properties to each other and to certain properties of physical systems.² By “mental properties” I mean properties that involve phenomenal consciousness and/or intentionality. A phenomenal property is such that, in Nagel’s famous phrase, there is something it is like to have it; for example, *feeling dizzy*. An intentional property is one which involves some kind of aboutness or reference; for example, *thinking about Vienna*.³ Emergent property dualists differ from older dualist traditions in that their ontologies don’t include mental entities or substances, only mental properties that are instantiated in certain physical systems (e.g., you and me).

There are two varieties of Emergent Property Dualism; interactive and epiphenomenal. Interactive property dualists posit *sui generis* mental laws and causal relations that are thought to be required to account for mental processes and for certain physical phenomena; for example, to account for intelligent behavior. Epiphenomenal property dualists agree with physicalists that physical phenomena can be explained entirely within physics but hold that fundamental mental properties are needed to account for the very existence of mental phenomena since these cannot be explained physically. Interactive Property Dualism holds that there are horizontal (as well as vertical) laws connecting mental to physical properties. Epiphenomenal dualists think that there are only

² I will assume that these laws are contingent; i.e., not metaphysically necessary. If laws are taken to be metaphysically necessary then it is difficult to state the difference between Physicalism and Emergent Dualism since then both would hold that configurations of physical property instantiations metaphysically necessitate mental property instantiations. The physicalist adds that the mental properties are nothing over and above the physical properties while the dualist says that they are distinct, but it is difficult to say what this comes to. Almog 2002 holds a view on which mind and body are distinct but there is a necessary connection between them but I find the view implausible and even hard to consistently articulate.

³ The natures of and relationship between phenomenal consciousness and intentionality are vast and much discussed topics. In this paper I am not assuming any specific view about how they are connected and whether it is possible for one to be instantiated without the other.

vertical laws connecting mental properties to either mental or physical properties.⁴

If I am permitted to use a theological metaphor I could say that on the physicalist view to create the world all God needed to do was to create a fundamental physical ontology and fundamental laws, and distribute elements of the ontology as an initial condition – the Big Bang – in space-time and then let the universe evolve in accordance with those laws. On the emergent property dualist view, God didn't rest until s/he also created mental properties and laws connecting them to each other and to certain (perhaps very complicated) physical properties of physical systems.⁵

The Emergence of Consciousness

⁴ Emergent Property Dualism was advocated by the “British Emergentists” including C.D. Broad 1951. Recently the view has gotten more popular among philosophers; Chalmers' 1996 book *The Conscious Mind* is a major influence in the resurgence of Dualism. Chalmers on his weblog *fragments of consciousness* on September 26, 2005

(http://fragments.consc.net/djc/2005/09/jaegwon_kim_com.html) cites the following (recent or long term) converts to Dualism: Joseph Almog, Torin Alter, George Bealer, Laurence Bonjour, Paul Boghossian, Tyler Burge, Tim Crane, John Foster, Brie Gertler, George Graham, W.D. Hart, Ted Honderich, Terry Horgan, Steven Horst, Saul Kripke, Harold Langsam, E.J. Lowe, Kirk Ludwig, Trenton Merricks, Martine Nida-Rumelin, Adam Pautz, David Pitt, Alvin Plantinga, Howard Robinson, William Robinson, Gregg Rosenberg, A.D. Smith, and Richard Swinburne, Stephen White. The majority of philosophers working on the Mind-Body problem are still (?) physicalists.

⁵ An ontological view distinct from the two I am discussing and which deserves more than this footnote is what David Chalmers calls “Type F Monism” (sometimes it is also referred to as “Russellian Monism”). According to type F Monism, the most fundamental properties are both mental and physical in that they possess a physical dispositional nature and a mental categorical nature. So the view is neither physicalist nor emergentist. As I see it the main problem with this view is that it is not easy to see what the categorical mental features of fundamental physical properties (e.g., being an electromagnetic field value) can possibly be or how the configuration of such proto mental properties can result in genuine mental properties (i.e., a sensation of red).

According to contemporary cosmology, the early universe – a few minutes after the Big Bang – was a very dense, very hot soup of elementary particles and radiation. As the universe expanded, this soup evolved in accordance with the laws producing novel arrangements and modifications of the fundamental physical ontology. Thus appeared stars, planets, oceans, single cell plants and animals, and so on. For each of these kinds there was a time at which there were no things of that kind and a later time in which there were things of that kind and, for some kinds, a period of time when it was not determinate whether or not that kind was instantiated.

Even if we don't know all the details, it is still very plausible (in this most physicalists and dualists are in agreement) that the physical ontology and the physical laws are sufficient in principle to account for stars, planets, oceans, plants, up until at least single cell animals.⁶ That is, it looks plausible that each of these kinds and their properties are constituted by and realized in the arrangement of entirely physical ontology. To create all these kinds all God had to do is to create the initial physical conditions and the laws.

At some point in time the first glimmerings of consciousness occurred. Consciousness might have appeared first in fishes, or mammals, or Homo sapiens. Or perhaps, as David Chalmers suggests, proto-conscious states are associated even with very simple physical systems. If this is so then at some point these proto-conscious states combined to constitute a genuinely conscious state.⁷ But

⁶ An alternative view is Pluralism, according to which various special sciences deal with entities and properties that are distinct from each other and from physical ones. Chemistry, biology, psychology each would quantify over distinct properties that are connected to physical properties (and to each other) by contingent law. Since, in my opinion, this is not a very plausible view I am not going to discuss it further here.

⁷ Chalmers 1996, pp. 297-99. If this is right, God had to create more than the physical ontology just to create ordinary physical things. This view, which is a version of Epiphenomenal Dualism, is not in itself the same as Type F Monism. The proto-conscious states, according to this view, are connected to physical states via contingent law; rather than being the categorical bases of the dispositional physical states, as in Type F Monism. On the Type F Monist view, the creation of physical and consciousness properties are of a piece, since proto-consciousness is supposed to provide the categorical bases for dispositional

whenever it was that consciousness first emerged,⁸ Physicalism and Dualism have very different accounts of that emergence. According to Physicalism, consciousness, like other biological phenomena, is constituted by or realized in certain processes involving arrangements of fundamental physical entities. In contrast, dualists claim that consciousness is a phenomenon that is entirely distinct from physical phenomena. According to emergent dualists, matter, fields, etc., no matter how complicated, are not metaphysically sufficient for consciousness. Its emergence requires the existence of vertical laws that link it to physical properties. At some point in the history of the universe such physical properties were first instantiated and via those laws consciousness emerged.

Most physicalists (and some epiphenomenalist dualists as well) think that the fundamental physical laws specify for every physical state how it will evolve or the chances of various possible evolutions. If the fundamental physical laws are of the latter sort – as some versions of Quantum Mechanics say – then the state of the universe at any time and the laws don't completely determine what will emerge and so there is a strong sense in which a new configuration of physical elements may be novel. (But even if the dynamical laws are deterministic, as other versions of Quantum Mechanics have it, the evolution of physical ontology has continually produced and continues to produce “new” physical phenomena.) In any case, on both the dualist and physicalist version of this view, the evolution of physical phenomena is explained by a purely physical ontology. Dualists who hold this view think that mental properties are *epiphenomenal* with respect to physical properties. On this view, the fact that I have a pounding head-ache is not causally relevant to my behavior and in particular to my subsequently taking an aspirin. This is rather implausible but it is forced on a dualist who holds the causal

physical properties. The two views agree in that the proto-mental properties and laws had been instantiated since the Big Bang; and that the emergence of bona fide consciousness requires special, complex arrangements of the proto-mental properties.

⁸ Of course, according to some Theist Dualists, consciousness and the rest of the universe were created all at once, and therefore they would reject the evolutionary framework in which I discuss these issues. This, however, will leave the basic metaphysical issues unchanged.

closure of physics and thinks that pervasive overdetermination of behavior by both mental and physical causes is implausible.⁹

Interactionist dualists, on the other hand, claim that the physical is not causally closed and that, in addition to the *vertical* mental-physical laws that bind mental and physical phenomena together, there are also *horizontal* mental-physical laws that fill in gaps or supersede physical laws and so they take it that with the emergence of consciousness there evolve novel *physical*, as well as mental arrangements that cannot be accounted for purely physically. Another way of putting the difference between the two version of Dualism is that if the first version (Non-Interactive, or Epiphenomenal Dualism) is correct then if God had created the purely physical ontology and neglected to add the mental phenomena and mental laws then the world might have evolved in a way that matched the physicalist account as far as the arrangements of physical ontology is concerned but its living beings would have been what the philosophy of consciousness literature calls “zombies”.¹⁰ On the Interactive Dualists view, on the other hand, depending on how the horizontal laws operate, by omitting the mental phenomena and mental laws, God might have created a world that evolves in a way that is even physically different from ours.

Arguments against Physicalism

The contemporary mind-body problem is the problem of determining which of these fundamental ontologies – if either – is correct. It is surely true that most people’s intuitions favor the dualist’s ontology. We have no idea how purely physical phenomena – the whirling of atoms in the void or undulations of fields, or patterns of neuronal activity can result in mental phenomena, and more

⁹ Jaegwon Kim 1988 has developed this point at length as a problem for any version of Property Dualism.

¹⁰ For those only familiar with the zombies of B-movies: philosophical “zombies” are beings that are physically identical to human beings – they move like us, apparently speak and behave intelligently – but they completely lack phenomenal experience; in another bit of philosophical terminology, introduced by Nagel 1974, “there is nothing it is like” to be one of them.

specifically in consciousness. Dualist philosophers have fashioned various arguments from these considerations and I will address one such argument, David Chalmers' Conceivability Argument shortly in more detail.

Before that I will briefly discuss another consideration that has sometimes been brought against Physicalism. Some dualists have argued for a dualist ontology by claiming that some *physical phenomena* cannot be adequately explained solely in terms of physical laws and events; in other words, they have appealed to phenomena that seem to indicate that physics cannot be causally complete. Descartes produced arguments of this type and more recently so has Carl Popper and John Eccles.

According to Descartes, language use, mathematical computation, and more generally rational behavior could not be explained physically. The reason was that these behaviors exhibit a kind of "creativity" or "novelty" that he thought merely physical – which he conceived as mechanical – causes could not produce. The point is very familiar. There are meaningful sentences that we can produce and understand that have never been produced before and mathematical problems that are solved for the first time. Descartes thought that no mere "machine" could accomplish these feats. To a large extent this line of argument has been abandoned in the face of developments in the cognitive sciences that suggest models of how purely mechanical devices – computers and connectionist machines – can produce linguistic behavior and solve mathematical problems. However, it should be noted that there are no satisfactory accounts of how meaningfulness or intentionality can be physically realized, nor are there satisfactory accounts of the general all purpose reasoning that human beings engage in. But, of course, it isn't as if positing nonphysical properties or entities in any way advanced the theoretical understanding of these phenomena.

Another kind of role for mentality in producing physical effects has been suggested in the literature on interpreting quantum mechanics. It has been suggested – e.g., by Eugene Wigner, John Wheeler and David Chalmers¹¹ – that conscious observation is required for the "collapse" of the quantum mechanical

¹¹ For an in-depth discussion see Loewer 2003.

state of a system and thus for the appearance of the usual macroscopic world. If this were so then the physical laws would be incomplete and mental laws and causes would be involved in the evolution of physical systems. It is said that John Wheeler actually attempted to “measure” the effects of observation on quantum systems.

However, the connections between consciousness and quantum mechanics that Wigner et al. suggested have lost their appeal. Purely physical phenomena are sufficient to underlie the “collapse” of wave functions and there are various accounts of quantum mechanics now that make no reference to observation at all in formulating their laws.

While arguments for interactive dualism deserve more than the few paragraphs above I think it is pretty much a consensus among philosophers that these arguments are not persuasive and so dualists in recent times have mostly appealed to a very different line of thought. This line of thought also can be found in Descartes. The heart of the argument is Descartes’ idea that since we can clearly and distinctly conceive of a body existing without mind (a zombie), and vice versa, it is possible for them to exist separately, and so bodily, i.e., physical phenomena are not metaphysically sufficient for consciousness.¹² Recently sophisticated arguments along these lines have been proposed by Kripke 1972, Nagel 1974, Robinson 1993, Chalmers 1996, Jackson 1998, White 2005, and

¹² Descartes’ original argument was meant to show the distinction of mental and physical *substances*. Contemporary dualists are mostly *property* dualists. Another difference is that many contemporary dualists think that, whereas phenomenal properties are non-physical, intentional states (or those intentional states that don’t involve phenomenal concepts) are physical. The conceivability arguments intend to show only that phenomenal states are non-physical – they are silent about intentional states. Descartes himself thought that his conceivability argument proves that both intentional states (thoughts) *and* phenomenal states are non-physical. The connection between intentionality and consciousness and philosophers’ changing view about their relationship is a very interesting topic that I can’t go into.

Nida-Rümelin 2005.¹³ I will concentrate here on Chalmers, who formulates one of the most sophisticated versions of the argument.

After considering the argument I will show that it is unsound.¹⁴ My diagnosis of the problem is that the Conceivability Arguments turn on a fact unappreciated by its proponents: the peculiarity of phenomenal concepts. Advocates of the argument fail to recognize the special role these concepts play in our cognitive architecture and they mistakenly explain the puzzles these special concepts create in terms of ontology. In other words, they posit distinct mental entities when positing distinct concepts would have sufficed. If this is right then there is good reason to think that the radical ontological novelty of consciousness is an illusion. I will argue for that view in the conclusion.

Chalmers' Conceivability Argument

Chalmers' argument has three crucial premises. The first concerns Physicalism. Physicalism requires that all true statements, including phenomenal statements like 'Dave is experiencing a yellow sensation at t', be necessitated by truths expressed in the language of physics.¹⁵ In this all physicalists agree. Chalmers argues furthermore that this necessitation must itself be *a priori* and that

¹³ There is an altogether different argument proposed by Bealer 1997 that targets functionalism about mental states. Since many contemporary physicalists are functionalists this is an important argument for a physicalist to address; however, proper treatment of this very complex argument would require a separate paper.

¹⁴ Other versions of the argument can be given a similar response. See Balog 1998 for an application of this type of answer to each of the Conceivability Arguments.

¹⁵ This roughly follows from the definition of physicalism stated at the beginning of the paper originating from Lewis 1983, with a slight complication. This definition is not strictly equivalent to physicalism. Statements that make reference to special kinds of property—to put it crudely, negative, and global properties—are not necessitated by the full physical description of the world; they are only necessitated by the conjunction of the full physical description of the world together with the statement that it *is* the full fundamental description of the world. However, this issue will not make a difference for the rest of this paper so I will ignore it.

such a priori truths must be grounded in the nature of phenomenal and physical concepts. This is called the *A Priori Entailment Thesis* and is a powerful and quite controversial thesis. There are some interesting considerations in favor of it discussed in Chalmers 1996 (and in a number of subsequent papers by Chalmers) and Jackson 1993 and 1998, not the least of which is that the thesis links metaphysical modality and conceptual necessity in an intuitive and elegant way.

The second crucial premise derives from the observation that, even for an ideal knower who possesses mental concepts, it is possible to know all the physical facts concerning the distribution of particles and fields and the fundamental laws of physics, and yet fail to know the distribution of mental/consciousness properties, e.g., fail to know that Dave is experiencing a yellow sensation at *t*. Chalmers argues that an ideal knower, in possession of all the relevant concepts and a complete knowledge of the physical would be, at least in principle, in a position to figure out the distribution of stars, planets, oceans, cups of coffee, animals, and so on – but not where and when consciousness is instantiated. Hence the conceivability of zombies. The third premise of the argument simply is the claim that human beings are sometimes conscious (i.e., that eliminativism about consciousness is false). It follows from these three premises that consciousness is not physical, which also means that zombies are possible. In its own way, Chalmers' argument, like Descartes', also proceeds from the conceivability of zombies to their possibility.

Let's state the argument more concisely.

- (1) If Physicalism is true, then for any true T, statements of the form
 $P \rightarrow T$
are conceptual truths.
(P is the complete fundamental physical description of the world including the fundamental physical laws and also a statement to the effect that it is complete).
- (2) There is some true statement S_i to the effect that phenomenal conscious experience occurs (eliminativism about phenomenal experience is false).
- (3) If S_i is a phenomenal statement, then ' $P \rightarrow S_i$ ' is not a conceptual truth.

So

- (4) Physicalism is false.

The Zombie Refutation

There are many responses to the Conceivability Argument; the one I will discuss here shows that the Conceivability Argument is self-undermining; with the addition of some very plausible premises one can derive a contradiction from it. My diagnosis is that that premise 1, the A Priori Entailment Thesis is false. I will argue that at least phenomenal statements (I want to remain non-committal with regard to non-phenomenal statements) are *not* a priori entailed by the complete physical truth about the world. This doesn't mean that physicalism is true. Establishing that requires separate argument. However, if I am right, then the conceivability of zombies is compatible with their impossibility, and so a major argument for dualism is refuted.

Suppose that Chalmers' argument is sound. It would be true then that physical facts do not necessitate phenomenal facts. And it would follow that there is a possible world that is exactly like our world physically, but in which no phenomenal, or other, nonphysical, facts obtain. Of course, I make this

assumption only for the sake of a reductio. If Physicalism is true, as I think it is, then such a world is impossible, or rather, merely *conceptually* possible. But my strategy is to show that the very assumption that there is such a world undermines the argument that leads to positing its existence in the first place.

In the world we are imagining there exists a zombie-Chalmers, physically just like Chalmers, but not the subject of any phenomenal states. This conceptually possible zombie will formulate an argument that parallels the Conceivability Argument word by word only to reach the conclusion that Physicalism is false in the zombie-world. What are we to make of this argument? First of all, plausibly, zombie-Chalmers has intentional states. When he talks, his words are not mere meaningless sounds. Moreover, I argue that it is plausible to assume that zombie-Chalmers' intentional states are identical with Chalmers' intentional states except for intentional states that, in Chalmers, involve phenomenal concepts. Those of zombie-Chalmers' intentional states that, in Chalmers, involve phenomenal concepts refer to states of affairs present in zombie-Chalmers' world (presumably his own brain states).¹⁶ On this view, zombie-Chalmers' argument will be just as meaningful as Chalmers', though not quite *identical* to it. Although the argument is *word by word* identical to Chalmers' argument, some of the words (those that express phenomenal concepts in Chalmers' language) have different meanings in Chalmers' and zombie-Chalmers' mouths. I mark these words with a '+'. 'Pain+', for example, stands for a term of zombie-Chalmers that corresponds to Chalmers' term 'pain'.

Zombie-Chalmers' argument will go like this:

(1*) If Physicalism is true, then for any true T, statements of the form

¹⁶ I argue for all this in much more detail in Balog 1999.

P->T

are conceptual truths.

(P is the complete fundamental physical description of the world including the fundamental physical laws and also a statement to the effect that it is complete).

- (2*) There is some true statement S_i^+ to the effect that phenomenal⁺ conscious experience occurs (eliminativism about phenomenal⁺ experience is false).
- (3*) If S_i is a phenomenal⁺ statement, then 'P-> S_i^+ ' is not a conceptual truth.

So

- (4*) Physicalism is false.

We know, however, that the dualist conclusion of zombie-Chalmers' argument is, by hypothesis, false in the zombie-world. Consequently, we know that zombie-Chalmers' argument cannot be sound. Since, given that it is meaningful, it is clearly valid, one of its premises has to be false. However, because arguably each premise of zombie-Chalmers' argument is true if the corresponding premise in Chalmers' argument is true, one of the premises of Chalmers' argument must be false as well.¹⁷

The Zombie Refutation also helps diagnose where things went wrong. Since the other premises are extremely plausible, the fault must lie with premise 1, the *A Priori Entailment Thesis*. This shows that the conceivability of zombies does not have to arise from a feature specific to phenomenal consciousness, but that, more plausibly, it arises from a certain peculiarity of our phenomenal *concepts*. Zombies are conceivable because our phenomenal concepts refer *directly*, (and not via a physical, functional, or behavioral mode of presentation). The physicalist who – like myself – accepts the conceivability of zombies – has to argue that this peculiarity, that is, referring to a property *directly*, is in principle

¹⁷ See Balog 1999.

perfectly compatible with the hypothesis that they pick out a *physical property*.¹⁸ If that is so, i.e., if phenomenal concepts refer directly to physical properties, then there is a perfect explanation of why the *A Priori Entailment Thesis* is inapplicable to statements involving these concepts. It is therefore imperative for physicalists to work out a – physicalist – theory of phenomenal concepts. Such a theory needs to provide a detailed and convincing account of how a purely physical concept can refer directly to a purely physical property.

The constitutional account of phenomenal concepts

There is an account that attempts to do exactly that. I can only sketch the main idea here.¹⁹ I start by stating some criteria that a good physicalist theory of phenomenal concepts must satisfy.

Here are the minimal desiderata for a physicalist theory of phenomenal concepts:

- a) explain away the conceivability of zombies; that is, explain why the conceivability of zombies, far from being incompatible with Physicalism, is *to be expected* from the physicalist standpoint,
- b) account for apparent incorrigibility – or rather infallibility – of certain judgments involving certain phenomenal concepts – e.g., my judging ‘phenomenal red is occurring right now’,
- c) account for the way in which we seem to know our experiences – i.e., not by inference but by immediate acquaintance,
- d) account for the fact that our awareness of our own occurrent experience seems to give a substantial insight into its nature.

¹⁸ Physicalists can be non-committal about whether phenomenal properties are complex physical properties or higher level, e.g., functional properties, realized by complex physical properties. In the rest of the paper when I talk about phenomenal concepts directly referring to “physical properties” I want the reader to understand “physical or functional properties”.

¹⁹ Balog (forthcoming) elaborates the theory in much more detail. Similar ideas are proposed in Papineau 2002 and Block (forthcoming).

These desiderata suggest that a successful account of phenomenal concepts will have to posit a very intimate connection between conscious states and the concepts we form of them. Loar²⁰ suggested the idea that phenomenal concepts are direct recognitional demonstrative concepts. I think he means by this that when a person is having a particular experience she can deploy a mental demonstrative to immediately demonstrate that experience and that in some way the mode of presentation associated with the demonstrative involves the experience. Loar's suggestion is that the experience itself serves as a mode of presentation which, presumably, guides the demonstrative to the demonstrated experience. Loar doesn't explain how this works or what exactly "mode of presentation" comes to here. My account here will try to fill in the details.

The idea is that (certain) phenomenal concepts are partly constituted by the phenomenal experiences they refer to, and it is, at least in part, *in virtue of* being so constituted that they have the reference they have. On this view, a current phenomenal experience is literally *part* of the token concept currently applied to it, and the experience partly determines what the concept refers to. To be more precise, this is not a theory of phenomenal concepts in general. It is a theory of a subclass of phenomenal concepts, or certain characteristic applications of phenomenal concepts, if you will. What I have in mind is that there is a difference between phenomenal concepts that I apply to my own experience or imagination as it occurs and ones that I apply to your experiences, or to my past or future experiences. The first kind, like my concept 'this buzzing sound' as I listen to it, or my concept 'that shade of red' as I try to imagine what the sky looked like in the morning, I will call "direct phenomenal concept". The second kind, like the concept 'pain' as I apply it to your tooth-ache, I'll call "indirect phenomenal concept". Indirect phenomenal concepts are derivative on direct phenomenal concepts.

The account is then that *direct* phenomenal concepts are partly constituted by the experiences they refer to. If this is so then of course phenomenal concepts are direct also in the sense that they don't have any functional, physical, etc.

²⁰ Loar 1997.

analyses. This explains why zombies are conceivable. If there is no analytical connection between physical and phenomenal concepts, not even via the mediation of, e.g., functional concepts then there won't be any conceptual contradiction in the zombie scenario. In fact, the conceivability of zombies is not an embarrassing anomaly for Physicalism on this view, but rather one of the expected consequences of a physicalist theory of phenomenal concepts.

The intimate connection that this theory posits between a phenomenal concept and the experience it refers to also explains the incorrigibility of certain phenomenal judgments. A token of a direct phenomenal concept will contain a token of the experience it refers to and so certain judgments involving the concept will be infallible. For example, a phenomenal concept may refer to a particular type of visual experience, say the experience typically caused by seeing a red object in ordinary light, etc. – call this type of experience reddish – by being constituted in part by a particular token of that type of experience. If I form the judgment

‘This experience is reddish’

where ‘this experience’ refers to and is partly constituted by an experience of mine that is also partly constitutive of my concept 'reddish' then the judgment cannot fail to be true.

The constitutional account helps to explain some other puzzling features of conscious states: that our knowledge of our own conscious states (unlike of the conscious states of others) is via direct acquaintance with them, and that this acquaintance seems to provide us a special insight into the nature of these states. Take the case of perception. We know about the presence of a red object by perceiving its redness. Our perception of redness mediates our knowledge of the presence of the red object, by providing evidence for judging that a red object is present. The perception of redness and redness itself are distinct. Judging that an object is red is an indirect judgment based on the data provided by perception.

Now take the case of the judgments that we form about our own phenomenal states, say the judgment that I am undergoing a reddish experience. It seems that I am immediately acquainted with this experience and my judgment is

not based on anything except the experience itself. This is well accounted for by the constitutional account of phenomenal concepts. If phenomenal concepts are partly constituted by phenomenal states, our knowledge of the presence of these states (at least in our “inner” way of thinking of them) is not mediated by something distinct from these states. It is in some sense direct and not mediated by any independent evidence – rather the judgment contains its own evidence. Also, the insight such judgments afford us into the nature of conscious states is insight into what it is like to have such states. This insight is inevitable on the constitutional account. You couldn’t form the appropriate judgments unless you had the experience.

Phenomenal consciousness: emergent or physical?

The constitutional account of phenomenal concepts seems quite successful in explaining the puzzles surrounding phenomenal consciousness in a way that is fully compatible with Physicalism. There is nothing incoherent about the notion of phenomenal *concepts* being physical properties of the brain partly constituted by phenomenal *properties* that are themselves physical properties of the brain. Furthermore, this account shows why, on the assumption of Physicalism, we shouldn’t expect to have a revelatory metaphysics of consciousness. Zombies are conceivable, on this account, not because consciousness is non-physical, but because phenomenal concepts refer to them directly. For the same reason, phenomenal concepts don’t have revealing analyses, and scientific accounts of phenomenal properties will always seem puzzling. So far, so good.

But the account is also compatible with dualism. Doesn’t this show that the account is actually no help for Physicalism? Not in my view, for two reasons. First, there are arguments *for* Physicalism that carry a lot of weight.²¹ This provides a dialectic in which the constitutional account helps dispelling apparent problems with a view that has independent plausibility. Second, dualism is quite implausible and metaphysically awkward in a number of related ways.

²¹ Papineau 1995; Loewer 1995.

First, it is in tension with the nomological/causal completeness of physics. As we have seen, the nomological/causal completeness of physics is the doctrine that all physical change can be accounted for causally and nomologically – to the extent it can be accounted for at all – without going outside the physical realm. According to this principle, the falling of a tree (which consists of motions of particles that compose the tree) at t can be accounted for entirely – to the extent it can be accounted for at time t_0 – in terms of the state of the universe (or a sufficiently large part of that state) at t_0 and the fundamental laws of physics. Adding other descriptions of events at t_0 , including mental descriptions – whether or not these are necessitated by the physical descriptions – add nothing to specifying whether the tree will fall or what the probability is of the tree falling. There is good – but not conclusive – reason to think that the causal/nomological completeness of physics is true. The fundamental physical laws really do seem to cover all physical phenomena including human neurological phenomena and bodily behavior. At least there are no cases where it is plausible that anything physical violates the fundamental physical laws.²² Even some dualists (e.g., Chalmers) seem to accept it.

But if it is true then it is difficult to see what the causal role of mental properties is in the dualist ontology. There seem to be just two, equally implausible possibilities. One is that there are no horizontal laws connecting a mental property instantiation to a subsequent physical (or mental) property instantiation. This is epiphenomenalism. The other possibility is that there are horizontal laws but they are always compatible with the physical laws that connect physical states with subsequent physical property instantiations. This would involve massive *causal overdetermination*. Jaegwon Kim has argued forcefully against either of these possibilities.²³ The first is problematic since it is

²² For a comprehensive history of the interaction between the development of the sciences and views on the causal/nomological completeness principle see Papineau 2001.

²³ Although, somewhat ironically, in his recent book *Physicalism or Something Near Enough* Kim seems to end up with a kind of epiphenomenalist dualist view since he claims that phenomenal properties are not ontologically reducible to

difficult to see how we can know anything about the instantiation of properties that have no effects²⁴, but more worrisome, it just seems crazy to think that all our mental states have no physical effects. Massive overdetermination also seems enormously implausible. Why would God have made all the horizontal laws connecting the mental with the physical if in fact they make no difference to what happens (and would make no difference in any physically possible world)?

A second, and related problem is that, on both varieties of dualism, it is *metaphysically*, though probably not nomologically possible for phenomenal properties to be completely separated from physical action. Pain, e.g., might systematically result in “inappropriate” action, say, instead of avoidance it might result in behavior that seeks out pain; or, alternately, there might be no systematic connection between phenomenal experiences and action at all. Random and inappropriate phenomenal experiences might be paired up with unrelated actions and lives. If the mental and the physical are radically distinct then it must be possible for them to interact in all manners of ways, including unusual, “senseless”, “weird” ways. It is not clear that these “weird” scenarios are even coherent. Furthermore, the possible weirdness doesn’t even stop at phenomenal experiences getting disconnected from action. According to dualism, there are metaphysically possible worlds, e.g., in which phenomenal properties are connected by law *not* with biological organisms but inanimate objects of various sizes. So on the dualist view my car might be conscious or Mount Everest might.

Third, Emergent Property Dualism has to account for why psychophysical correlations occur even though phenomenal properties are not metaphysically determined by the physical; to explain those correlations nomological relations are posited. But the kind of laws, both vertical and

physical properties and that physics is causally complete. Given that he has so strongly opposed causal overdetermination it appears that he considers phenomenal properties epiphenomenal.

²⁴ One might argue that a person is *acquainted* with her phenomenal states and that this relation is not a causal one. But this seems to just put a label on the mystery.

horizontal, that Emergent Property Dualism posits are found no where else in nature. The relationship between configurations of physical entities and properties and, for example, geological and biological phenomena are not like vertical laws at all. Certain configurations of physical entities and properties literally *constitute* or *realize* geological and biological phenomena. There is no need for an “extra” law to connect the two. This doesn’t mean that biological properties are *identical* to physical properties. The reason is that certain biological properties are *multiply* realizable, not only by various physical configurations, but by configurations of ontologies that are different from the actual physical ontology and that satisfy laws that are different from the actual laws of physics. Underlying this is the widely accepted view that biological properties are individuated in terms of their causal/nomological relations. A possible entity that looks, walks, etc. like a duck may not be a duck (if it lacks the DNA of a duck) but it is plausibly *alive* in virtue of instantiating the causal/nomological profile that individuates living creatures.

Dualism, on the other hand, posits genuine laws connecting physical and phenomenal property instantiations. There are two ways the dualist can think of these laws. Depending on whether a dualist thinks of conscious properties as primitive, or as being realized by complexes of proto-conscious properties, different, but equally damaging difficulties arise. If the dualist holds that conscious properties are primitive then she will have to posit laws that connect enormously complex physical properties with simple phenomenal properties, like the sensation of a particular shade of red. She will have to posit not just a few, but a whole multitude of such fundamental laws, corresponding to each primitive conscious property. These laws are different from any laws of nature we know from science. This doesn’t show that the theory is false, only that it is implausible. She might think, on the other hand, that conscious properties are not primitive but are realized by complexes of proto-conscious properties. In this case the laws would connect complex properties with complex properties; so no implausibility on this count. However, in this case the dualist owes us an account of *how* exactly proto-conscious properties combine to produce the familiar conscious properties. The nature of these complex properties is entirely obscure and it

doesn't seem likely that there is a realization theory in the offing to clear up the mystery. The situation, *mutatis mutandis*, is very similar concerning the horizontal, diachronical laws interactionist dualists posit to connect mental and physical property instantiations at different times.

Fourth, although modern day dualists focus their attention solely on the possibility of worlds exactly like ours physically, but lacking any phenomenal properties instantiated, and not on the converse, that is, the possibility of worlds exactly like ours phenomenally, but lacking in any physical properties instantiated, it appears that a dualist would have to condone the existence of purely phenomenal worlds.²⁵ It is barely intelligible what a world like that would be like.

A final note. The constitutional account, as I said, is compatible with dualism. As a matter of fact, there is a dualist version of it proposed by David Chalmers²⁶. However, there are serious problems that a dualist version of the constitutional account faces. Chalmers summarizes the view like this: "In particular, I will take it that... the content of a phenomenal concept and a corresponding phenomenal belief, is partly *constituted* by an underlying phenomenal quality, in that the content will mirror the quality (picking out instances of the quality in all epistemic possibilities." (p. 14) Several questions arise. What is the constitution relation in the dualist framework and how does that relation succeed in determining reference? Chalmers treats constitution as in no need of explanation except in so far as whatever it is a phenomenal concept succeeds in picking out in every (conceptually) possible world is the experience type the constituent experience is a token of. But any token experience falls under several experience types: e.g., experience of red, experience of dark red, of red 218, etc. What determines which experience type a given phenomenal concept refers to? It is also difficult to understand how something non-physical can be partly constitutive of something else – the phenomenal concept – that,

²⁵Descartes actually did in the *Meditations* (cf. Cottingham, et. al 1984).

²⁶ Chalmers 2003.

presumably, will be involved in causal transactions via the judgments that contain them.

So, while the constitutional account is compatible with dualism, dualism is quite unattractive on its own, and the chief reasons people end up embracing dualism in the first place are the very ones the constitutional account helps disarming. All this points to the conclusion that the radical ontological novelty of conscious phenomena is an *illusion*. What is novel in the evolution of minds is our first person phenomenal *concepts*, but the novelty here is not ontological. There are no novel properties and laws involved in our mental lives – the familiar physical entities, properties, and laws will do. With the constitutional account we have an explanation of why, even though conscious phenomena are physical, we have an almost irresistible tendency to treat them as ontologically novel relative to the physical. But with the constitutional account we also have a reason to resist this tendency.

Reference:

- Almog, Joseph, 2002. *What Am I? Descartes and the Mind-Body Problem*, Oxford University Press.
- Balog, Katalin. 1998. *Conceivability Arguments*. Ph.D. diss., Rutgers University.
- Balog, Katalin (forthcoming) “Mental Quotation: a Cure for the Mind-Body Problem”.
- Balog, Katalin, 1999. “Conceivability, Possibility, and the Mind-Body Problem”, *The Philosophical Review*, Vol 108, No 4.
- Bealer, George, 1997. “Self-Consciousness”, *The Philosophical Review*, Vol. 106, Issue 1.
- Block, N. (forthcoming). Max Black's objection to mind-body identity. *Oxford Review of Metaphysics* 3.
- Broad, C.D., 1951. *The Mind and its Place in Nature*, Routledge & Kegan Paul, New York.
- Chalmers, David. 1996. *The Conscious Mind*. New York: Oxford University Press.
- Chalmers, David, 2002. “Consciousness and its Place in Nature.”, in (D. Chalmers, ed) *Consciousness and its Place in Nature*, Oxford University Press.
- Chalmers, David, 2003. “The Content and Epistemology of Phenomenal Belief”, in Q. Smith and A. Jokic (eds), *Consciousness: New Philosophical Perspectives*, Oxford.

- Cottingham, John, Robert Stoothoff, and Dugald Murdoch, eds. 1984. *The Philosophical Writings of Descartes*. Cambridge: Cambridge University Press.
- Davidson, Donald, 1980. "Mental Events", in Donald Davidson, *Actions and Events*, Oxford: Clarendon Press.
- Jackson, Frank. 1993. Armchair Metaphysics. In *Philosophy in Mind*, ed. M. Michael and John O'Leary-Hawthorne. Dordrecht: Kluwer.
- Jackson, Frank. 1998. *From Metaphysics to Ethics*. New York: Oxford University Press.
- Kim, Jaegwon, 1988. *Mind in a Physical World*. Cambridge: MIT Press/Bradford Books.
- Kim, Jaegwon, 2005. *Physicalism, Or Something Near Enough*, Princeton University Press.
- Kripke, Saul. 1972. *Naming and Necessity*. Cambridge: Harvard University Press.
- Lewis, David. 1983. New Work for a Theory of Universals. *Australasian Journal of Philosophy* 61:343-77.
- Loar, Brian. 1997. Phenomenal States. In *The Nature of Consciousness*, ed. Ned Block, Owen Flanagan, and Güven Güzeldere, 597-617. Cambridge: MIT Press. (Revised version of Loar 1990).
- Loewer, Barry, 1995. "An Argument for Strong Supervenience", in (E. Savellos & U. Yalcin, eds.) *Supervenience: New Essays*. Cambridge University Press.
- Loewer, Barry, 2001. "From Physics to Physicalism", in (K. Gillett, B. Loewer, eds.) *Physicalism and its Discontents*, Cambridge University Press.
- Loewer, Barry, 2003. "Consciousness and Quantum Mechanics", in *Consciousness: New Philosophical Perspectives* (Q. Smith and A. Jokic, eds), Oxford: Oxford University Press.
- Melnyk, Andrew. 2003 *A Physicalist Manifesto: Thoroughly Modern Materialism*, Cambridge UP.
- Nagel, Thomas. 1974. What Is It Like to Be a Bat? *Philosophical Review* 83:435-50.
- Nida-Rümelin, Martine. 2005. "Grasping Phenomenal Properties", forthcoming in Torin Alter and Sven Walter (eds.) *Phenomenal Knowledge and Phenomenal Concepts*, Oxford, OUP.
- Papineau, David. *Philosophical Naturalism*, Oxford: Blackwell Publishers, 1993.
- Papineau, David, 1995. "Arguments for Supervenience and Physical Realization." In (E. Savellos & U. Yalcin, eds.) *Supervenience: New Essays*. Cambridge University Press.
- Papineau, David, 2001. "The Rise of Physicalism", in (K. Gillett, B. Loewer, eds.) *Physicalism and its Discontents*, Cambridge University Press.
- Papineau, David, 2002. *Thinking about consciousness*, Oxford UP.
- Robinson, Howard. 1993. The Anti-materialist Strategy and the Knowledge Argument. In *Objections to Physicalism*, ed. Howard Robinson. Oxford: Oxford University Press.

White, Stephen. 2005. "The Argument for the Semantic Premise", forthcoming in Torin Alter and Sven Walter (eds.) *Phenomenal Knowledge and Phenomenal Concepts*, Oxford, OUP.