in *Sage Encyclopaedia of the Mind* (ed.) H. Pashler (Sage Reference 2013)

# Emergence

## Introduction

The term 'emergence' is used in a variety of (often incompatible) ways in the philosophic and scientific literature. However, emergentism is always a thesis about *properties*, and most versions share the following components: emergent properties are (i) dependent upon, (ii) determined by and (iii) not deducible from, basic physics. A highly diverse range of properties have been held, at various times, and by various philosophers and scientists, to be emergent. Examples include mental properties, most often consciousness; chemical bonding; and ordered patters such as tornados in chaotic systems such as the weather. To say of such properties that they are emergent is to say something about the way they are related to the particles, properties and laws of physics. This entry first clarifies the related notions of emergence and non-deducibility, then distinguishes three distinct forms of emergence with respect to the basic physical level. Let us begin with clarification of the key notions.

Emergent properties are:

(i) Instantiated only by complex physical particulars;

(ii) Determined by structural physical properties of their bearers;

(iii) *Non-deducible* from physics.

Emergentists about the mind, for instance, think that at least some mental properties (e.g. beliefs, desires, sensory experiences) satisfy (i)-(iii). Focus on this kind of emergence. Component (i) contrasts emergentism with Cartesian substance dualism – emergentists deny any non-physical 'mental substance' in which our mental lives take place. Mental properties are possessed only by things that are fully composed of physical parts (e.g. brains).

Component (ii) holds that instantiation of certain structural physical properties is sufficient for instantiation of emergent properties, and entails that you and I cannot differ mentally unless we also differ physically. Components (i) and (ii) are common to most extant theories of mind, emergentist or not, and are often jointly referred to as the *supervenience* of mental properties on the physical. What distinguishes emergentism from these other positions is component (iii). Non-deducibility makes emergent properties 'surprising', 'novel', and 'unexpected' with respect to the physical properties they emerge from – no scientist could *deduce* your mental state merely from physical knowledge about your brain, body, environment, and so on. The broad consensus on defining emergence in terms of non-deducibility hides significant differences. Just as there are several ways in which you can fail to have blue eyes, so there are several ways in which a property can fail to be deducible from physics.

## Deducibility

Arguably the only way to deduce a property from physical properties and laws is to complete a 'functional reduction' of the property to be deduced. Understand a *functional property* to be a property defined by its causes and effects. Being a mousetrap is a functional property, shared by many physically diverse machines – mousetraps differ in physical constitution, and they trap mice in a range of different ways. All that those machines have *in common* is that when you input a live mouse, they output a dead (or in some cases live, but captive) mouse. The property of being a mousetrap is what all mousetraps have in common, and so is defined by what mousetraps *do* – by the way they *function*. Suppose we are trying to deduce a property Q from physical properties, entities and laws. According to Jaegwon Kim's theory of functional reduction, we must:

(a) 'Work Q into shape' for deduction by conceiving it as a functional property, i.e. 'Q = the property of being such as to play causal role R'.

(b) Find a physical property P that plays causal role R, and explain how it does so, in terms of basic physical laws and properties.

If we can complete these two steps, we are in a position to deduce which things have property Q. An example will help. Water is highly *cohesive*, compared to chemically similar compounds such as hydrogen sulphide ($H_2S$). For instance, $H_2O$ boils at 100°C, but $H_2S$ boils at -60.2°C, so $H_2S$ is gaseous at room temperature, whereas $H_2O$ is liquid. Let property Q = the cohesiveness of water, and suppose we want to deduce Q from basic physical laws and properties alone. First, we must understand cohesiveness in functional terms; let's say cohesiveness = the property of having mutually attractive parts. Now if we can use physics to predict whether and to what extent *x*'s parts attract each other, we will thereby have deduced *x*'s cohesiveness. We know that water has $H_2O$ molecules as parts. To deduce the cohesiveness of water from physics, we need to show that $H_2O$ molecules attract each other to a degree that explains (for instance) why water is a liquid at room temperature. It turns out that we can do exactly this. Due to their shape, $H_2O$ molecules are electrostatic dipoles. Coulomb's law tells us that $H_2O$ molecules will exert an attractive force on each other; this is known as 'hydrogen bonding'. $H_2S$ molecules do not form hydrogen bonds, and so hydrogen sulphide is less cohesive than water. The deduction outlined above is a *reduction* because it shows us that the cohesiveness of water is *nothing over and above* the properties and relations of its molecules.

There are several ways for a property to be non-deducible, corresponding to the ways in which functional reductions can fail. Deducibility fails for a property P if:

1. P is not exhausted by its functional role, *or*
2. P has a causal role that is not occupied by a physical property, *or*
3. Neither (1) nor (2) but *we* cannot complete the deduction of P from physics due to our own conceptual and/or computational limitations.

There may be other forms of non-deducibility, but conditions (1), (2) and (3) are at least independently *sufficient* for non-deducibility. Condition (1) will be met by properties with fundamentally qualitative natures (whether or not they also have functional roles). Such properties cannot be deduced by functional reduction because step (a) of the reduction cannot be completed. Condition (2) will be met by any property which has novel, irreducible causal powers with respect to the physical. Such properties cannot be deduced because step (b) cannot be completed. Conditions (1) and (2) are each sufficient for non-deducibility *in principle*; condition (3), on the other hand, will be met by any property which is in principle deducible (a property which meets neither of the two other conditions), but is *in practice* non-deducible due to certain facts about ourselves, which facts typically vary from case to case.

Given the preliminary characterisation of emergence given in the introduction to this article, there will be as many kinds of emergence as there are ways for deducibility to fail. We can usefully divide them, however, into two categories: 'ontological', and 'epistemological'. Ontology is the study of what there is; correspondingly, ontological emergentists think that emergent properties are non-deducible because they are 'something over and above' the physical. A supervenient property meeting either of conditions (1) or (2) will be ontologically

emergent. Properties that meet (1) need not meet (2), and vice-versa. Furthermore, (1) and (2) are compatible, in that there might be a property that met both conditions, so counting as doubly non-deducible. For such a property, neither step (a) nor (b) of a functional reduction can be completed. In this entry, the term 'weak ontological emergence' is used for properties which supervene on the physical and meet (1) but *not* (2); and 'strong ontological emergence' for supervenient properties that meet (2), regardless of whether they meet (1). Epistemology is the study of knowledge; correspondingly, epistemological emergentists think that emergent properties are non-deducible because of limitations in our cognitive and/or conceptual abilities. A supervenient property meeting condition (3) will be epistemologically emergent. Certain properties which are nothing over and above the physical might be non-deducible for a range of reasons. For instance, our way of thinking about a property might make it very difficult for us to functionally conceive it; or the complexity of the physics involved might make it in practice impossible for us to deduce it. We may now appeal to the three different ways in which deducibility can fail in order to more precisely characterise three distinct kinds of emergence.

## Three kinds of emergence

We will first give a broad characterisation of the two kinds of ontological emergence, and through this characterise epistemological emergence. Think back to the case of the cohesiveness of water. Cohesiveness is (a) exhausted by its causal role, (b) *physically realized,* in that physical properties play the causal role that defines it. It is these facts that enable us to deduce the cohesiveness of water from physics. The deducibility of cohesiveness makes it transparent how it is related to the physical properties of $H_2O$, and the physcial laws that govern its behaviour. Once those properties and laws are fixed, we get cohesiveness *for free,* as it is fully realized by hydrogen bonding. Physical realization of a property P thus

explains *why* P supervenes on the physical. Things are different with ontologically emergent properties. Consider a property P1 which meets condition (1) for non-deducibility. Let P1 be a qualitative property which is defined by its *qualitative* nature. P1 is not physically realized, since it is not defined by a functional role. The supervenience of P1 on the physical therefore requires 'bridge laws' connecting it to its physical base. Now consider a property P2 which meets condition (2) for non-deducibility. P2 may well be functionally defined, but it has no physical realizer. We also need bridge laws to ground the supervenience of P2. Ontological emergentists (both weak and strong) thus hold, in addition to (i)-(iii) above, that emergent properties are:

(iv) Not physically realized, and connected to the physical by irreducible 'bridge laws' that are not laws of physics.

Suppose mental properties to be ontologically emergent, and that God has created all the basic physical entities, and made the laws of physics true. His work is not yet complete – He must still make the bridge laws true, thereby making it so that some physical things have mental properties. Because there are bridge laws, true in the actual world, relating physical and mental properties, it is not possible *in the actual world* for two physically indistinguishable individuals to differ mentally. However, there are *possible* worlds that are complete physical duplicates of this one, but at which the bridge laws do *not* hold, and so which have no mental properties. This makes any kind of ontological emergence inconsistent with *physicalism*. Physicalists think that everything that exists is nothing over and above the physical, so they must hold that any possible world that is physically indistinguishable from our world is indistinguishable in all other respects as well. If any properties at our world are ontologically emergent, then physicalism is false.

Properties contribute causal powers to the things that possess them – for instance, the momentum of a brick, but not its colour, contributes to the brick the power to break windows. Components (i)-(iv) of emergence are common to both weak and strong emergence, and so do not distinguish emergent properties which do, from those which do not, contribute *novel* causal powers to their bearers. Let us add a fifth component to the existing four. Emergent properties:

(v) contribute powers over and above those contributed by the physical properties they emerge from.

This contribution of additional powers is typically called 'downwards causation', and violates the *causal closure* of the physical domain (of which more presently). We may now more precisely characterise our three kinds of emergence:

Strong ontological emergence = all of (i)-(v)

Weak ontological emergence = (i)-(iv); not (v)

Epistemological emergence = (i)-(iii); neither (iv) nor (v)

Both strongly and weakly ontologically emergent properties violate physicalism, and require bridge laws to connect them to the physical properties they supervene upon. Strongly emergent properties do, while weakly emergent properties do not, contribute causal powers over and above those of the physical properties they supervene on. Epistemologically emergent properties, by contrast, are nothing over and above the physical, do not require bridge laws, and are non-deducible solely for practical reasons. The remainder of this entry

gives examples of properties which have at one time or another been taken to be emergent in each of the senses characterised above.

1. Strong ontological emergence

Suppose mental properties are strongly emergent. If we try to predict how brains behave merely by summing the causal contributions of individual neurons, then we will 'miss out' the extra powers contributed by the mental properties, and get the predictions wrong. By contrast, if we can explain the behaviour of the brain just by reference to physical laws and properties, then mental properties do not exert a downwards causal influence. A common position in the philosophic literature is that science has already progressed far enough to show that there is no downwards causation from anything non-physical to anything physical. Physics forms a closed, complete system, such that everything that happens within the physical domain can be fully explained without reference to anything outside it. If physics is causally closed, no properties are strongly emergent. Even if there is now ample evidence to support closure, it was not always so. The physics of C. D. Broad's day, for instance, could not explain all the functional characteristics of complex chemicals, hence Broad's theory that those characteristics were strongly emergent. It seemed to Broad as though chemical compounds had functional properties that were not physically realized, but emerged according to irreducible bridge-laws, and which made causal contributions in addition to those of their physical components.

By way of illustration, imagine that the cohesiveness of water is strongly emergent. An easy way to imagine this is to suppose that $H_2O$ molecules are not dipolar, so that there is no hydrogen bonding in water, but that water is nonetheless highly cohesive – i.e. is liquid at room temperature, has a higher boiling point than $H_2S$, etc. The attractive forces between

molecules will not be due to physical force-generating properties (e.g. Coulomb forces due to charge) but rather due to the emergent cohesiveness of water. Suppose that we have physical theories that fully explain and predict the behaviour of hydrogen and oxygen atoms in isolation. Those theories will predict that water is a *gas* at 20°C, just as our current theories tell us that water *would* be a gas at 20°C if its molecules were not dipolar. Note that we can still complete step (a) of our deduction, for even if cohesiveness were strongly emergent, we could still conceive it functionally. What we could not do, if cohesiveness were strongly emergent, is find a physical realizer for it. The success of modern physics in reductively explaining the characteristics of complex chemicals is no small part of the reason why Broad's emergentism is no longer taken seriously.

2. Weak ontological emergence

Arguably the most plausible candidates for weak ontological emergence are *phenomenal properties.* Consider Frank Jackson's famous example of Mary the colour scientist, who learns all the physical facts about colour perception while locked inside a black and white room, without *seeing* anything coloured. Suppose Mary is very clever, and also knows all the facts that *are* deducible from physics. Intuitively, when she first sees a red tomato, she learns something – "what it is like" (to borrow Thomas Nagel's phrase) to see red. Such examples convince most philosophers that phenomenal properties are not deducible from physics. It is also widely agreed that such properties are non-deducible because is very difficult (perhaps impossible) to *conceive* of phenomenal properties such as the visual redness of a red tomato, or the painfulness of a pain, in purely functional terms. Understanding pain as a state caused by tissue damage, which causes you to say 'ouch', and so on, seems to leave out the most important part – its painfulness. Emergentists such as David Chalmers say this is because phenomenal *properties* are importantly different to functional properties such as

cohesiveness, in that what is essential to them is the way they *feel*, and not the way they *behave*. The phenomenal emerges from the physical according to bridge laws, but does not add anything to it causally. On this view, there is a possible world physically identical to this one throughout its history – at which a physical doppelgänger of you is reading this article – but at which nothing is conscious. Weak ontological emergentists think that emergent properties are non-deducible because they are not functional properties, and this is why we cannot conceive them as such.

3. Epistemological emergence

'Type-B physicalists' (this term is due to Chalmers) agree that phenomenal properties are not deducible from physics, and that this is due to our inability to conceive them in functional terms, but deny that they are ontologically distinct from the physical. On this view, painfulness is a functional property, similar to cohesiveness. Our reluctance to accept this is explained by reference to the special nature of phenomenal *concepts*, rather than the *properties* those concepts represent. We make a mistake when we think consciousness could not be physical – the kind of mistake Lois Lane makes when she thinks Clark Kent could not be Superman. Phenomenal concepts are distinctive ways of thinking *about the brain*, which fool us into thinking that functional analyses of consciousness leave something out. Type-B physicalists think that phenomenal properties are non-deducible because phenomenal concepts make it very difficult for us to conceive of those properties in functional terms. The concept of pain, for instance, makes it very difficult for us to see pain as a property exhausted by its function. Being in pain, however, is as functional as being a mousetrap. It is our way of *thinking* about pain that makes it non-deducible, not the nature of pain itself. Hence the form of emergence here is epistemological.

Type-B physicalism is not the only kind of epistemological emergence. Not all mental properties are difficult to conceive functionally – beliefs and desires, for instance, seem more amenable to such analyses than phenomenal properties. Still, one might say, even these properties are non-deducible *in practice*, due to the extreme complexity of the physics involved. The forms of emergence considered thus far are *synchronic* – phenomenal properties, for instance, are taken to be non-deducible from the physical properties brains have *at the same time*. Weather systems provide an example of *diachronic* epistemological emergence (as do other systems whose dynamics are chaotic). Since the weather at any given time is extremely sensitive to slight changes in earlier conditions (the 'butterfly effect'), it is impossible for us to deduce exactly when, in the future, a weather pattern such as a tornado will appear. The property of being a tornado is functional, and tornados are nothing over and above the air molecules that compose them. Still, we cannot deduce where or when they will occur. Such patters are therefore epistemologically emergent with respect to physical conditions *at earlier times*. In general, epistemological emergence about a domain of properties differs from ontological emergence in that it explains the non-deducibility of emergent properties by reference to human nature, rather than the nature of the emergent properties. In the scientific literature, 'emergence' is almost always used in its epistemological sense, and care must be taken not to conflate this usage with the far more controversial ontological variants discussed above.

## Conclusion

Progress in science has reduced the appeal of strong ontological emergence, through mounting evidence that all causal powers that physical things possess are contributed solely by their physical properties, according to physical laws. Weak ontological emergence about consciousness, however, is more popular, and lively philosophical debates over the relative

merits of this position and type-B physicalism are ongoing. Since the laws of physic comes out the same either way, it is difficult to see how any amount of scientific progress could help decide this issue. By far the liveliest area of current research on emergence concerns the various epistemological kinds, since these are consistent with the widespread commitment to physicalism among both philosophers and scientists.

## Further Reading

Beckermann, A., Flohr, H. and Kim, J. (eds.) 1992. *Emergence or Reduction: Essays on the Prospects of Nonreductive Physicalism*. (Berlin: de Gruyter)

Broad, C. D. 1925. *The Mind and its Place in Nature*. (Routledge and Kegan Paul)

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

Chalmers, D. 2003. 'Consciousness and its Place in Nature', in Stich, S. and Warfield, F. *The Blackwell Guide to Philosophy of Mind* (Oxford: Blackwell).

Hohwy, J and Kallestrup, J. 2008. *Being Reduced: New Essays on Reduction, Explanation and Causation*. (Oxford University Press)

Jackson, F. 1982. 'Epiphenomenal Qualia' *Philosophical Quarterly*, 32: pp. 127-36.

Kim, J. 2005. *Physicalism, Or Something Near Enough*. (Princeton University Press)

Metzinger, T. (ed.) 2000. *Neural Correlates of Consciousness*. (MIT Press)

Nagel, T. 1974. 'What is it Like to be a Bat?' *Philosophical Review*, 83: pp. 435-50.

Papineau, D. 2002. *Thinking About Consciousness*. (Oxford University Press)