Two Aspects of Emergence

Peter Wyss

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

I. Introduction

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

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

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

II. The Legacy of Classical Emergentism

The resurgent interest in emergence during the past decade has produced not only a staggering amount of published work (more than seventy publications since 1995), but also a confusing variety of emergentisms, all tailor–made to specific purposes. Now that the dust is settling after the renaissance of emergence, there are attempts to chart and classify the various ideas, and a proper re–evaluation has only just begun (see, e.g., Gillett, 2002; O'Connor & Wong, 2002; Silberstein, 2002; Stephan, 1999b; Van Gulick, 2001). Although it is common today to present a coherent picture of classical emergentism (doubtlessly coating the current brands in authority and dignity), this conceals the fact that emergentism in its heyday was as diverse as it is

today. Now the diversity is truly enormous: it ranges from the cognitive emergence of new ways of thinking to the emergence of individuals with free will that survive death. Rather than providing a survey of the current conceptions, let me revisit the stalemate between vitalism and mechanism of which emergentism was the product, followed by a brief look at the current context of revival. I will then extract two central ontological posits and the characteristic attitude of emergence. First, here are simplified versions of vitalism and mechanism.

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

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

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

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

Furthermore, physicalism is either reductive or non–reductive. Resembling mechanism in its rigor, *reductive* physicalism regards the mental as reducible to the physical, often in the sense that mental properties are identical with, or 'nothing but', physical properties. (In a weaker sense, reduction means that the mental is explainable in terms of the physical.) *Non–reductive* physicalism denies the identity of mental and physical properties. Instead, it is suggested that mental properties supervene on physical properties. Here is a rough and ready characterisation of supervenience. For two families of properties *P* and *Q*, *P* supervenes on *Q*, if two things that are indiscernible with respect to *Q* are indiscernible with respect to *P* (cf. Kim, 1998, pp. 9-10). Alternatively, an instance of *P* supervenes on an instance of *Q*, if there can be no change in *P* without corresponding change in *Q* (though there may be *Q*–change without *P*–change). For a while non–reductive physicalists thought that mind–body supervenience could establish a position tight enough to go through as physicalism while being loose enough as *not* to entail reduction. But doubts have been raised recently whether supervenience can in fact substantiate these hopes (see, e.g., Heil, 1998; Kim, 2003a). It is striking that interest in emergence resumed when enthusiasm for supervenience started to subside roughly ten years ago.

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

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

The first feature is the *distinctiveness* of emergent properties. The idea is that an emergent property *P* individuates the thing that has it as something new—a piece of matter that instantiates vitality is no longer a *mere* piece of matter but *also* an organism (Alexander, 1920, ii, pp. 45f.). As it were, *P* confers a new identity on the thing that has it. Combined with the idea that nothing is real unless it has causal powers— that is, having causal powers is the mark of reality (Alexander, 1920, ii, p. 8)—, *P* is distinctive in the sense that it makes a causal difference to the thing that has *P*, that is, *P* endows it with new causal powers. (Otherwise, we would not even know that something has *P*.) Hence epiphenomenalism is rejected. Distinctiveness also means that the causal powers of emergent properties are irreducible and fundamental. The thought is this: if the causal powers of emergent properties were reducible to (or identical with) the causal powers of their base properties, they would be neither new nor distinct. As a consequence, the causality of emergent properties is inexplicable in terms of, and theoretically unpredictable from, those of their base properties.

The second feature is the *dependence* of emergent properties. The idea that life and mind depend on some basis makes emergentism a naturalist position. Dependence has a temporal (diachronic) and a non–temporal (synchronic) sense. Lloyd Morgan writes:

> '[I]f by Vitalism we give expression to the fact that living matter has certain distinctive properties, it may be freely accepted; but […] if by it we imply that these properties neither are nor can be *the outcome of evolution*, it should be politely rejected […]' (1899, p. 196, my emphases).

The message is clear: we should renounce miraculous extra–bits in our cosmology, if only because we have no idea of how matter, life, and mind relate (see Lloyd Morgan, 1923, pp. 12f.). For emergentism, all there is has originated *from within*, or, to use Alexander's phrase, 'blossomed out' of, nature (1920, ii, p. 52). The idea that everything is ultimately made of some one stuff is compatible with mechanism and physicalism. However, if all change is the result of a regrouping of components that have a few intrinsic properties such as mass or charge and are governed by a couple of basic laws, then it is hard to see how anything new could ever appear (cf. Broad, 1925, p. 76; Lloyd Morgan, 1923, p. 2).

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

First of all, there is epistemic modesty and humility. Alexander writes:

> 'The existence of emergent qualities […] is something to be noted, as some would say, under the compulsion of brute empirical fact, or, as I should prefer to say in less harsh terms, to be accepted with the "natural piety" of the investigator. It admits no explanation' (1920, ii, pp. 46f.).

There are questions, he thinks, 'we have no right to ask' (1922, p. 620), as there are facts we must accept in the mood of 'reverent temper' (*op. cit.*, p. 621). Natural piety, in other words, is 'the habit of knowing when to stop in asking questions of nature' (*op. cit.*, p. 609). By embracing a stance of 'deliberate innocence', or 'strenuous *naïveté*', we prevent ourselves from asking silly questions or filling the explanatory void with some purpose–built entity (Alexander, quoted by Muirhead, 1939). When faced with explanatory and purported ontological gaps, emergentism implores us to stay cool and relax.

Second, there is a resistance to excessively optimistic scientism and explanatory chauvinism. The successes of mechanistic explanation and the reductive strategy do not carry over to the cases of life and consciousness. Aiming at explaining the complex in terms of the simple certainly is a valuable heuristic. But its downside is a

tendency to over–simplify facts. As Broad sharply remarks of mechanism, it 'of course reaches its wildest absurdities in the attempts […] to treat mental phenomena mechanistically' (1925, p. 77). Besides, as participants in the vitalism/mechanism debate had access to the *same* biological knowledge, their radically different outlooks must have been driven not by facts, but by certain explanatory pre–conceptions. Emergentism tells us not to let our metaphysics be compelled by our favourite explanatory strategies.

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

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

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

IV. Emergence at Work: a Twofold Problem for Supervenience

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

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

In contrast, non–reductive physicalists hold that mental properties *supervene* on physical properties without being reducible or identifiable with them. They think that mind–body supervenience secures a minimal physicalism, because supervenience incorporates mind–body dependence and hence the primacy of the physical. Here is the first problem: emergentism also accepts mind–body supervenience and thus seems to *coincide* with non–reductive physicalism (or minimal physicalism). Both emergentists and non–reductive physicalists are confident they can keep mental properties irreducibly distinct from, yet somehow grounded in, physical properties. Together they resist being wheedled into accepting that *A*s really are 'nothing but' *B*s; or that *A*s are much better understood in terms of *B*s; or that the *B*s explain the *A*s; or that there are no *A*s as we know them. Because of this, some philosophers think there are no, or only marginal, differences between emergentism and non–reductive physicalism (see, e.g., Crane, 2001; Kim, 1993; Pereboom, 2002). Supervenience is too weak a criterion to separate emergentism from non–reductive (minimal) physicalism.

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

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

Since unexplained supervenience is accepted as a 'brute fact', and, in contrast to physicalists, they do not subscribe to the thesis of explanatory completeness,

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

V. So, Why Does Emergence Reappear?

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

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

To sum up. I suggest that part of the *appeal* of emergence is its expression of an attitude of metaphysical humility, which is fostered and motivated by discontent or disquiet about extreme positions such as physicalism. In turn, some of the re–thinking in the current metaphysics of mind is due to the resilient reappearance of emergence. The ‘metaphysical discomfort’ (see Kim, 1999, p. 28) that some physicalists feel

towards emergence does not reflect troublesome aspects of emergence, but reveals doubts about physicalism itself.

Peter Wyss
Birkbeck College
University of London

References

Alexander, S. (1920). *Space, Time, and Deity. The Gifford Lectures at Glasgow 1916-1918 (2 Vols.).* London: Macmillan.

Alexander, S. (1922). Natural Piety. *The Hibbert Journal, 20*, 609-621.

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

Broad, C. D. (1933). *Examination of McTaggart's Philosophy, Vol. I, 'The 'Nature' of a Continuant' (pp. 264- 278).* Cambridge: Cambridge University Press.

Crane, T. (2001). The Significance of Emergence. In C. Gillet & B. Loewer (Eds.), *Physicalism and Its Discontents* (pp. 207-224). Cambridge: Cambridge University Press.

Gillett, C. (2002). The Varieties of Emergence: Their Purposes, Obligations, and Importance. *Grazer Philosophische Studien, 65*, 95-121.

Heil, J. (1998). Supervenience Deconstructed. *European Journal of Philosophy, 6*, 146-155.

Hoernle, R. F. A. (1918). Mechanism and Vitalism. *The Philosophical Review, 27*, 628-645.

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

Horgan, T. (1994). Physicalism. In S. Guttenplan (Ed.), *A Companion to the Philosophy of Mind* (pp. 471-479). Oxford: Blackwell.

Kim, J. (1993). The Non–Reductivist's Troubles with Mental Causation. In J. Heil & A. Mele (Eds.), *Mental Causation* (pp. 189-210). Oxford: Oxford University Press.

Kim, J. (1996). *Philosophy of Mind*. Boulder, CO: Westview Press.

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

Kim, J. (1999). Making Sense of Emergence. *Philosophical Studies*, 95, 3-36.

Kim, J. (2001). Mental Causation and Consciousness: The Two Mind-Body Problems for the Physicalist. In C. Gillett & B. Loewer (Eds.), *Physicalism and Its Discontent*s (pp. 271-283). Cambridge: Cambridge University Press.

Kim, J. (2003a). Supervenience, Emergence, Realisation, Reduction. In M. J. Loux & D. W. Zimmerman (Eds.), *The Oxford Handbook of Metaphysics* (pp. 556-584). New York: Oxford University Press.

Kim, J. (2003b). *Being Realistic about Emergence*. Unpublished Manuscript. Lloyd Morgan, C. (1895). Naturalism. *The Monist, 6*, 76-90.

Lloyd Morgan, C. (1899). Vitalism. *The Monist*, *9*, 179-196.

Lloyd Morgan, C. (1923). *Emergent Evolution. The Gifford Lectures at St. Andrews 1922*. London: Williams and Norgate.

Loewer, B. (2001). From Physics to Physicalism. In C. Gillett & B. Loewer (Eds.), *Physicalism and Its Discontents* (pp. 37-56). Cambridge: Cambridge University Press.

Lowe, E. J. (2000). Causal Closure Principles and Emergentism. *Philosophy, 75*, 571-585.

Muirhead, J. H. (1939). Samuel Alexander. *Philosophy, 14*, 3-14.

Nagel, E. (1961). *The Structure of Science. Problems in the Logic of Scientific Explanation*. London: Routledge & Kegan Paul.

O'Connor, T., & Wong, H. Y. (2002). *The Metaphysics of Emergence*. Unpublished Manuscript.

Papineau, D. (2001). The Rise of Physicalism. In C. Gillett & B. Loewer (Eds.), *Physicalism and Its Discontents* (pp. 3-36). Cambridge: Cambridge University Press.

Pereboom, D. (2002). Robust Non-reductive Materialism. *Journal of Philosophy*, *99*, 499-531.

Shelley, M. (1818). *Frankenstein, or the Modern Prometheus*. Oxford: Oxford University Press (Reprint 1992).

Silberstein, M. (2002). Reduction, Emergence, and Explanation. In P. Machamer & M. Silberstein (Eds.), *The Blackwell Guide to the Philosophy of Science* (pp. 80-107). Oxford: Blackwell.

Stephan, A. (1999a). *Emergenz. Von der Unvorhersagbarkeit zur Selbstorganisation*. Dresden: Dresden University Press.

Stephan, A. (1999b). Varieties of Emergentism. *Evolution and Cognition, 5*, 49-59.

Van Gulick, R. (2001). Reduction, Emergence, and Other Recent Options on the Mind/Body Problem. *Journal of Consciousness Studies*, *8*, 1-34.