(from *Erkenntnis* 46, No. 3 (1997), 305-314)

Armchair Arguments Against Emergentism

Achim Stephan

Many philosophers associate 'Emergentism' with a philosophical doctrine which had its heydays in the nineteen-twenties and is therefore a position deserving historical interest only. Recent work by Ansgar Beckermann (1992), Terence Horgan (1993), and Jaegwon Kim (1992, 1993a) shows, however, that theories of emergence are valuable for the philosophy of mind, particularly for the evaluation of nonreductive physicalism.

Particularly Jaegwon Kim, in a number of papers, has developed an intricate argument against nonreductive physicalism. Among other things, he claims that nonreductive physicalism is a special form of emergentism. If this is so, then every argument against emergentism is also an argument against nonreductive physicalism. And worse, according to Kim, emergentism faces a hopeless dilemma when confronted with the problem of mental causation: Either emergentism must give up the claim that mental properties are real properties, or it must accept that mental properties have downward causal powers, which has the implication that the physical realm is not causally closed.

There are a few philosophers willing to bite the bullet: Terence Horgan, for example, suggests that we carefully explore mental irrealism as a possible physicalist position (see his 1993, 581). On the other hand there are philosophers who postulate some kind of downward causation, as Roger Sperry has done for years (e. g. in 1969, 1980, and 1991) and Timothy O'Connor very recently (1994, 98). I intend to show that emergentists need not choose between mental irrealism and downward causation. They can interpret mental causation as a case of supervenient causation - similar to the local reductionist.[1] Kim himself once favored this way of interpreting mental causation.[2]

---

[1] So-called 'local reductions' are generated by biconditional bridge laws restricted to *specific* species and structures, whereas unrestricted bridge laws sanction 'global' or 'uniform' reductions (see Kim 1996, 233 f.).

[2] Very recently, however, Kim has questioned whether supervenient causation is enough to guarantee physicalism; I will come to this point in Section 4.

This paper has five parts. In view of the immense variety of emergentist doctrines, I will begin with a concise description of the characteristic features of emergence theory relevant to Kim's attack against nonreductive physicalism. Second, I'll argue that emergentists can see mental causation as supervenient causation without making implausible assumptions. In Part Three, I'll discuss Kim's claim that emergentists don't have the option of treating mental causation as supervenient causation. In Part Four I'll take up Kim's recent worries about supervenient causation, and suggest how to handle mental causation as a case of 'superdupervenient' causation. This option, however, would definitely be barred to emergentists. And, finally, I raise the issue whether macro-properties must superdupervene on their micro-base-properties in order to be causally efficacious.

1. Emergentism

Theories of Emergence have many faces (see Stephan 1992), but when debating nonreductive physicalism, only so-called *synchronic property emergentism* is relevant. Compared to other theories of emergence, property emergentism claims (i) the emergence of certain macro-properties in complex systems (and not, for example, the emergence of forces, effects, laws, events, or systems as such). (ii) *Synchronic* emergentism focuses on the idea that macro-properties of complex systems are irreducible to the properties and relations of their parts; it is not concerned - as is so-called *diachronic* emergentism - with new qualities or forces that might come up unpredictably during the course of evolution.

Synchronic property emergentism and reductive physicalism agree on several claims:

(1) Both synchronic property emergentism and reductive physicalism proceed from a basic physical ontology. The parts out of which all natural systems are made up are physical, and basic properties and forces are physical.

(2) Emergentists and reductionists agree that physical systems may have *systemic* properties; this means that systems may have properties that no individual part of the system possesses. Systemic properties are held to be *real* properties, at least as real as the basic properties.

(3) Emergentism and reductionism share the idea of *mereological supervenience*: Systemic properties supervene on the properties and relations of the parts of the system. In other words, there can be no difference in the systemic properties without there being a difference in the properties and relations of the parts.

(4) In contrast to the reductionist, the emergentist divides systemic properties into two groups: *resultant* systemic properties and irreducible, i.e., *emergent* systemic properties. The claim is made that there are emergent properties, namely those that, unlike resultant properties, cannot be reduced to the properties and relations of the parts of the system.

Horgan points out just this distinction between the ontological relation of supervenience and the much more ambitious explanatory relation of what he calls 'superdupervenience'. The latter relation holds between two property families, if the ontological relation of supervenience between them can be accounted for physically (cf. Horgan 1993, 560, 566, 577-81). To put it in Horgan's terms: emergentism claims that there are systemic properties which do not *superduper-vene* on the properties and relations of the system's parts, although they supervene mereologically on them.

2. Mental Causation

If you concede that mental properties have causal powers, while accepting at the same time the causal closure of the physical domain, then you must consider the causal role of mental properties to be somehow dependent on the causal role of physical properties. Since multiple realizability of mental properties counts against their identification with physical properties, the only choice that remains is to interpret mental causation as a case of supervenient causation. According to Kim (1984), supervenient causation is a type of macro-causation that allows the ascription of a sufficiently strong causal role to systemic properties, without treating them as mere fictions:

A macrocausal relation between two events, x's having F and y's having G, where F and G are macroproperties will be called a supervenient causal relation if and only if x's having F supervenes on x's having m(F), y's having G

supervenes on y's having m(G), where m(F) and m(G) are microproperties relative to F and G, and there is an appropriate causal connection between x's having m(F) and y's having m(G) (cf. Kim 1984, 262).

A central question to the emergentist then is whether or not mental causation can be interpreted as supervenient causation.[3] The first feature of emergentism, physical monism, is explicitly part of the definition of supervenient causation, and so is the third characteristic, namely mereological supervenience of the systemic to the micro-properties. Moreover, says Kim, supervenient causation appears to guarantee the reality of supervenient properties: „supervenient causal relations ... are among the ones that are 'real'" (1984, 265). But, the emergentist's claim that certain macro-properties are irreducible may be problematic. This feature is not implied by the relation of supervenient causation. Let us see, whether they are at least compatible.

If you hold a less ambitious concept of reduction (e.g. only that bridge laws exist between macro- and micro-properties), mereological supervenience of the former on the latter may already allow the reduction of systemic properties to properties and relations of the system's parts. However, this would render the four characteristics which make up emergentism inconsistent among themselves, and you could dispense with the argument via mental causation. Kim, however, has claimed that emergentists base their claims on a very strong notion of reduction (1992, 124-27). Reduction, in their ambitious sense, means to make the very bridge laws or the ontological relation of supervenience intelligible. Since the relation of supervenience alone does not provide an explanatory relation between supervenient and subvenient properties, neither does the relation of supervenient causation. Thus, the emergentist can have both: the irreducibility of mental properties, and their entering into supervenient causation.

[3] The question whether supervenient causation is sufficient for causation has been raised by Brian McLaughlin (1983) and Lynne Rudder Baker (1993). I will not address this issue, but see the discussion below.

3. Armchair arguments[4] I

Jaegwon Kim, however, has argued strongly against the view that emergentists could interpret mental causation as supervenient causation. His argument against the compatibility of the two positions can be put in the following way (cf. Kim 1984; 1992, 134-36; and 1993b, 203-6):

(1) Emergent mental properties are real properties (the thesis of emergent realism).
(2) To be a real property is to have causal powers ('Alexander's Dictum'[5]).
(3) Emergent mental properties are irreducible (the irreducibility thesis).
(4) To be an irreducible property, is to be a genuine new addition to the ontology of the world (the ontology thesis).
(5) To be an irreducible real property, is to have new and irreducible causal powers. (From 2 and 4.)
(6) Emergent mental properties have irreducible causal powers. (From 1, 3, and 5.)
(7) A property which can only be a supervenient cause has no genuine and irreducible causal powers; its causal role is reducible to the causal role of the subvenient physical properties and relations (reducibility of supervenient causation).
(8) Thus: Emergentism is not compatible with the idea of supervenient causation. (From 6 and 7.)

At first sight Kim's argument seems to be extremely plausible. A closer look, however, reveals several flaws. Premises (1) and (3) can be granted, because they just repeat the second and fourth features of synchronic property emergentism. Alexander's Dictum might be disputed, though; and if so, then the power of

[4] Some have claimed that it is merely an empirical question whether emergentism does hold or not; e.g. Meehl and Sellars (1956). McLaughlin, for example, has argued that advances in science, not philosophical criticism, led to the fall of British Emergentism (1992, 90).

[5] The thesis that to be a real property is to have causal powers goes back to Samuel Alexander (cf. 1920, Vol. 2, p. 8); it was dubbed 'Alexander's Dictum' by Jaegwon Kim.

Kim's argument can be seen to wane. However, the emergentist's gain would be costly, for she would be forced to treat mental properties as epiphenomena. To avoid this consequence, it would be better to accept premise (2), i.e. to accept that mental properties have causal powers. Obviously, premise (5) is central to Kim's argument; he states it as follows:

> „To be real, Alexander has said, is to have causal powers; *to be real, new, and irreducible, therefore, must be to have new, irreducible causal powers*" (1993b, 204; cf. also 1992, 135).

However, Kim's conclusion (5) is not compelling. Since the application of Alexander's Dictum upon *irreducible* properties seems to warrant two different propositions, namely:

(5a) To be irreducible and real, is to be irreducible and to have causal powers.

(5b) To be irreducible and real, is to have irreducible causal powers.

Proposition (5b) alone corresponds to Kim's (5). (5a) does not contradict (7), if it is possible that an irreducible property might possess reducible causal powers. To avoid this possibility, Kim would need some further premise in order to exclude variant (5a). E. g.:

(4*) If a property's causal powers are reducible, then this property is not irreducible.

Premise (4*) along with the irreducibility claim of the property, implies the irreducibility of the property's causal powers. Kim does not explicitly claim something like (4*); however, it seems that he did assume such a premise implicitly - having in mind an argument somewhat like the following:

(4a) If property P's causal powers are reducible, then property P does not make a genuinely new causal contribution.

(4b) If property P does not make a genuinely new causal contribution, then P is not a genuinely new property.

(4c) If property P is not a genuinely new property, then P is not irreducible.

(4*) Thus: If property P's causal powers are reducible, then P is not irreducible.

If you add this formally valid argument to Kim's argument, then premise (5a) is excluded, and Kim's argument against emergentism would then be conclusive. However, this argument is itself problematic. Though premise (4a) seems to be acceptable, either premise (4b) or premise (4c), which corresponds to the (ontological) premise (4) of Kim's argument, turn out to be its Achilles' heel. Which one, depends on the reading of 'genuinely new property'. Kim introduces this notion in accordance with Lloyd Morgan (1923, 64 f.). The property of weighing exactly 174.34556899999 lbs., Kim says, might be a 'new' property, a property that has never been exemplified till now; however, it would not be a *genuinely* novel one in the sense the emergentist has in mind. To be genuinely new the property must also be irreducible (see 1996, 227 f.). And, as we have seen, to the emergentist 'irreducible' always means '*explanatorily* irreducible'.[6] If we adopt this reading of 'genuinely new', (4c) seems to be warranted.

But now, what about premise (4b)? Let P be a property mereologically supervenient on microproperties m_i; to make a genuinely new causal contribution for P would be to make a causal contribution different from the causal contribution the subvenient microproperties m_i make. Thus, the claim that property P does not make a genuinely new causal contribution is a plain ontological claim, it is not an explanatory claim. Moreover, it does not license the further claim that P is not a genuinely new property, if 'genuinely new' means to be *explanatorily* irreducible. Premise (4b) only seems to be acceptable, if we do not adopt the emergentist's reading of 'genuinely new'. Then, however, a less ambitious reading of 'genuinely new' is mingled with its more ambitious reading within (4c), and this would render premise (4c) unwarranted. In either way, Kim's argument against emergentism seems to fail at the point of shifting from a less ambitious to a more ambitious notion of reducibility. His argument just lacks the

[6] Remember that reduction, in the emergentist's ambitious sense, means to *explain* the ontological relation of supervenience which is presupposed between macro- and micro-properties. Thus, we should distinguish the relation of being ontologically reducible (via bridge laws or the relation of supervenience) and the explanatory relation of making the ontological relations itself intelligible. A macro-property, therefore, might be ontologically reducible to a system's micro-properties without being explanatorily reducible to them. No modal relation will by itself guarantee an explanatory relation.

last and decisive brick; but the missing premise - „if a property is reducible ontologically (via bridge laws or the relation of supervenience), then it is reducible explanatorily (by making those relations themselves intelligible)“ - is hardly plausible. For the missing premise to be true, would mean that each relation of supervenience could be made intelligible. This is not the case.[7]

In accordance, proposition (7) needs the distinction between an ontological and an explanatory reading of irreducibility. Kim's statements on supervenient causation and the phrasing of (7) suggest that he argues ontologically. So premise (7) might be accepted by most readers. Interpreted in an explanatory manner, (7) comes out false, since the existence of supervenient causal relations does not by itself imply explanatory reductions to subvenient causal relations.

Thus we have the following situation: The second order property of 'being irreducible' is attributed to both properties *and* causal relations. (i) If we interpret both of them ontologically, then Kim's argument against emergentism appears to be formally valid. The emergentist's claim (premise 3), however, is not represented adequately, since synchronic property emergentism reads 'irreducible' explanatorily. (ii) If we interpret 'irreducible' in the explanatory sense for properties, then the emergentists position is grasped adequately, but then, a plausible argument for premise (5) is missing. (iii) If we use 'irreducible' in both cases explanatorily, again Kim's argument is valid formally, and the emergentists position is grasped correctly; premise (7), however, would then be false. Thus Kim's argument leaves enough room for the emergentist: He may interpret - like the reductionist - mental causation as supervenient causation.

4. Armchair arguments II

Kim's arguments against nonreductive physicalism are not easy to tackle. In his postscript to „Supervenience and Mind“ Kim questions whether supervenient causation is really sufficient for treating mental causation physicalistically: „Is it a robust enough relation to vindicate the causal efficacy of the mental?“ (1993c, 359). His main worry seems to be that supervenient causes

[7] See, e. g., Levine's 'explanatory gap'-argument (1993, 130-134).

may just turn out to be epiphenomenal. A relation, however, that would be strong enough to exclude epiphenomenalism, has to be one that would treat the mental at least in the same way as ordinary physicalistically explainable macrophenomena. A relation that would be sufficient for this purpose is what I would like to call *superdupervenient causation*. I introduce this notion following Lycan's and Horgan's expression of *superdupervenience*: superdupervenience is ontological supervenience that is robustly explainable in a materialistically explainable way (cf. Horgan 1993, 566). This would amount to the following:

A macrocausal relation between two events, x's having F and y's having G, where F and G are macroproperties will be called a *superdupervenient* causal relation if and only if x's having F superdupervenes on x's having m(F), y's having G superdupervenes on y's having m(G), where m(F) and m(G) are microproperties relative to F and G, and there is an appropriate causal connection between x's having m(F) and y's having m(G).

It should be clear that if nothing less than the relation of superdupervenient causation is robust enough to vindicate the causal efficacy of the mental, then emergentists have no chance to account for mental causation within a materialistic frame. The relation of mereological superdupervenience between physical and mental properties is exactly what is neglected by emergentism.

5. An open question

I have distinguished the possibility of mental phenomena being supervenient causes from the possibility of their being superdupervenient causes. The first allows emergentists - against Kim - to treat mental phenomena as emergent phenomena that play a genuine causal role without exerting a downward influence; the second, however, would not allow emergent phenomena to be causally efficacious this way. If superdupervenient causation is required for causal efficacy within a materialistic frame, then emergentism indeed faces Kim's dilemma. Emergentists, however, might question whether superdupervenient causation is so required. They might maintain that supervenient causation suffices for causation; since they can allow that emergent phenomena are supervenient, they can, then, maintain as well that mental phenomena are causally efficacious.

Thus, to reject the emergentists's claim that mental phenomena are both emergent and causally efficacious, Kim's early view that supervenient causation suffices for causation would have to be rejected.

Supervenient causation may not suffice for causation.[8] However, it won't do simply for Kim to maintain that it fails to do so. For his purpose is to argue not merely that we lack an account of how emergent properties could be causally efficacious without downward causation, but, rather, that emergent properties could not be causally efficacious without downward causation. He seems to assume that for mental properties to be causally efficacious without exerting a downward influence, they would have to be explanatorily reducible, that is, they would have to superdupervene. But, then, he owes us a case for the claim that macro-properties must superdupervene on their micro-bases to be causally efficacious without exerting a downward influence. Neither he nor anyone else, for that matter, has made such a case. It thus remains an open question whether emergent phenomena can be causally efficacious without downward causation. Kim has failed to block a pass through the horns of his dilemma.[9]

Literature

Alexander, S.: 1920, *Space, Time, and Deity. 2 Vols*., Macmillan, London.

Baker, L. R.: 1993, 'Metaphysics and mental Causation', in J. Heil, A. Mele (eds.) *Mental Causation*, Clarendon Press, Oxford, pp. 75-95.

Beckermann, A.: 1992, 'Supervenience, Emergence, and Reduction', in A. Beckermann, H. Flohr, J. Kim (eds.) *Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism,* Walter de Gruyter, Berlin, New York, pp. 94-118.

Horgan, T.: 1993, 'From Supervenience to Superdupervenience: Meeting the Demands of a Material World', *Mind* 102, pp. 555-586.

Kim, J.: 1984, 'Epiphenomenal and Supervenient Causation', *Midwest Studies in Philosophy* IX, pp. 257-270. Reprinted in: J. Kim, *Supervenience and Mind*, Cambridge University Press, Cambridge 1993, pp. 92-108.

[8] See the references in footnote 3.

[9] I am very grateful to Ansgar Beckermann, Jaegwon Kim and Brian McLaughlin for their helpful comments on earlier drafts of this paper.

Kim, J.: 1989, 'The Myth of Nonreductive Materialism', *Proceedings and Addresses of the American Philosophical Association* 63, pp. 31-47. Reprinted in: J. Kim, *Supervenience and Mind*, Cambridge University Press, Cambridge 1993, pp. 265-284.

Kim, J.: 1992, ''Downward Causation' in Emergentism and Nonreductive Materialism', in A. Beckermann, H. Flohr, J. Kim (eds.) *Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism,* Walter de Gruyter, Berlin, New York, pp. 119-138.

Kim, J. 1993a, *Supervenience and Mind,* Cambridge University Press, Cambridge.

Kim, J. 1993b, 'The Nonreductivist's Troubles with Mental Causation', in J. Heil, A. Mele (eds.) *Mental Causation*, Clarendon Press, Oxford, pp. 189-210. Repr. in: J. Kim, *Supervenience and Mind*, Cambridge University Press, Cambridge 1993, pp. 336-357.

Kim, J.: 1993c, 'Postscripts on Mental Causation', in J. Kim, *Supervenience and Mind*, Cambridge University Press, Cambridge 1993, pp. 358-367.

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

Lloyd Morgan, C.: 1923, *Emergent Evolution*, Henry Holt, New York; Williams and Norgate, London.

McLaughlin, B.: 1983, 'Event Supervenience and Supervenient Causation', *The Southern Journal of Philosophy* 22, *Supplement: Supervenience*, pp. 71-91.

McLaughlin, B.: 1992, 'The Rise and Fall of British Emergentism', in A. Beckermann, H. Flohr, J. Kim (eds.) *Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism,* Walter de Gruyter, Berlin, New York, pp. 49-93.

Meehl, P. E., Sellars, W.: 1956, 'The Concept of Emergence', in H. Feigl, M. Scriven (eds.) *The Foundations of Science and the Concepts of Psychology and Psychoanalysis. Minnesota Studies in the Philosophy of Science. Vol. 1*, University of Minnesota Press, Minneapolis, pp. 239-352.

O'Connor, T.: 1994, 'Emergent Properties', *American Philosophical Quarterly* 31, pp. 92-104.

Sperry, R.: 1969, 'A Modified Concept of Consciousness', *Psychological Review* 76, pp. 532-536.

Sperry, R.: 1980, 'Mind-Brain Interaction: Mentalism, Yes; Dualism, No', *Neuroscience* 5, pp. 195-206.

Sperry, R.: 1991, 'In Defense of Mentalism and Emergent Interaction', *The Journal of Mind and Behavior* 12, pp. 221-245.

Stephan, A.: 1992, 'Emergence - A Systematic View on its Historical Facets', in A. Beckermann, H. Flohr, J. Kim (eds.) *Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism,* Walter de Gruyter, Berlin, New York, pp. 25-48.