Donald Davidson introduces supervenience to the philosophy of mind in his "Mental Events". He famously writes that "mental characteristics are in some sense dependent, or supervenient, on physical characteristics" (1970:214). Curiously, however, there has been little effort to explicate what Davidson means by supervenience; philosophers typically assume Jaegwon Kim's conception of supervenience. My aim here is to explicate the passages in which Davidson discusses supervenience. I argue that Davidson's supervenience is very different from the one assumed in contemporary philosophy of mind, and is not dependence in the sense of some deeper metaphysical relation.

2. The characterization of supervenience

Davidson characterizes supervenience in several places. In "Mental Events", he writes:
Supervenience might be taken to mean that there cannot be two events alike in all physical respects but differing in some mental respect, or that an object cannot alter in some mental respects without altering in some physical respects (1970:214).

The first part of the sentence is a characterization in terms of indiscernibility, namely, that "there cannot be two events alike in all physical respects but differing in some mental respect", that is, there cannot be two events that are physically indiscernible but mentally discernible. The second part is another characterization, one which construes it as a type of covariance: "an object cannot alter in some mental respects without altering in some physical respects", that is, mental changes co-vary with physical changes.

In later writings, Davidson provides additional covariance definitions, the gist of which is that any mental difference between objects must be accompanied by a physical difference. In "Reply to Harry Lewis", he writes:

The notion of supervenience, as I have used it, is best thought of as a relation between a predicate and a set of predicates in a language: a predicate \( p \) is supervenient on a set of predicates \( S \) if for every pair of objects such that \( p \) is true of one and not of the other there is a predicate in \( S \) that is true of one and not of the other. (1985:242)

And in "Thinking Causes" he makes a similar claim:

The idea I had in mind is, I think, most economically expressed as follows: a predicate \( p \) is supervenient on a set of predicates \( S \) if and only if \( p \) does not distinguish any entities that cannot be distinguished by \( S \). (1993:4)

On a charitable reading, Davidson’s characterizations are all equivalent. Supervenience is a thesis about the relations between properties or characteristics or respects, e.g., mental and physical properties, which Davidson understand as predicates. These properties are ascribed to particulars such as events, objects and entities. To make things more explicit, let us take two sets of properties, \( R \) and \( S \). We can think of \( R \) as a set of mental predicates, and of \( S \) as a set of physical properties.

We would say that \( R \) supervenes on \( S \) just in case the following condition holds:

For every \( M \) of \( R \) and for every pair of objects (events, entities) \( x \) and \( y \), if for every \( P \) of \( S \), \( Px \leftrightarrow Py \) (i.e., \( x \) and \( y \) are \( S \)-indiscernible), then \( Mx \leftrightarrow My \) (i.e., \( x \) and \( y \) are \( M \)-indiscernible).
Let us compare this characterization to Kim's notions of supervenience. First, Kim (1984) famously distinguishes between a strong and a weak reading of this condition. On the strong reading, the condition applies to every pair of possible objects x and y, even if they inhabit "different worlds". On the weak reading, it applies to every pair of objects belonging to the same world (any world), but need not apply to objects across worlds. In "Thinking Causes", Davidson says that his version of supervenience is of the weak sort:

Kim himself (correctly, I think) finds my version of supervenience very close to his 'weak' supervenience, and as not entailing connecting laws. (1993:4, n. 4)

Second, Kim demonstrates that under assumptions of closure of S, strong and weak supervenience are equivalent, respectively, to strong and weak entailment $P^* \rightarrow M$ principles, where $P^*$ is a maximal S-property. Unlike Kim, Davidson never explicates supervenience in terms of entailment conditionals $P^* \rightarrow M$. Third, Kim also introduces a notion of global supervenience, which, arguably, fits better with the thesis of externalism; Davidson, we recall, is an outspoken proponent of externalism. But it turns out that global supervenience is an intricate notion that comes with very different versions. Instead of invoking global supervenience, it is sufficient to use the individual notions but not limit S to monadic, micro, local, or intrinsic properties; it could include causal relations with the physical environment, and even bits of causal history. S could even include physical properties of remote objects if such properties are indeed relevant to the ascription of mental properties.

Davidson's statement about his notion of supervenience being weak is puzzling. The problem with weak supervenience is that it does not support dependence. Weak supervenience is consistent with the scenario in which my counterpart and I have exactly the same physical properties, but different mental properties. But then the mental difference is not due to our physical properties, since
nothing related to our physical makeup, including past and present causal relations with their environment, differs. It would thus seem that there are mental properties that do not depend on physical properties.

Davidson himself rules out such scenarios. He maintains that counterfactual scenarios like the Twin-Earth and Swampman thought-experiments, the mental differences are accompanied by physical differences:

> What I take Burge’s and Putnam’s imagined cases to show (and what I think the Swampman example shows more directly) is that people who are in all relevant physical respects similar (or ‘identical’ in the necktie sense) can differ in what they mean or think… But of course there is something different about them, even in the physical world; their causal histories are different. (1987:32-33)

Weak supervenience, however, lacks the modal force to support these psychophysical dependencies. Why, then, does Davidson invoke weak supervenience? One answer might be that mental properties do not strongly supervene on the intrinsic physical properties of agents, as the Twin-Earth and Swampman examples show. The mental only weakly supervenes on intrinsic physical properties: two agents of the same “world” that are physically indiscernible are also mentally indiscernible (see Davidson 1995).

However, this proposal will not do. It is true that the mental does not strongly supervenes on intrinsic physical properties, but it does strongly supervenes on intrinsic and extrinsic physical properties. It thus makes more sense to use strong supervenience over intrinsic and extrinsic physical properties, which reflects dependence, rather than using weak supervenience over intrinsic physical properties alone, which does not reflect dependence. One way or another, it seems that the weakness of Davidson’s supervenience does not stem from the usual "across-worlds" relations.
3. Supervenience as a philosophical concept

Davidson does not say much about the philosophical import of supervenience, but it is clear that his views on this are unique. One respect in which they are unusual has to do with the philosophical role of supervenience. Supervenience is widely upheld by non-reductive monists: those who maintain that every mental event is a physical event, but deny psychophysical laws. Many who espouse versions of this view worry that it is insufficiently "materialistic". The concern is that non-reductive monism says too little about the relations between mental and physical properties. Although it denies that mental properties are physical properties, it imposes no alternative constraints on the attribution of mental properties. Monism ensures that every object with mental properties also has physical properties, but, beyond that, anything goes: monism is consistent with the possibility that my physical counterpart has no mentality whatsoever, while my cup of coffee does. Surely a monistic, to say nothing of materialistic, doctrine that allows such wild attributions is worthless. Something must be done to close this gap. And this is where supervenience kicks in. The role of supervenience is to put more significant constraints on the psychophysical relations between mental and physical properties, without reducing mental properties to physical properties. Supervenience, on this account, is a secondary thesis that makes non-reductive monism materialistically kosher.

"Mental Events" gives the impression that supervenience plays this legitimizing role in Davidson's philosophy. After presenting the tenets of anomalous monism, Davidson immediately introduces supervenience, saying that "although the position I describe denies there are psychophysical laws, it is consistent with the view that mental characteristics are in some sense dependent, or supervenient, on physical
characteristics" (p. 214). From this we might conclude that Davidson, too, feels obliged to complement his monistic thesis about events with a substantive and positive thesis about the psychophysical relations between predicates. But later on, in his "Reply to Harry Lewis" (1985) and in "Thinking Causes" (1993), it turns out that this is not how Davidson sees the role of supervenience. Declaring that "supervenience in any form implies monism; but it does not imply either definitional or nomological reduction", Davidson reveals that he invoked supervenience to demonstrate that anomalous monism is consistent: "So if (non-reductive) supervenience is consistent (as the syntax-semantics example proves it is) so is AM [anomalous monism]" (1993:5).

Contrary to first impressions, then, supervenience is not a secondary thesis intended to correct the deficiencies of the primary doctrine of non-reductive monism. There is no evidence that Davidson deems his anomalous monism to be in need of reinforcement, whereas we do have evidence that Davidson does not take supervenience to provide such reinforcement. Davidson, of course, does resist the idea that mental properties float freely, as it were, over the physical domain, and does take supervenience as asserting that the mental depends on the physical realm. But this claim about dependency is not made as a substantive addition to anomalous monism. Rather, supervenience is used both to help establish monism and the consistency of anomalous monism. Davidson deploys supervenience once again in "Thinking Causes", this time to secure the causal relevance of mental properties. The claim made is that supervenience entails that an event's mental properties make a difference to its causal relations.

Davidson’s supervenience is also unique with respect to the notion of dependence. Most philosophers, following Kim, maintain that mind-body
supervenience is grounded in some deeper *metaphysical* relation. The idea is that any $P^* \rightarrow M$ conditional reflects the dependence of $M$ on $P^*$, and this dependence is a metaphysical determination relation, e.g., identity, constitution, emergence, or realization, which underlies and explains the supervenience relations. It is thus not surprising that, in the context of supervenience, the notions of dependence and determination are often used interchangeably. The implicit assumption is that $M$ depends on $P^*$ by virtue of $M$'s being determined by $P^*$, whereas determination is understood as a metaphysical determination.

Davidson's notion of dependence is different. The idea that the application of a mental predicate is grounded in some metaphysical determination of the mental by a fixed physical basis is foreign to Davidson's approach. He never hints that the mental depends on the physical by virtue of some metaphysical determination relation; and certainly does not introduce the more familiar determination relations to substantiate his supervenience thesis. In fact, in "Thinking Causes", the main argument for the causal relevance of mental properties suggests that supervenience is *not* such a determination relation. He asserts: "supervenience as I have defined it does, as we have seen, imply that if two events differ in their psychological properties, they differ in their causal properties (which we assume to be causally efficacious). If supervenience holds, psychological properties make a difference to the causal relations of an event, for they matter to the physical properties, and the physical properties matter to causal relations" (1993:14). But it is apparent that "make a difference" cannot be understood to mean "determine" in a metaphysical sense. For it refers to the mental-to-physical direction, whereas the pertinent metaphysical relation is in the physical-to-mental direction. It is most unlikely that Davidson would take supervenience to point to metaphysical determination of the mental by the physical,
and still claim that supervenience implies that mental properties "matter to the physical properties".

We see that Davidson does not uphold the idea that supervenience reflects metaphysical physical-to-mental determination or dependence relation. It seems, moreover, that he also rejects the idea that dependence (of mental on the physical) grounds or accounts for supervenience. If anything, it is the other way around. Davidson says that "supervenience gives a sense to the notion of dependence here, enough sense anyway to show that mental properties make a causal difference" (1993:14). So it is not that dependence accounts for supervenience, but, if anything, dependence is explicated in terms of the supervenience of the mental on the physical.

Lastly, it is telling that Davidson invokes supervenience in causal contexts. In discussing the Twin-Earth and Swampman cases, Davidson insists that "of course there is something different about them, even in the physical world; their causal histories are different". He later describes supervenience as implying that "mental properties make a causal difference". And he links supervenience with the causal nature of the mental, stating that "Kim, as we noted, thinks my version of supervenience implies that all mental properties could be withdrawn from the world and this would make no difference to causal relations; but this supposition turned out to be incompatible with my understanding of supervenience" (1993:14); and that "[s]upervenience as I defined it is consistent with… the assumption that there are no psychophysical laws… It is not even slightly plausible that there are no important general causal connections between mental and physical properties of events. I have always held that there are such connections" (1993:14).
4. Summary

Let us sum up the distinctive features of Davidson’s supervenience. Supervenience is not a secondary thesis the objective of which is to reinforce anomalous monism. It is not explicated by some deeper metaphysical determination or dependence relation, but if anything, it is supervenience that gives cogency to the notion of dependence. And it has something to do with the "causal connections between mental and physical properties of events". In addition, Davidson characterizes supervenience in terms of indiscernibility or covariance and not in terms of the entailment $P^* \to M$ conditionals, and declares that his supervenience is of the weak kind. Whether we can we extract from these remarks a cohesive notion of supervenience, and whether this notion can be reconciled with anomalism is something I will discuss elsewhere.

Oron Shagrir
Department of Philosophy
Program for Cognitive Science
The Hebrew University of Jerusalem

References:


