Supervenience and Anomalism are Compatible

Oron SHAGRIR[†]

ABSTRACT

I explore a Davidsonian proposal for the reconciliation of two theses. One is the supervenience of the mental on the physical, the other is the anomalism of the mental. The gist of the proposal is that supervenience and anomalism are theses about interpretation. Starting with supervenience, the claim is that it should not be understood in terms of deeper metaphysical relations, but as a constraint on the relations between the applications of physical and mental predicates. Regarding anomalism, the claim is that laws have to satisfy certain counterfactual cases, in which an interpreter evaluates her past attributions in the light of new pieces of evidence. The proposed reconciliation is that supervenience entails that an interpreter will always attribute the same mental predicates to two individuals with the same physical states. However, supervenience does not imply that the interpreter cannot revise her past attributions to the two individuals.

1. Introduction

In "Mental Events", Donald Davidson famously argues that the mental is anomalous, namely, that "there are no strict deterministic laws on the basis of which mental events can be predicted and explained" (1970, 208); in particular, he argues, there are no strict psychophysical laws. Davidson also contends that the denial of psychophysical laws is consistent with a supervenience thesis, namely, that "mental characteristics are in some sense dependent, or supervenient, on physical characteristics" (1970, 214).

[†] Departments of Philosophy and Cognitive Science, The Hebrew University of Jerusalem. E-mail: oron.shagrir@gmail.com

Jaegwon Kim challenges this compatibility claim. He argues that no notion of supervenience can secure dependence and anomalism at the same time. Weak supervenience, according to Kim, does not establish dependence, whereas strong supervenience, of the mental on the physical, entails strict psychophysical laws. These laws are of the P*→M form, i.e., *"Necessarily, for any x, if x has P* then x has M*", where P* is a maximal physical property and M is a mental property.

Those who responded to Kim tend to accept strong supervenience. They agree with him that strong supervenience leads to strict P*→M conditionals that are necessary truths (at least in being counterfactual supportive). But they deny that these are laws; they argue that the maximal P* do not fit, for one reason or another, with the standard of laws. My aim here is to explore a different proposal for the reconciliation of supervenience and anomalism. The gist of the proposal is that supervenience and anomalism are theses about interpretation, i.e., the process by which we attribute mental *predicates*.

Starting with supervenience (Section 2), I argue that Davidson's notion of supervenience must be very different from the one assumed in contemporary philosophy of mind. Unlike Kim, Davidson does not assume that supervenience points to a deeper metaphysical relation between properties, one that grounds and accounts for the supervenience relations. Instead, Davidson's supervenience should be understood as an indiscernibility constraint on the relation between the applications of mental and physical predicates.

Next, I propose some reconstruction of the anomalism thesis, according to which psychophysical (and psychological) statements do not satisfy certain counterfactual cases (Section 3). I focus on the cases (called here the M-variance cases) in which we hold fixed the application of physical predicates to someone, and

reconsider our mental attributions to him in the same situation but in light of new evidence. The claim is that there are cases in which the interpreter revises her attribution of mental predicates. The source of the claim is the thesis that the interpretation process is *open-ended* in the sense that at no point does the evidence remove all alternative ascriptions.

The last section is about compatibility (Section 4). The proposal, in a nutshell, is that supervenience implies that whenever an interpreter attributes to two individuals the same physical state, P, she attributes to them the same mental state. But supervenience does not entail that the interpreter cannot, at a later point, revise her past attributions, making her decide that it would have been better to attribute a different mental state to individuals with P. Davidsonian supervenience, as I will put it, requires the satisfaction of some counterfactual cases (called here the M-invariance cases) but does not require the satisfaction of M-variance cases. Though the proposal might be attractive to the Davidsonian, it has its limitations, which I discuss in closing.

My aim here is not to advance a proposal that is exactly Davidson's in every detail. The proposal is Davidsonian in two senses. First, I follow some of the claims and arguments that Davidson advances with respect to supervenience, laws, and anomalism (even though we have to take into account that Davidson says different things about supervenience and anomalism in different places). Second, I locate the proposal within Davidson's overall philosophical outlook. In these senses, the proposal can be comfortably characterized as Davidsonian.

2. *Supervenience*

Davidson discusses supervenience in several places, but in none of them does he provide a comprehensive account. I discuss these passages at some length elsewhere (Shagrir, 2010). Here I highlight some of the distinctive features of Davidson's supervenience (Section 2.1) and then reformulate a Davidsonian conception of supervenience (Section 2.2).

2.1 Davidson on supervenience. Davidson's supervenience has several unique features. Let me start with the characterization of supervenience. Davidson conceives of supervenience as a thesis about the relations between properties or characteristics or respects – e.g., mental and physical properties – which he understands as *predicates* (following Davidson's convention (1993) I sometimes use the term "properties" on the understanding that we are talking about predicates). Davidson characterizes these relations in terms of *indiscernibility* ("there cannot be two events alike in all physical respects but differing in some mental respect" (1970, 214)) or *covariance* ("an object cannot alter in some mental respects without altering in some physical respects" (1970, 214).[1] Unlike Kim, Davidson never explicates the supervenience relations in terms of entailment conditionals P*→M; the P*s in the conditionals refer to maximal properties as "the strongest consistent properties constructible" (Kim 1984, 58) in the set of (physical) properties by Boolean operations and negation.[2]

---

[1] See also Davidson (1985, 242; 1993, 4). Elsewhere I point out that, on a charitable reading, these characterizations are equivalent (Shagrir, 2009).

[2] Whether Davidson thinks that supervenience implies the P*→M conditionals is an interesting question I discuss elsewhere (Shagrir, 2010).

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 properties 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 ↔ Py (i.e., x and y are S-indiscernible), then Mx ↔ My (i.e., x and y are R-indiscernible).

Given Davidson's externalism, the set S need not be limited 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 the physical properties of remote objects if such properties were indeed relevant to the ascription of mental properties. [3] Thus when talking about a maximal S-property we should understand it in the "global" sense of a complete "world-perspective" in terms of S-properties.[4]

A second feature has to do with the modal strength of supervenience. Kim (1984) famously distinguishes between strong and weak supervenience. On the strong reading the supervenience condition applies to any pair of possible objects, whereas on the weak reading it applies to any pair of objects, both of which inhabit the same world. 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)

It is hard to see how weak supervenience allows anomalism and yet secures dependence at the same time; it is not clear why Davidson thinks that weak supervenience can even support the kind of psychophysical dependencies that he

---

[3] See Davidson (1990a) for a list of the relevant factors in the attribution of mental states.

[4] For a precise characterization see Stalnaker (1996) and Sider (1999); another way of defining these "global" maximal properties is via closure under quantification. Sider (1999) proves that when taking P* to be maximal in this global sense, the strong entailment condition is equivalent to strong global supervenience.

wishes to maintain. Davidson holds that statements about psychophysical relations have some modal force in the sense that they are counterfactual supportive. When discussing the Twin-Earth and Swampman thought experiments, he insists that 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. (1987a, 32-33)

But then it is puzzling how weak supervenience fits with this statement. If these counterfactual cases belong to "different worlds," then weak supervenience lacks the modal force to support the pertinent psychophysical dependencies (i.e., that the mental difference *is* accompanied by a physical difference). And if these counterfactual cases belong to the same world, then weak supervenience entails that P*→M conditionals are counterfactual supportive; it is then hard to see how weak supervenience can block the move to connecting laws. Why, then, does Davidson promote the weaker version of supervenience? I return to this issue in Section 4.

A third feature concerns 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 also deny psychophysical laws. Many who advocate versions of this view are concerned that it is insufficiently "materialistic." The worry is that non-reductive monism says too little about the relations between mental and physical properties. The role of supervenience in this context is to put significant constraints on the psychophysical relations between mental and physical properties without reducing mental properties to physical properties.

One might get the impression that supervenience plays the same legitimizing role in Davidson's anomalous monism.[5] But 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). Davidson, of course, does resist the idea that mental properties float freely, as it were, over the physical domain, and he 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 to help establish both monism *and* the consistency of anomalous monism.

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. This understanding underlies the conviction that psychophysical supervenience must be of the strong kind. The idea is that a strong P*→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 of the mental on the physical. 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

[5] This is Kim's impression: "Clearly, mind-body supervenience is a net addition to anomalous monism. By adopting it, Davidson has substantially strengthened his position on the mind-body problem" (2003, 130).

being determined by P*, whereas determination is understood as metaphysical determination.[6]

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 he certainly does not introduce the more familiar determination relations to substantiate his supervenience thesis.[7] In fact, in "Thinking Causes," Davidson seems to say that it is the other way around. It is supervenience that grounds or accounts for the dependence of the mental on the physical: "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; if anything, dependence is explicated in terms of the supervenience of the mental on the physical.[8]

Lastly, Davidson invokes supervenience in *causal* contexts. In discussing the Twin-Earth and Swampman cases, Davidson insists that "of course there is *something*

---

[6] See Kim who writes that "mind-body supervenience… points to the existence of a dependency relation" (1998, 10), which grounds or accounts for it.

[7] Davidson says in "The Material Mind" that "there is a sense in which the physical characteristics of an event (or object or state) *determine* the psychological characteristics; in G.E. Moore's view, psychological concepts are *supervenient* on physical concepts". But he then makes clear that determinacy here is nothing more than indiscernibility: "It is impossible for two events (objects, states) to agree in all their physical characteristics… and to differ in their psychological characteristics" (1973a, 253).

[8] See also his main argument for the causal relevance of mental properties, where Davidson asserts that "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 "make a difference" here refers to the mental-to-physical direction, whereas the pertinent metaphysical relation is in the physical-to-mental direction.

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 "supervenience as I define it is consistent with… the assumption that there are no psycho-physical 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).

Let me sum up the distinctive features of Davidson's supervenience. Supervenience is characterized in terms of indiscernibility or covariance and not in terms of the entailment P*→M conditionals; in addition, at some point Davidson declares that this condition has the modal force of weak supervenience. Supervenience is a philosophical thesis whose objective is not to reinforce anomalous monism but rather to demonstrate its consistency. It is not explicated by some deeper metaphysical determination or dependence relation between properties; if anything, it is supervenience that gives cogency to the notion of dependence. And, lastly, supervenience has something to do with the "causal connections between mental and physical properties of events."

2.2. Redefining supervenience. What can we make out of all these features? Do they combine into a cogent notion of supervenience? Admittedly, I am not sure there is such a notion. But I think it is safe to say that Davidson's supervenience is not a

metaphysical thesis about some deeper metaphysical dependence relation (though, of course, the thesis does not rule out such a dependence relation).[9] Supervenience, I suggest, is a thesis about the relations between physical properties and the procedures of interpretation: the procedures by which an interpreter attributes mental predicates to others on the basis of what they say and do.[10] The thesis is that the relations can be characterized by an indiscernibility condition, namely, that an attribution of mental difference is always accompanied by some physical difference. Put differently, supervenience is the statement that an interpreter ascribes the same mental properties to objects or events unless there is some physical difference about them, i.e., in their physical makeup, environment, or causal histories. If there is no such physical difference, there is no mental difference.[11]

But what exactly is the source of the supervenience relations? What is it that grounds these relations, if not metaphysical determination? One possible answer is that supervenience is a thesis about the linkage between the rules and principles of the application procedures of mental and physical predicates. The application of both mental and physical properties of events relates to causal connections between these events. These causal connections can be characterized in two divergent ways. On the nomological conception, where there is a causal relation, there is a law.[12] In particular, causal relations are underwritten by strict physical laws that impel us to ascribe exactly the same physical predicates to two objects that have exactly the same causal

---

[9] I refer here to deeper metaphysical relation between *properties*; as we saw above, Davidson states that supervenience implies a deeper metaphysical relation between events, i.e., of identity.

[10] Davidson (1973b; 1974; 1990b); see Child (1994, chapter 1) for a comprehensive analysis of the relations between interpretation and mentality; see also LePore and Ludwig (2005, part II and chapter 22).

[11] See Campbell (2000) who argues that supervenience is related to interpretation, and that it is of the weak variety.

[12] See Davidson (1967; 1995b).

interactions with the physical environment. The other characterization is in terms of the crucial role such causal connections play in fixing the content of thoughts, beliefs, and other attitudes. The public and intersubjective character of the mental prompts the interpreter to ascribe the same mental properties to two individuals that have exactly the same causal interactions with their environments. Supervenience can be seen as linking the two characterizations of events. It is the claim that the physical (type) identity that is dictated by laws is accompanied by a mental (type) identity that is constrained by intersubjectivity.[13] A mental difference that is rooted in some difference in the causal interaction of individuals with their environment must be accompanied by a physical difference that is rooted in the (strict) physical laws that cover these causal relations.

What exactly is the modal force of this no-mental-difference-without-a-physical-difference condition? The suggestion is that supervenience is strong in one sense and weak in another. It is strong in preserving psychophysical relations across worlds – "worlds" here are no more than counterfactual situations, as in the Twin-Earth and Swampman examples. In these examples, we saw, Davidson insists that the mental difference among counterparts is accompanied by a physical difference. Supervenience is weak in the sense that it does not preserve the psychophysical relations "across scenarios" but only "within scenarios." We can think of these scenarios as points of view from which an interpreter assigns extensions to predicates. They are triplets <W,S,T>, where W is a "possible" context, S is an interpreter, and T is time of attribution. These scenarios differ *not* in the meaning ("secondary

[13] It should be noted, however, that supervenience is *not* the claim that mental states are public, or accessible to the interpreter, viz., that an interpreter can tell, under favorable conditions, what someone thinks and believes (Shagrir, 2009).

intensions") we assign to predicates but rather in the amount of *evidence* available to an interpreter at this point (more on this in section 3.2).

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 properties 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, for every scenario C_i, for every pair of worlds W_j and W_k, viewed from this scenario, and for every pair of objects x in W_j and y in W_k, the following condition holds: if, for every P of S, $Px \leftrightarrow Py$, then $Mx \leftrightarrow My$.

I return to this supervenience condition below (Section 4). But let me elucidate the condition with an example. It's 10:00 o'clock on Sunday morning. Mary is trying to figure out what Frank wants to do. Two games, one baseball, the other soccer, are being broadcast on TV. Mary notices that Frank is watching the soccer game; she knows his tendency to watch soccer; he even told her that he intends to watch this game. On the basis of the evidence she has, Mary attributes to Frank the mental state M, i.e., a preference to watch soccer over baseball. Supervenience is just the claim that Mary must attribute the same mental state, i.e., a preference to watch soccer over baseball, to a possible individual, Dave, if Frank and Dave are physically indiscernible.

In another scenario Mary, who is no sports fan, departs to her lab. Returning home at noon, she finds Frank watching the baseball game. On being asked, Frank replies that he deliberately tricked Mary and had intended to watch the baseball game. In light of this new piece of evidence, Mary has good reason to revise her previous assessment of Frank's state of mind. She *now* says that it is better to attribute to Frank not the mental state M but, rather, the mental state M', i.e., the preference to watch baseball over soccer. In other words, Mary now says that Frank's state of mind, at 10:00, was not M, but M'. Supervenience compels Mary to attribute to Dave the same

mental state, M', assuming that Dave and Frank are (still) physically indiscernible. Supervenience does *not* entail, however, that the change in mental attribution (from M to M'), across scenarios, is accompanied by a change in physical attribution (e.g., from P to P'). It might be the case that Mary has no reason (at noon) to doubt the past attribution of P, i.e. that Frank had P at 10:00. What she says (at noon) is that although Frank (and Dave) was in a physical state P, he nonetheless had M' and not M.

To summarize, the supervenience condition is strong "within scenarios". From the perspective of Mary-on-Sunday-at-10, all possible individuals with P have M. And from the perspective of Mary-on-Sunday-at-noon all possible individuals with P have M'. The supervenience condition is weak "across scenarios" in the sense that there might be a mental difference, from M to M', that is not accompanied by a physical difference. I should emphasize, however, that the weak, interpretation-dependent, aspect of supervenience does not entail compatibility with anomalism. To achieve compatibility we also have to reconsider the thesis of anomalism.

3. Anomalism

My aim in this section is to advance an understanding of anomalism that emerges from several crucial sentences in "Mental Events" where Davidson advances an argument for psychophysical anomalism:

> No matter how we patch and fit the non-mental conditions, we always find the need for an additional condition (provided he *notices*, *understands*, etc.) that is mental in character (1970, 217); Belief and desires issue in behavior only as modified and mediated by further beliefs and desires, attitudes and attendings, without limit (1970, 217); The point is rather that when we use the concepts of beliefs, desire, and the rest, we must stand prepared, as the evidence accumulates, to adjust our theory in the light of considerations of overall cogency: the constitutive ideal of rationality partly controls each phase in the evolution of what must be an evolving theory (1970, 222-223).

As I see it, these sentences encapsulate an argument that consists of three premises. One premise is the claim that the interpretation process is *open-ended*. Open-endedness is a sort of indeterminacy thesis. It means that the interpretation process always leaves room for different attributions of mental states. There is no moment, now or in the future, in which the interpreter can tell that all the other ascriptions are inconsistent with the evidence available at that moment. This, at any rate, is the way I understand the saying that interpretation is "an evolving theory" modified and mediated "without limit."

A second premise, implicit in the argument, is that a general statement is a law only if it satisfies certain sorts of counterfactual cases. Put in the psychophysical context, the requirement amounts to the satisfaction of two conditions: the M-invariance condition and the M-variance condition. Roughly, the M-invariance condition consists of the usual counterfactual cases such as Twin-Earth and Swampman. The M-variance condition includes cases ("scenarios") in which we reconsider our attributions in the same situation in light of new pieces of evidence.

The third premise is the non-satisfaction of the M-variance condition. This premise is arguably implied by open-endedness and is manifested by the claim that any non-trivial system of interpretation rules allows the adjustment of past attributions in the light of new pieces of evidence; that, as interpreters, "we must stand prepared, as the evidence accumulates, to adjust our theory in the light of considerations of overall cogency." The conclusion is anomalism, i.e., the claim that (synchronic) psychophysical statements are not laws.

Let me clarify what is *not* claimed here. First, I am not suggesting that the only way to understand these "Mental Events" passages is in terms of open-endedness. I am also not suggesting that these passages are the most indicative for the thesis of

anomalism. Davidson advances different arguments for anomalism in different places, and there are different understandings of these arguments.[14] Second, I am not providing detailed arguments for the premises and, hence, a full-fledged argument for anomalism. The overall concern here is not the validity of anomalism but its compatibility with supervenience. My aim is an understanding of anomalism that fits the Davidsonian philosophical outlook.

In what follows, I explicate each premise. In section 3.1, I discuss the idea of open-endedness and distinguish it from other indeterminacy theses; in section 3.2, I discuss the notion of lawlikeness, and in section 3.3, I describe the move from open-endedness to the non-satisfaction of the M-variance condition. Later on, in Section 4, I suggest that supervenience entails the satisfaction of the M-invariance condition and, yet, that supervenience does not entail the satisfaction of the M-variance condition; hence it is compatible with anomalism.

3.1 Open-endedness. Open-endedness is the thesis that there is no moment along the interpretation process in which the interpreter can rule out the possibility of alternative ascriptions. Open-endedness can be seen as comprising two claims. One is a within-theory indeterminacy claim, namely, that an interpretation scheme ("psychological theory") does not fix a single ascription of mental states. There are always alternative ascriptions that are inconsistent with the current ascriptions, but that are consistent with the available evidence. The other claim is that this indeterminacy is not a local phenomenon. The accumulation of evidence is an ongoing process that stretches into the future (think of Mary who returns home at

[14] See, e.g., Kim (1985), Child (1994) and Yalowitz (1998).

noon, finding Frank watching the baseball game).[15] Yet at no point, now or in the future, does the evidence remove every single alternative ascription.

When we talk about the interpretability of mental states, we mean interpretability by *us*, namely, a *human* interpreter.[16] Yet the claim of open-endedness applies not (only) to ill-informed interpreters but to ideal, fully informed, interpreters. In particular, open-endedness is consistent with the ideal that an interpreter should (and would) attribute the correct mental state M to someone, when the circumstances are favorable. The claim, rather, is that even in ideal circumstances, the interpreter cannot rule out that other ascriptions, i.e., not-M, are also consistent with the evidence she has (I return to the issue of ideal interpretation in the last section).

The idea of open-endedness can be explicated by contrasting the applications of mental and physical properties. According to Davidson, an ideal physical theory is a *closed* system of quantitative laws on the basis of which it is possible to predict and explain a physical event. That is, a physical theory completely (or up to chance) fixes the physical properties of an object, at time *t*, provided there is sufficient evidence, e.g., if we know enough about the physical state of the world at time *t*-1. In this respect, a physical theory is *not* open-ended. The attribution of physical properties *from within* a theory is determined and fixed by laws and sufficient evidence. With this system of laws, the (probability of the) occurrence of any physical event can be derived. This does not mean that we always make such derivations. The point is that a constitutive ideal of the physical domain is that something's physical properties *could* be derived, at least in principle, as long as enough evidence has been accumulated.

---

[15] Davidson writes that "it is always possible, of course, to improve one's understanding of another, by enlarging the data base, by adding another dose of sympathy or imagination, or by learning more about the things the subject knows about" (1994, 232).

[16] See Davidson (1974); for discussion see Child (1994, 24-26).

When it comes to the ascription of mental predicates, the situation is different. No system of interpretation rules fixes the ascription of mental predicates. The rules of interpretation – any set of rules – always leave room for different attributions. Thus the scheme of interpretation is not only indeterminate relative to other schemes of application. Open-endedness, which is unique to the mental domain, refers to the indeterminacy arising *from within* an interpretation scheme. It is the assertion that at any point during the interpretation process there are several ascriptions consistent with someone's behavior, environment, and history. No matter how much evidence has been accumulated, the interpreter is never in a position to exclude all but one ascription.

One could complain that this contrast, between the mental and the physical, relies on an understanding of physical theories that is quite idealized. On this understanding, a physical theory is a closed system of laws that completely fixes the physical properties of an object. In real physics, however, it is seldom the case that *all* properties can be fixed by a physical theory (plus computing machinery). We find open-endedness in the ascription of physical properties too. Hence, open-endedness cannot be the basis of the anomalism of the mental; for, otherwise, the physical domain is anomalous too, and this conflicts with Davidson's own view that anomalism is a distinctive feature of the mental.

Admittedly, the assumption that a physical system is such a "comprehensive closed system" (Davidson 1970, 223) is questionable. My aim here, however, is not to defend this claim, but to highlight its centrality in the argument for anomalism. Davidson says that "there are laws drawing on concepts from the same conceptual domain and upon which there is no improving in point of precision and comprehensiveness. We urged in the last section that such laws occur in the physical

sciences" (1970, 223). These laws, Davidson says elsewhere, are closed "in the sense of requiring no ceteris paribus clauses; they were to come as close to allowing the unconditional prediction of the event to be explained as the perhaps irreducibility probabilistic character of physics allows" (1987b, 111). As Child puts it, "the occurrence (or the chance) of any event is precisely determined by the physical character of its antecedents, in accordance with laws which make up a general, quantitative system" (1994, 61). This "*system* of laws" (Child 1994, 61) is nothing like anything we find in the realm of the mental, which does not "constitute a closed system" (Davidson 1970, 224).[17] Again, I admit that this picture of physics is contentious, but my overall aim here is the substantiation of a Davidsonian understanding of anomalism.

Another objection is that the open-endedness of the mental is inconsistent with the closure of the physical domain. The objection is this: Let us assume that Mary is also a knowledgeable scientist. With the help of our ideal closed system of quantitative laws, she can derive *any* future physical state from Frank's physical state, P*, at least up to chance. She could derive any future piece of evidence and, on the basis of it, rule out any ~M ascription to Frank.

But the closure of the physical domain alone does not allow such derivations. First, Mary cannot compute *all* future states at once. Even were she able, at 10 a.m., to compute the state of the physical world at, say, noon, there would always be a further movement, state of the environment, etc., that Mary hasn't yet computed. From her point of view, there can always be another piece of evidence that is relevant to her

[17] For further discussion see Child (1994, 61); see also Davidson (1991, 215-217; 1997, 71).

past attributions.[18] Second, even were she able to compute the state of the physical world at noon, it does not follow, without further argument, that Mary could deduce what would be the best attribution of *mental* states at that time. For attribution of mental states, Mary will look at Frank's circumstances, speech, and behavior, and I certainly do not assume that this kind of evidence – e.g., Frank *watching* the baseball game – is entirely physical (non-semantic).

Another way to explicate open-endedness is by distinguishing it from other grades of indeterminacy. One grade of indeterminacy emerges from the benign fact that we cannot always accurately attribute mental states to another. There are borderline cases, indefiniteness, and vagueness.[19] In addition, we often make mistakes, miscalculate, mishandle a crucial piece of evidence, or misjudge the situation. We sometimes do not even follow the guides of interpretation rules (e.g., not applying the principle of charity). Open-endedness, however, is present even in cases in which our judgment is impeccable. Open-endedness is the claim that even in the impeccable cases the interpreter cannot rule out other ascriptions. She cannot rule out the possibility of future evidence indicating a better ascription.

A second case of indeterminacy is related to competing schemes, namely, empirically equivalent but different theories. A theory is indeterminate by past evidence and, arguably, by all *possible* evidence, even in the broadest sense.[20]

---

[18] To be more precise, the derivability of any future state P does not imply decidability, namely, of whether an arbitrary event will occur or not, e.g., the earth changing its orbit; for a relevant and very useful discussion of undecidability/ incompleteness results in the context of physical systems, see Pitowsky (1996).

[19] See Child (1994, 69-70).

[20] This kind of indeterminism is emphasized by Quine (1970, 178-179; 1987, 10); for a recent discussion see Lepore and Ludwig (2005, chapter 15, especially pp. 223-227). Quine, however, talks about *logically incompatible* theories and distinguishes "underdetermination of theories" that turns up in both mental and physical domains, and "indeterminacy of translation" that turns up only in the mental domain. Quine famously argues that in the undetermination case there is a fact of the matter about the

Davidson clearly upholds the likelihood of competing schemes of interpretation. He grants the possibility of different interpretative theories that apply equally well to all possible evidence; this is the situation that Davidson associates with indeterminacy of interpretation. But Davidson thinks that this kind of indeterminacy also turns up in the physical domain. His example is the different scales for measuring temperature – Centigrade, Fahrenheit, etc. – that can be seen as competing schemes that apply equally well to all possible evidence.[21] Davidson thus concludes that this kind of indeterminacy cannot supply a reason for anomalism.[22]

The issue of indeterminacy of interpretation is notoriously difficult and controversial. The important point, for our purposes, is that it differs from open-endedness. The indeterminacy of competing schemes is consistent with the derivation, from within theory, of the occurrence (or chance) of every event. This is arguably the case in physical theories: given enough evidence, we have a within-theory derivation (up to chance) of every physical event. Open-endedness is the denial of this within-theory derivation of (mental) events. Open-endedness is the assertion that no interpretation scheme fixes the mental properties once and for all.[23] It is the claim that the very same interpretation scheme ("theory") is always consistent with several different ascriptions of mental states.[24]

---

correctness of theories – e.g., one is correct and the other is not – whereas in the indeterminacy case there is none.

[21] See Davidson (1973a, 257; 1991, 214-215; 1997, 79). For further discussion and criticism see Child (1994, 70-72), Rawling (2003, 97-105), and Lepore and Ludwig (2005, 243-247).

[22] Davidson (1991:215).

[23] See Davidson, who states in a response to Dennett that "the same [psychological] system can be used to support different predictions" (1997:81).

[24] There is, of course, no fine line between competing attributions that belong to competing schemes and those belonging to the same scheme. By indeterminacy arising *from within* an interpretation scheme I mean that the attributions cannot be seen as merely belonging to different schemes.

Open-endedness is closely related to what William Child (1994) describes as the uncodifiability of rationality, namely, that "there is no fixed set of rules or principles from which, together with a statement of the circumstances of any particular case, we could deductively derive a complete, detailed specification of what one ought to do or think in that case" (1994, 58-59). Open-endedness might come down, under certain assumptions, to the same claim, but its focus is different. Child emphasizes the non-deducibility/derivability of mental properties; that there is no fixed set of strict principles "for deductively deriving a specification of S's *mental properties*" (1994, 60). The focus of open-endedness is on the claim that there are no means, deduction or others, by which an interpreter can ever exclude other alternative ascriptions as being inconsistent with the circumstances. No matter how much evidence is being accumulated, the interpreter will never be in a position to say that her ascriptions are the only ones consistent with the evidence.[25]

Another claim of indeterminacy has to do with an anti-realist view about the mental, a view on which there is no fact of the matter as to what someone thinks, believes, and prefers. Open-endedness does not entail this view. Open-endedness is not the claim that mentality is constituted by interpretation, or is interpretation-

[25] The difference in focus turns up in the understanding of anomalism. Child seems to think that the (ideal) interpreter can justify her attributions; what matters is that this justification is not a matter of derivation from a system of laws. I deny that such a justification can be provided (this failure is manifested in the claim that statements that involve mental predicates do not satisfy certain counterfactuals; see section 3.2). The difference also turns up in the reconciliation of supervenience and anomalism. Child (1994, 74-78) accepts P*→M supervenience laws, but denies that they constitute a fixed system on the basis of which we can *deduce* the mental properties of someone in new circumstances. My approach is very different (Section 4). Elsewhere I discuss and criticize at length Child's proposal of reconciliation (Shagrir, 2009). I argue that accepting the P*→M endangers both the argument for monism and the argument for anomalism (as reconstructed by Child).

dependent.[26] That Frank has the mental state M (say) and not M' might be determined by some objective, non-epistemic, interpretation-independent facts. Open-endedness is a claim about the interpretation process itself. The claim is that even when ascribing M to Frank, the interpreter cannot exclude all other ascriptions as being inconsistent with the evidence. Thus while open-endedness is compatible with anti-realism, it does not itself give us reason to reject realism.

Lastly, there is the claim that content (and meaning) is not fully determined by past and present circumstances, since some relevant determinant pieces lie in the future.[27] Open-endedness is consistent with this view, but is not a commitment to it. The claim of open-endedness is not that relevant determinant evidential pieces lie in the future. It might well be that all the relevant determinants lie in the past and present. The claim, rather, is that even in the latter cases, the interpreter, even an ideal one, cannot tell that there are no other ascriptions consistent with the evidence she has (and so, I shall claim, she cannot tell that no revision is forthcoming).

3.2 Laws. Let us return to Mary and Frank at 10:00 o'clock on Sunday morning. Assume that Mary, being a knowledgeable scientist, can tell Frank's physical condition and movements, his physical environment, bits of his physical past, and so on, perhaps even his maximal physical state P* (note that "maximal" here includes Frank's entire relations with the physical world). Besides being a knowledgeable scientist, Mary also has admirable interpretative skills. Operating under favorable

[26] See Child (1994, 47-55) for the distinction between the thesis that interpretation is constitutive of mentality and one that it is not.

[27] Jane Heal, for example, understands Davidson this way: "[O]n Davidson's story… contents… await determination in light of the later pieces. The patterns to which we must look in assigning meaning are the patterns spread out across time, including the future, and not merely across space. Thus features of later utterances, features which fix what overall patterns they and the earlier ones can form, are partly constitutive of the earlier utterances having the meaning they do" (1997, 193).

circumstances, Mary attributes to Frank the mental state M, i.e., a preference to watch soccer over baseball. But what would drive her to generalize and judge $\forall x(Px \rightarrow Mx)$ as a law, whereas P is some physical property, perhaps even Frank's maximal physical state P*?

My aim here is to advance a certain understanding of laws according to which P→M conditionals should support two sorts of counterfactual cases; I call them the M-invariance and the M-variance cases. One motive of this requirement is internal to the argument. It is the link between open-endedness claim and the thesis of anomalism. The idea is that *if* Davidson aims to establish anomalism through open-endedness, as suggested here, then it is reasonable to assume that he upholds this kind of requirement. Another motivation is the way Davidson treats laws, at least in some places. According to Davidson, "lawlike statements are general statements that support counterfactual and subjunctive claims, and are supported by their instances" (1970, 217). In particular, the P→M conditional is a law only if it is counterfactual supportive. Davidson also says that "to judge a statement lawlike or illegal is not to decide its truth outright; relative to the acceptance of a general statement on the basis of instances, ruling it lawlike must be a priori" (1970, 216). This does not mean that finding the correlation between P and M is a priori (it is obviously an empirical matter). It means that the leap from the truth of the P→M conditional in the current circumstances to its truth in counterfactual cases is, at least in part, a conceptual matter.[28] The suggestion, in other words, is that when judging statements as laws, we take into account the special features of our theories and the way we apply them

[28] See also pp. 216 ff. and the conclusion that "in all these respects nomologicality is much like analyticity, as one might expect since both are linked to meaning" (1970, 218).

against the evidence (I do not argue, however, that these passages necessitate the conception of laws suggested here).

My proposal is that the P→M conditional should support two sorts of counterfactuals. It would be useful to think of these counterfactual cases as lying along two dimensions, the M-invariance and the M-variance. We would say that P→M satisfies the M-invariance condition just in case it is true in all the M-invariance cases. We would say that P→M satisfies the M-variance condition just in case it is true in all the M-variance cases.

We can think of the M-invariance dimension as including the usual counterfactual situations such as the Twin-Earth and Swampman cases, in which Frank's doppelgangers have the same intrinsic physical properties as Frank but still do not have Frank's mental property M. It also includes counterfactual situations in which Frank's doppelganger Dave does have P. Along this dimension we hold the correlation between M and P constant in the actual world (in W_0), i.e., we hold that Frank is in the extension of P and of M (on Sunday morning at 10:00), and that P→M is true in Mary's world, W_0. What we query is whether the correlation between P and M also holds for all counterparts in these counterfactual situations: whether Dave who has P also has M, and that Frank's doppelganger on Twin-Earth who has ~M, also has ~P.

We can think of the M-variance dimension as including scenarios or perspectives from which we reevaluate mental ascriptions. Intuitively, they are tools for evaluating what mental properties an interpreter would have attributed to someone with P on the basis of *additional* evidence. They are used to test, for example, whether it is best to attribute M to Frank when he has P on Sunday morning at 10:00. In these scenarios we hold constant that Frank has P on Sunday morning at 10:00 and

examine, on the basis of additional evidence, whether it is proper to attribute M to him.

We described scenarios as triplets <W,S,T>, where W is a possible context, S is an interpreter and T is time of attribution. The M-variance cases are points of view from which an interpreter assigns extensions to predicates. They share three common features. One is that the context is a future trail of the actual world: It includes as its history a duplicate of the current physical state of the world, W_0, and *its* past history. A second feature is that the context is a *possible* trail. It is possible in the sense that an interpreter cannot exclude it as being inconsistent with the current evidence she has; in our example it is the evidence available to Mary on Sunday at 10:00. A third feature is that the interpreter (at time T) has access to all the present and past evidence; in our example, the interpreter (Mary or another) has at least the amount of *evidence* available to Mary on Sunday at 10:00.

More intuitively, Mary ascribes mental states to Frank on the basis of his behavior, his relations with the physical environment, and causal history. Even when she takes into account all this information, Mary is not aware of his future behavior, relations with the environment, and so forth. When made aware of these – say, by waiting long enough – she has at her disposal further pieces of evidence. In one scenario Mary returns home at noon to find Frank watching the baseball game. In another scenario she finds him watching the soccer game. What we query is what Mary should do, in light of this new piece of evidence: whether she should still say that it is better to attribute to individuals with P, like Frank at 10:00, the mental state M.

The main conceptual difference between the two dimensions is this: Along the M-invariance dimension we test whether P→M holds from the $<W_0$, Mary, Sunday-

at-10:00>-perspective that Frank has P and M on Sunday at 10:00 (in W_0). We examine, that is, whether the correlation between P and M holds in other, counterfactual situations, from the perspective that it holds in the actual world (in W_0). Along the M-variance dimension we do not view the counterfactual situations through the <W_0, Mary, Sunday-at-10:00>-prism, but, intuitively, we reevaluate Mary's M-ascription to Frank, at the time he had P, from other perspectives (which are duplicates of W_0 up to Sunday morning at 10:00). We examine from these perspectives Frank's situation on Sunday at 10:00 in W_0.[29]

The M-variance scenarios, however, are *not* contexts in which we change the schemes of applications and meaning.[30] The relevant schemes of applications used in all scenarios are the ones used by Mary on Sunday morning at 10:00, and such terms as 'baseball', 'soccer', 'water', and 'H_2O' have the same meaning they have on Mary's lips. Respectively, the contents of mental states that are associated with these terms are similar across scenarios. Scenarios differ from one another in the amount of evidence available to the interpreter when applying these terms.[31]

---

[29] It would be interesting to compare this "two-dimensional system" to the one proposed by David Chalmers (2004). To highlight just a few points: The M-invariance dimension is seemingly similar to Chalmers's metaphysical dimension with respect to the list of counterfactual cases. There are differences, however, with respect to the truth-values in some of the cases: Chalmers thinks that there are physical twins that have no mental states at all (zombies), whereas Davidson argues that there cannot be mental difference without some physical difference. The M-variance dimension is different from Chalmers's epistemic dimension. Most importantly, the M-variance cases are not scenarios in which we change the secondary intensions of terms (see next paragraph and next footnote). Thus while one can agree with Chalmers that satisfying the scenarios along *his* epistemic dimension ("the diagonal") is not a requirement for being a law, one can hold that satisfying scenarios along the M-variance dimension is a requirement for being a law.

[30] Davidson explicitly indicates that laws are tested relative to fixed schemes of application (and meaning); see, e.g., Davidson (1970, 222; 1991, 215).

[31] There is, of course, no fine line between theory and evidence, and, hence, a point at which a certain amount of additional evidence affects theory (and meaning). The criterion is that when the theory (or meaning) changes, we no longer have a scenario.

One could wonder about the difference between psychological and physical predicates: Why impose the variance condition with respect to psychological predicates, but not with respect to physical predicates? The reply is that we also impose variance conditions in the case of physical predicates. Physical statements (i.e., that include no mental predicates) differ from psychological statements in that they do satisfy the pertinent variance conditions. Consider, for example, the statement *water is H_2O* (or the statement *if something is H_2O then it is water*). This statement satisfies the variance condition only if it is true in all scenarios along the WATER-variance dimension. However, there are good reasons to think that *water is H_2O* does satisfy the variance conditions (*if* it is true at all). Mary can in principle, even if never in practice, deduce it from the (ideal) physical theory and the entire evidence up to this point. Assuming the closure of the physical domain, she can deduce that the liquid she calls 'water' is made up of molecules of H_2O. Thus there cannot be a scenario in which Mary revises her identification of water and H_2O. For this scenario must be inconsistent with the physical theory and evidence on the basis of which Mary can deduce that water *is* H_2O.

3.3 From open-endedness to anomalism. How do we arrive at psychophysical anomalism from open-endedness? Why, to be more specific, not count the P→M conditionals as laws? The last premise of the argument is that the M-variance condition is defeated by the open-ended nature of interpretation. The reason is this: As time passes, Mary is constantly exposed to more pieces of evidence. Some pieces of evidence might make her even more convinced that her attribution of M to Frank is the correct one. Other pieces might encourage her to revise her interpretation, deciding that the more fitting attribution is not M, but M'. Other future evidence might

bring Mary to change her mind yet again, deciding that the more fitting property is M'' (or M again), and so on. Thus Mary must always stand prepared, as the evidence accumulates, to adjust her past mental attributions. There will always be a scenario in which Mary revises her attribution to Frank.

Consider the scenario in which Mary returns home at noon to find Frank watching the baseball game. What should she do, in light of this new piece of evidence? Mary has no reason to doubt her past attribution of P, i.e., that Frank had P at 10:00. The attribution of P was fixed, via a comprehensive closed system of laws, by the complete physical state before 10:00. Nothing that happened after 10:00 gives her any reason to change this attribution. However, Mary has good reason to revise her previous assessment of Frank's state of mind. In light of the new evidence, she can *now* say that it is better to attribute to individuals with P the mental state M' rather than M. In particular, she can say that although Frank was in physical state P at 10:00, he nonetheless had M', i.e., he preferred, at 10 o'clock, to watch baseball over soccer.

I have so far explained why open-endedness defeats the M-variance condition. The *argument* for this claim is roughly this: Open-endedness asserts that at each moment, at 10:00 or after, Mary takes into account that Frank's having ~M is consistent with the rules of interpretation and the available evidence at the time. It thus follows that there is a scenario in which we attribute to Frank, at the time he had P, a different mental state, say M'; we judge that although Frank had P at 10:00, it is better to attribute to him M' and not M. If there were no such scenario, then Mary could tell (on Sunday morning) that there are no forthcoming circumstances under which she would revise her ascriptions from M to M'. This means that Mary could tell that there are no circumstances under which the best ascription would be M': She could tell that any such circumstances would be inconsistent with the application

schemes she uses and the evidence she has. But given the publicity constraint on mental ascriptions, this would show that M' is not an option at all. For if Mary knows that there are no favorable circumstances under which she will ascribe to Frank M', then M' is not publicly detectable and, hence, cannot possibly be ascribed to Frank. But then we lose open-endedness.

One could point out that this conclusion conflicts with the idea of radical interpretation, namely, that "what a fully informed interpreter could learn about what a speaker [believes] is all there is to learn" (Davidson 1983, 148). If Mary is an ideal interpreter (say, the circumstances are favorable at 10 o'clock), then the evidence she has at this point is enough ("all there is to learn") to tell that Frank has M rather than M'. No further pieces of evidence, now or in the future, can change this attribution. But this seems to conflict with the claim that Mary must stand prepared to revise her current attributions, and, in particular, that there is a scenario in which the (additional) evidence would drive Mary to say that Frank had M' rather than M at 10 o'clock.

The notion of an ideal ("radical") interpreter is notoriously difficult, perhaps even contradictory.[32] There are also different ways to understand this notion. On a weaker understanding, Mary, as a fully informed interpreter, should (and would) attribute the correct mental state, M, to Frank; no further evidence can change this attribution. Mary can even justify her M-attribution, over an M'-attribution, with respect to these circumstances. On a stronger understanding, Mary can tell, in addition, that she is "fully informed"; she can tell that the circumstances are favorable.

[32] See Lepore and Ludwig (2005, chapter 15) who claim that the idea of radical interpretation is suspect. They argue that there are facts that determine which attribution is correct, but that these facts are beyond the reach of the (ideal) interpreter. This is because "the evidence available to the radical interpreter, together with the constraints he can legitimately bring to bear on his task, genuinely underdetermine the theories he can confirm" (p. 222). While Lepore and Ludwig invoke their argument in the context of competing schemes, it seems to me that they equally apply to the context of competing attributions within a scheme.

She can, thus, justify her M-attribution to Frank, once and for all.[33] Now, the claim that Mary must always stand prepared to revise her current attributions is surely in tension with the stronger understanding of an ideal interpreter; for Mary *knows* that no revision is forthcoming. But the claim is consistent with the weaker understanding of an ideal interpreter; for this understanding does require such knowledge.[34]

To summarize the argument for anomalism, it starts with the premise that the interpretation process is open-ended. It proceeds with the (implicit) premise that the satisfaction of the M-variance condition by P→M conditionals is necessary for being a law. And it concludes with the premise that open-endedness entails that P→M conditionals do not satisfy the M-variance condition. We thus get psychophysical anomalism: the P→M conditionals are not laws.

4. Compatibility?

Can we reconcile the theses of supervenience and anomalism? Kim famously argues that if supervenience is a notion of dependence, then the answer is negative. If the supervenience of the mental on the physical is strong enough to secure dependence, then it implies strict psychophysical laws of the P*→M form (again, P* might be a "global" property including Frank's entire relations with the physical world). Kim does not suggest that these P*→M conditionals can be the basis of a useful and

---

[33] LePore and Ludwig (2005) apparently associate radical interpretation with the latter or even a stronger requirement, one according to which a fully informed interpreter could *justify* her ascriptions. The ascription is justified not only with respect to the available evidence (which is the weaker requirement), but with respect to all possible evidence (this is not to say that the interpreter knows all possible evidence; it is to say that she knows that no further piece of evidence, whatever it might be, will defeat the current ascription).

[34] On the weaker understanding, Mary might be unable to point to other ascriptions consistent with the evidence; some of these ascriptions might even be inaccessible to her.

productive scientific enterprise. His claim is that, metaphysically speaking, the P*→M conditionals are metaphysically necessary and threaten the thesis of anomalism.[35]

Davidsonians address the challenge in two different ways. They can deny that the P*→M conditionals have the form of laws. They can say that the maximal P* do not fit into lawlike statements; indeed, they might challenge Kim to give us just one single real-world psychophysical law.[36] I discuss these responses elsewhere and argue that they are in tension with central Davidsonian theses.[37] Another way to meet Kim's challenge is to deny that the pertinent P*→M conditionals have the modal force of laws, e.g., that they support the right kind of counterfactuals. Davidson seems to take this route when saying that his notion of supervenience is "weak". I will now examine the potential of this proposal in light of what has been said so far about supervenience and anomalism.

4.1. Compatibility. It was suggested here that supervenience is a thesis about the relations between the applications of mental and physical predicates. Supervenience is the claim that any difference in mental ascription is accompanied by a difference in physical ascription. Putting it in terms of evidence, the claim is that the evidence pertinent to mental ascription is related in some way to the evidence pertinent to physical ascription. Evidence for mental difference is always accompanied by evidence for physical difference. But does this statement entail that the P*→M conditional is a law? The answer seems to be negative when we take "worlds" to refer

[35] See Kim (1984; 1990). Elsewhere I discuss Kim's argument and the responses to it (Shagrir, 2009).
[36] For Davidsonian responses of this sort see Davidson (1993, 1995a), Child (1993, 1994) and Shea (2003); see also Putnam (1973), who appeals to considerations of complexity.
[37] Shagrir (2009); see also Kim (2003).

to the M-invariance cases, and "scenarios" to refer to the M-variance cases. The answer is negative simply because the supervenience condition is weak along the M-variance dimension.

Let us start with the M-invariance dimension. Supervenience entails that if, relative to a scenario, there is a correlation between a mental property, say M, and a physical property, say P*, in one world, then this correlation holds across worlds. Supervenience ensures that if relative to the <W_0, Mary, Sunday-at-10:00>-perspective, Frank has M and P* (in W_0), then P*→M holds along the M-invariance dimension. Supervenience, in other words, rules out a case in which Frank, who has P*, has M, whereas Dave, who has P*, has M', e.g., the preference to watch baseball over soccer. Supervenience ensures that when viewed from the same perspective, two possible objects that differ in their mental properties also differ in some physical properties. This feature of supervenience is in accord with Davidson's discussion of the thought experiments, where he insists that there is something different in the physical properties of Davidson and Swampman (who differ in their mental states), and something different in the physical properties of Oscar and his Twin-Earth doppelganger.

That P*→M holds along the M-invariance dimension is not due to a specific deeper metaphysical dependence relation of the mental on the physical. It simply reflects the relations between the principles for applying mental and physical predicates. These principles, when put together, dictate that from a given perspective, and based on the same evidence, we cannot simultaneously ascribe to someone P* and M, and to someone else P* and ~M. The strong aspect of supervenience, which reflects these practices, entails that the correlations between M and P be preserved along the M-invariance dimension. Namely, that if Frank is having P* and M, then

any other possible object that has P* must have M. It is in this sense that M depends on P*.

How about the M-variance dimension? Supervenience, as defined here, does *not* entail that the correlations between M and P* must be preserved across scenarios. From a different perspective, upon having additional evidence, we might decide that it is better to attribute to Frank, at the time he had P*, the mental state M' and not M. We might even decide that Frank has no mental states at all. Supervenience is consistent with the P*→M conditional having different truth-values in different M-variance scenarios. It is consistent with P*→M not satisfying the M-variance condition. Thus supervenience is consistent with anomalism.

4.2 Objections. There are formidable objections to the proposed reconciliation. One objection is that the weak aspect of supervenience, along the M-variance dimension, is inconsistent with its being a dependence relation. If the mental is dependent on the physical, then the mental difference between Frank's having M in the $<W_0$, Mary, Sunday-at-10:00>-perspective and his having ~M in the Mary-returning-home-at-noon perspective must be accompanied by some physical difference. However, Frank has P* in both cases. It thus seems that the proposed notion of supervenience is no notion of dependence: the mental supervenes on the physical, even though it does not depend on it.

In response, I can only observe that there is a difference between the two cases, even in the physical world. Changing the attribution to objects with P*, from M to M', is always accompanied by a change in the physical world: the (centered) worlds, from which Mary ascribes to Frank mental states, are different. One physical world is the actual world on Sunday at 10:00, but the other is a counterfactual

situation in which upon returning home at noon, Mary finds Frank is watching baseball. Thus the mental difference is correlated with a physical difference: it is correlated with the difference between the physical worlds at the time the ascriptions are being made.

One could argue, then, that the notion of supervenience outlined here is equivalent to strong supervenience of the set of mental properties on the set of "super-maximal" physical properties. A super-maximal physical property P^+ includes the entire physical history, past, present and *(possible) future trail.* One could even argue that when Mary attributes M to Frank, she associates M not with P*, but with P^+. She might take into account that the physical future will be different than P^+, say $P^=$, i.e., that when she returns home at noon Frank will be in a physical state that correlates with M' (and this shows that M' is strongly correlated with $P^=$). But if we have strong supervenience, then we have laws. On the basis of P^+, Mary could rule out, on Sunday at 10, any ~M-ascriptions (and scenarios). She could tell that the psychophysical $P^+ \rightarrow$M conditional is a law.

My aim here is not to rule out any form of strong supervenience and laws. It might well be that M strongly supervenes, in some sense, on P^+, where P^+ is a complete, perhaps infinite, physical trail. It might also be that, in some sense, the $P^+ \rightarrow$M conditionals are laws. But we do not have strong supervenience and laws in the sense that we characterized these notions here. Let us start with supervenience, which was described earlier as a relation between the physical and mental predicates attributed to individuals. Can we really apply something like P^+ to Frank? Mary, being a knowledgeable scientist, attributes P* to Frank on Sunday at 10:00 (which is already an idealized move). She might also be able to derive, and so attribute to

Frank, *any* future physical state (which is even more idealized). But, as argued above, this does not entail that Mary can compute Frank's entire future trail P^+.

Furthermore, even if M supervenes strongly on P^+, it does not immediately follow that we have a $P^+ \rightarrow M$ law. Knowing Frank's P^+ suffices to rule out any ~M-scenario, but only if we assume that P^+ somehow includes all the relevant evidence for the attribution of Frank's mental states. But this assumption is no truism. Mary might be able to calculate Frank's physical state at noon without being able to tell what Frank says and does at that moment. Saying and acting are not assumed here to be physical properties. It might well be that in order to tell what Frank says and does, Mary has to confront Frank (or, perhaps, be told what he says and does).[38]

Lastly, one could object to the "epistemic" framework of the proposal. This framework is not likely to please the metaphysical realist who interprets supervenience as (reflecting) a metaphysical dependence relation, and who sees laws as objective features of the world. In replying, it should be remembered that the proposal does not appeal to the "usual" epistemic features of explanation and prediction (complexity, generality, simplicity, etc.). The proposal, rather, construes anomalism and supervenience as emerging, at least in part, from the constitutive principles of the interpretation process. This construal can hardly satisfy the metaphysical realists, but I hope I have shown that it is faithful to Davidson's philosophical outlook. And it was Davidson, after all, who introduced the theses to contemporary philosophy of mind.

5. Summary

[38] But see Heal (1997:176), who points out that radical interpretation "starts from information, *all* of which is non-semantic"; see also Davidson 1973b:128.

I have explored a proposal to reconcile the thesis of supervenience with the thesis of anomalism. The gist of the proposal is that both theses should be understood in the context of interpretation. Supervenience is a thesis about the relations between the application principles of mental and physical predicates: it is the statement that every difference in mental attribution is accompanied by a difference in physical attribution. It is not the statement, however, that a specific correlation between mental and physical predicates is not revisable. Anomalism is the claim that even when doing her best, the interpreter must always stand prepared to revise her past attributions in the light of good reason, i.e., new pieces of evidence. Thus anomalism amounts to saying that any specific correlation between mental and physical predicates is defeasible. And supervenience is compatible with this assertion.[*]

REFERENCES

Where reprinted versions of original papers are mentioned, page numberings refer to the former.

CAMPBELL, N. 2000, "Supervenience and Psycho-Physical Dependence", Dialogue: Canadian Philosophical Review 39, pp. 303-315.

CHALMERS, D. 2004, "Epistemic Two-Dimensional Semantics", Philosophical Studies 18, pp. 153-226.

CHILD, W. 1993, "Anomalism, Uncodifiability, and Psychophysical Relations", Philosophical Review 102, pp. 215-245.

---. 1994, *Causality, Interpretation and the Mind*, New York: Oxford University Press.

DAVIDSON, D. 1967, "Causal Relations", Journal of Philosophy 64, pp. 691-703. Reprinted in Davidson 1980, pp. 207-227.

---. 1970, "Mental Events", in L. Foster and J.W. Swanson, eds, *Experience and Theory*, London: Gerald Duckworth & Company Ltd., pp. 79-101. Reprinted in Davidson 1980, pp. 207-227.

---. 1973a, "The Material Mind", in P. Suppes, L. Henkin, A. Joja and G. C. Moisil eds., *Proceedings of the Fourth International Congress for Logic,*

[*] *Acknowledgment:* Thanks to Carl Posy, Ted Sider, Barry Smith, and three anonymous referees of *dialectica*. This research was supported by The Israel Science Foundation, grant 725/08.

Methodology, and Philosophy of Science, Amsterdam: North-Holland Publishing Company, pp. 709-722. Reprinted in Davidson 1980, pp. 245-259.
---. 1973b, "Radical Interpretation", dialectica 27, pp. 314-328. Reprinted in Davidson 1984, pp. 125-139.
---. 1974, "Belief and the Basis of Meaning", *Synthese* 27, pp. 309-323. Reprinted in Davidson 1984, pp. 141-154.
---. 1975, "Thought and Talk", in S. Guttenplan, ed., *Mind and Language*, Oxford: Clarendon Press, pp. 7-23. Reprinted in Davidson 1984, pp. 155-170.
---. 1980, *Essays on Actions and Events*, Oxford: Clarendon Press.
---. 1983, "A Coherence Theory of Truth and Knowledge", in D. Henrich (ed.), *Kant oder Hegel?*, Stuttgart: Klett-Cotta. Reprinted in Davidson 2001, pp. 137-153.
---. 1984, *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press.
---. 1985, "Reply to Harry Lewis," in B. Vermazen and M.B. Hintikka, eds., *Essays on Davidson: Actions and Events*, Oxford: Clarendon Press.
---. 1987a, "Knowing One's Own Mind", Proceedings and Addresses of the American Philosophical Association 60, pp. 441-458. Reprinted in Davidson 2001, pp. 15-38.
---. 1987b, "Problems in the Explanation of Action", in P. Pettit, R. Sylvan, & J. Norman, eds., *Metaphysics and Morality: Essays in Honour of J. J. C. Smart*. Oxford: Basil Blackwell, pp. 35–49. Reprinted in Davidson 2004, pp. 101-116.
---. 1990a. "Turing's Test", in K. Mohyeldin Said, W. Newton-Smith, R. Viale and K. Wilkes, eds., *Modeling the Mind*, New York: Oxford University Press. Reprinted in Davidson 2004, pp. 77-86.
---. 1990b, "Representation and Interpretation." in K. Mohyeldin Said, W. Newton-Smith, R. Viale and K. Wilkes, eds., *Modeling the Mind*, New York: Oxford University Press. Reprinted in Davidson 2004, pp. 87-99.
---. 1991, "Three Varieties of Knowledge", in A. Phillips Griffiths, ed., *A. J. Ayer Memorial Essays: Royal Institute of Philosophy Supplement 30*, Cambridge: Cambridge University Press, pp. 153-166. Reprinted in Davidson 2001, pp. 205-220.
---. 1993, "Thinking Causes", in J. Heil and A. Mele, eds., *Mental Causation*, Oxford: Clarendon Press, pp. 3-17.
---. 1994, "Donald Davidson", in S. Guttenplan, ed., *A Companion to the Philosophy of Mind*, Oxford: Blackwell Publishers, pp. 231-236.
---. 1995a, "Could There Be a Science of Rationality?", International Journal of Philosophical Studies 3, pp. 1-16. Reprinted in Davidson 2004.
---. 1995b, "Laws and Cause", dialectica 49, pp. 263-279. Reprinted in Davidson 2005, pp. 201-219.
---. 1997, "Indeterminism and Antirealism", in C. B. Kulp, ed., *Realism/Antirealism and Epistemology*, Lanham: Rowman and Littlefield Publishers, pp. 109-122. Reprinted in Davidson 2001, pp. 69-84.
---. 2001, *Subjective, Intersubjective, Objective*. Oxford: Clarendon Press.
---. 2004, *Problems of Rationality*. Oxford: Clarendon Press.
---. 2005, *Truth, Language and History*. Oxford: Clarendon Press.

HEAL, J. 1997, "Radical Interpretation", in B. Hale and C. Wright, eds., *A Companion to Philosophy of Language*, Malden: Blackwell Publishing, pp. 175-196.

KIM, J. 1984, "Concepts of Supervenience", Philosophy and Phenomenological Research 45, pp. 153-176. Reprinted in Kim 1993, pp. 53-78.

---. 1985, "Psychophysical Laws", in E. LePore and B. McLaughlin, eds., *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, Oxford: Basil Blackwell, pp. 369-386. Reprinted in Kim 1993, pp. 194-215.
---. 1990, "Supervenience as a Philosophical Concept", Metaphilosophy 1 pp. 1-27. Reprinted in Kim 1993, pp. 131-160.
---. 1993, *Supervenience and Mind*, Cambridge: Cambridge University Press.
---. 1998, "The Mind-Body Problem After Fifty Years", in A. O'Hear, ed., *Current Issues in Philosophy of Mind: Royal Institute of Philosophy Supplement 43*, Cambridge: Cambridge University Press, pp. 3-21.
---. 2003, "Davidson's Philosophy of Mind/Psychology", in K. Ludwig, ed., C*ontemporary Philosophy in Focus: Donald Davidson*, Cambridge: Cambridge University Press, pp. 113-136.

LEPORE, E. and LUDWIG, K. 2005, *Donald Davidson: Meaning, Truth, Language, and Reality*. New York: Oxford University Press.

PITOWSKY, I. (1996), "Laplace's Demon Consults an Oracle: The Computational Complexity of Predictions", Studies in the History and Philosophy of Modern Physics 27, pp. 161-180.

PUTNAM, H. 1973, "Reductionism and the Nature of Psychology", Cognition 2, pp. 131-149.
---. 1975, "The Meaning of 'Meaning'", in K. Gunderson, ed., *Language, Mind and Knowledge* (Minnesota Studies in the Philosophy of Science, VII), Minneapolis: University of Minnesota Press, pp. 131-193.

QUINE, W.V.O. 1970, "On the Reasons for Indeterminacy of Translation", *Journal of Philosophy* 67, pp. 178-183.
---. 1987, "Indeterminacy of Translation Again", Journal of Philosophy 84, pp. 5-10.

RAWLING, P. 2003, "Radical Interpretation", in K. Ludwig, ed., C*ontemporary Philosophy in Focus: Donald Davidson*, Cambridge: Cambridge University Press, pp. 85-112.

SHAGRIR, O. 2009, "Anomalism and Supervenience: A Critical Survey", Canadian Journal of Philosophy 39, pp. 237-272.
--- 2010, "Davidson's Notion of Supervenience", in Anat Biletzki, ed., *Hues of Philosophy – Essays in Memory of Ruth Manor*. London: College Publications.

SHEA, N. 2003, "Does Externalism Entail the Anomalism of the Mental?", Philosophical Quarterly 53, pp. 201-213.

SIDER, T.R. 1999, "Global Supervenience and Identity across Times and Worlds", Philosophy and Phenomenological Research 59, pp. 913-937.

STALNAKER, R. 1996, "Varieties of Supervenience", Philosophical Perspectives 10, pp. 221-241.

YALOWITZ, S. 1998, "Causation in the Argument for Anomalous Monism", Canadian Journal of Philosophy 28, pp. 183-226.