The Waning of Materialism

Edited by

ROBERT C. KOONS AND GEORGE BEALER

11

Modest Dualism[1]

Tyler Burge

I am grateful to Bernard Kobes for his sympathetic and insightful survey, 'Burge's Dualism.' He presents concise sketches of my methodology, my views on mind–body causation, and my argument from anti-individualism against materialist token identity theories. I will comment briefly on each of these three topics before centering on his primary contribution—his discussion of compositional materialism. Then I will present a new line of thought on the mind–body problem.

My methodology requires metaphysical *claims*, as distinguished from metaphysical speculation and heuristic guidance, to be grounded as closely as possible in specific knowledge that resides in explanations and judgments in science and common sense. The history of philosophy has shown that a metaphysics that either goes it on its own or takes off from insufficiently specific features of science and common sense tends to be wayward.

Our best understanding of causation comes from reflecting on good instances of causal explanation and causal attribution in the context of explanatory theories. Similarly, our best understanding of what sorts of things exist comes from reflecting on ontological commitments of explanations in science, or clear-cut judgments in common sense. Ontological commitments are what must be the case if such explanations or judgments are to come out true. Of course, both scientific explanations and common-sense judgments are fallible. And sometimes they are reduced to other forms of explanations or judgments. Science and common sense remain our best epistemic routes to understanding both causation and ontology.

Understanding mind–body and body–mind causation is philosophically important and difficult. But such understanding has regularly been distorted by antecedent metaphysical commitments. Thus Descartes' account of mind–body causation as a relation between two self-sufficient substances over-dramatized the problem, to the detriment of nearly all subsequent discussions. The claim of some modern materialists that mental events, properties, and kinds must be associated with a physical mechanism if they are to have causal upshot raises

[1] I thank Ned Block for several astute critical comments on an earlier draft.

the concern that mental attributes are in themselves irrelevant to mind–body or body–mind causation. Neither approach sufficiently grounds its metaphysics in what is known from psycho-physical explanation. We should start with our knowledge that there are causal interactions among psychological events and between psychological events and physical events, and try to understand the interactions in light of scientific explanations and common-sense judgments.[2]

I continue to think that my old argument against the materialist token identity theory is sound. Indeed, it seems to me decisive. Any given token physical neural event that is a candidate for being identical with a mental event could have been associated with different causal antecedents in the distal environment from those it in fact had. The different causal history could have been part of a pattern of physical or social environmental relations that constitutively determined a different kind of mental event, with different representational content. Mental events are different events if they have different representational contents. Such differences in distal causal history are not constitutively determinative of the identity and kind of the neural event. The same type and token neural event could have derived from various causal histories, of the sort that determine different mental events, type and token. So the given neural event could exist in a situation in which the mental event with which it is supposed to be identified did not occur, and in which some other mental event (with a different representational content) occurs instead. So the neural event is not identical with the original mental event. The argument is general. So token neural events are not identical with token (representational) mental events.[3]

Kobes expounds a closely related argument: token physical events in the brain are not constitutively dependent for having the natures that they have on relations to specific attributes in the physical or social environments. Representational mental events are constitutively dependent for having the natures that they have on such relations. Token events are different if their natures are. So token physical events in the brain and representational mental events are not identical.

I think that this argument is implicit in the argument that I gave against token identity theories in 'Individualism and the Mental.' The modal points made in the first argument were grounded in my view of the natures—or basic explanatory attributes—cited in the second. Kobes in effect emphasizes the different forms of individuation in different sciences, and emphasizes that the modal argument that I gave has its underpinnings in the different identity conditions of basic explanatory kinds referred to in the sciences.

[2] For elaboration of these points see my 'Mind–Body Causation and Explanatory Practice' and 'Postscript to "Mind–Body Causation and Explanatory Practice,"' both in Burge (2007).

[3] The argument was first given in 'Individualism and the Mental,' section IV, in Burge (2007). The argument is further defended and elaborated in 'Mind–Body Causation and Explanatory Practice.'

I focus now on Kobes's interesting and, I think, successful attempt to give some substance to my doubts about a materialist view that does not maintain *identities* between mental and physical states or events, but holds that mental states and events are *composed* of physical entities. Call this view *compositional materialism*. This view is less committal than the materialist token identity theory. A statue could be composed of marble. The statue is not identical with the marble. The same marble could have composed a different statue, or no statue at all. Then the marble would have existed but the statue would not have existed. The statue is materially composed of the marble. So identity is a stronger relation than material composition.

The natural sciences are built on the view that more complex material entities, such as planets, crystals, and plants are materially composed of material parts. Chemistry and physics were conceived in terms of material composition from the beginning. Some biologists once thought that explaining life would be a problem for compositional materialism in biology. But that view was never a majority position. The idea that plants are to be understood in terms of some soul or entelechy within them that is not composed of material entities has never had a serious foothold in western thinking, at least since early-modern science; nor should it have.

Because the framework of material composition has been explanatorily successful in the natural sciences, it is prima facie reasonable as a heuristic strategy to explore whether relations of material composition hold between psychological entities and neural or chemical entities. But *belief* in compositional materialism at this stage of inquiry seems to me not to be reasonable. The difference is between heuristic worth-trying and having reason for belief.

In the first place, the entities that psychology theorizes about (psychological states and events) are not perceived as material in the way that planets, salt crystals, and plants are. Although much of the *evidence* for psychology is physical movement conceived functionally (more specifically, as behavior), what psychology tries to understand and explain, and what it theorizes about, is not perceived or introspected as material.

In the second place, there are prima facie differences between psychology (sociology, economics, and so on), on one hand, and the natural sciences, on the other, that ground caution about *assuming* that material composition is the right relation between psychological events and brain events, merely on the basis of the success of the framework of material composition in the natural sciences. There are widely articulated problems in understanding how consciousness and qualitative states can be understood as materially composed. And there are attributions of reason in parts of psychology that have no analog in the natural sciences. Material composition is not a relation that grounds theorizing in mathematics or logic. As will emerge later in this chapter, I think that because psychology makes essential reference to structures from these disciplines, there is some question whether a framework of material composition applies

within relevant areas of psychology. In sum, the prima difficulties in construing psychology in terms of a material-compositional framework—difficulties that have concerned philosophers for centuries—are much more substantial than they ever were in biology.[4]

In the third place, there are large areas of psychology in which compositional materialism has literally no positive support in the explanations or methods of the science. Although there are limited correlations between neural events and low-level sensory states, which give some hope to a material-compositionalist view about them, the view has no positive support in science as applied to propositional thought, and nearly none as applied to consciousness.

I believe that there is no good reason to believe compositional materialism. I think, however, that it is the least implausible form of materialism about the mind. I will pursue critical discussion of the view as applied to *propositional thought*, an area where it seems to me to be particularly doubtful. Kobes concentrates his attention on this area, and I join him in this focus.

I like Kobes's approach. He explores ways the world could empirically turn out to be, avoiding strong commitment to any of the ways. Unlike most authors who have written on this topic, Kobes does not claim more than he or anyone has good and specific reason to believe.

Kobes introduces a highly knowledgeable mathematical archangel as a heuristic device. He proposes two scenarios. On both, the archangel can use facts about fundamental physical entities over all of space–time to compute chemical and biological facts, including neural facts, regarding a community of psychological subjects. On both scenarios, Kobes stipulates that assuming global supervenience of the representational aspects of psychology on the physical, and bracketing issues regarding phenomenal character, the archangel can also compute the representational facts over the community of subjects. The two scenarios differ over what neural states and events that systematically correlate with the representational facts are computed *from*.

On the first, the archangel can discern at the neural level, prior to attributing propositional attitudes to the subjects, a network of states and events that interact causally, and that are candidates for composing the intentional [representational] states and events to be attributed at a later stage of the computation.

The idea is that the archangel computes certain salient, discrete neural states and events in a causal network. At a later stage of the computation—a stage that invokes psychological kinds and principles—the already distinguished neural states and events can be seen as specifically correlating with representational

[4] There have been massive attempts in mainstream philosophy since the 1950s to show that worries about materialism are just mistaken. I think that in some cases the attempts succeed in showing that certain a priori arguments in favor of dualism are unsound. And in some cases, the attempts set out prima facie empirically possible materialist pictures regarding *certain* psychologically relevant phenomena. The attempts have, in my view, failed to give good reasons for believing any form of materialism about the mind.

states and events that are individuated anti-individualistically. And at this later stage the correlated neural states and events can be seen as materially composing the representational states and events.

On the second scenario, the archangel cannot compute, without using principles from representational psychology, neural events that form patterns that naturally correlate with the representational states and events. The archangel cannot, independently of psychological concepts and principles, identify a pattern of neural events that correlates structurally with states and events in the psychologically identified structure.

Kobe proposes, 'compositional materialism is false just in case the archangel would have to first recapitulate . . . intentional psychology, and only then seek neural event correlations.' In the second scenario, 'if neural correlates can be found at all for token mental events, the direction of metaphysical explanation [would be] from mental events and their causal patterns to the correlated neural events.' In the first scenario, just as the archangel first identifies 'chemical and physiological kinds and their instances by their physical components, distinctively arranged,' so the archangel identifies, without using psychological kinds or principles, distinctively arranged material items that can later be seen to correlate with and compose instances of psychological kinds.

The key issue for Kobes is whether there are principles, other than psychological ones, for independently demarcating boundaries of neural events and states that are the units out of which psychological events and states are (and are later seen to be) composed. Kobes holds that if demarcation of neural correlates is possible only with help of psychological principles, compositional materialism is false. I believe that his idea is that if ideal non-psychological explanations in science do not independently identify the material complexes that correlate with (and compose) psychological states and events, there is reason to believe that compositional materialism is not true. He holds that the explanatory question is empirically open.

A second issue, which Kobes takes to be independent, is this. Even if the archangel 'computes principled and determinate correlations between [representational] and neural events,' would the correlations do explanatory work analogous to that of compositions in familiar material sciences? If the correlations did not illuminate diachronic causal relations in a way analogous to diachronic relations among materially composed things, then correlation would not suffice to establish material composition. Kobes insightfully supposes, for example, that the neural correlates of psychological causes might be *causally relevant* to the neural correlates of a psychological effect of that cause, without being a neural *cause* of the neural correlate of the psychological effect. In such a case, neural causation would not line up with psychological causation in the way that would be required by compositional materialism.

Again Kobes holds that the facts are empirically unknown. Our actual neural and psychological theories do not provide us with relevant correlations or with

parallel causal structures—neural and psychological. Perhaps the situation is simply a product of our ignorance. Kobes's point is that nothing that we now know makes the situation postulated by compositional materialism an explanatorily illuminating or distinctively likely situation.

Kobes's suggestions are insightful, intriguing, and provocative. I do not see my way to the bottom of these issues. I will, however, hazard some tentative comments.

One doubt starts very far back. Kobes assumes that the archangel can compute biological facts from a starting point of 'only the complete facts about fundamental physical objects, events, fields, laws, and causes over all of space and time.' The archangel can identify 'chemical and physiological kinds and their instances by their physical components, distinctively arranged.'

In contrast to Kobes, I doubt that it is possible to identify the material components of biological kinds and compute the behavior of the components that causally correlates with the behavior of instances of biological kinds, relying entirely on entities' *distinctive arrangement and behavior as described by physics alone*. If, for example, the physical components of biological species or the physical mechanisms of phenotypic expression of genotypes could be identified purely by distinctive physical arrangement and physics-described behavior of physical components, then biological kinds and laws would be reducible to physical kinds and laws much more simply than they seem to be.[5] I think that to identify the physico-chemical components of biological kinds and compute the physico-chemical behavior of the components of biological kinds, one must make use of biological principles.

I am not certain that these points are correct. The issues are very abstract, and Kobes grants the archangel large powers. What seems to me hard to believe is that biological kinds in general coincide with independently identifiable, 'distinctive' patterns of arrangement and movement of physical particles. I find it doubtful that generalizations of physics that apply specifically to the physical components of distinctively biological transactions, can be derived from independent physical observations and independently identifiable principles of physics that govern 'distinctive' movement of those components. Biological kinds are individuated by theoretical notions in biology. I doubt that their components can be picked out as 'discrete' or salient patterns among entities identified in physics.

Like specifications of all kinds and principles, specifications of biological kinds and principles are inevitably intensional. They are not reducible to specifications of finite groupings of particles, for example. I doubt that an effective procedure can compute the component physical behavior of instances of the kinds governed by those principles—even relative to an intensional base of physics-kinds and physics-principles. I think that the relevantly distinct component kinds and movements are distinctive only in light of biological categorization and

[5] I know of no clear sense in which all of biology is reducible to physics.

explanation, not—or at least not in general—independently. The supervenience, and even the compositional dependence, of the subject matter of biology on the subject matter of physics, do not entail that one can *compute* the physical behavior distinctively associated with biological entities from the behavior of entities described in physics together with principles of physics.

Let us suppose that these points are correct. Then it would *not* be a sufficient condition for rejecting compositional materialism that the archangel could not, prior to using psychological kinds and principles, identify physical components of psychological kinds. It would *not* be a sufficient condition for rejecting compositional materialism that the archangel could not compute distinctive patterns of neural states and events that could be slotted into a psychological network, once representational psychological kinds were individuated in an anti-individualistic manner. For instances of biological kinds are *composed* of entities identified in physics, even though (by hypothesis) the distinctive components of biological kinds cannot be identified as salient and distinctive types through concepts and principles available only in physics.[6]

A second doubt concerns the alleged independence of Kobes's two worries about compositional materialist theories. Kobes takes the question whether the archangel can establish representational-neural correlations that show how neural events 'asymmetrically and synchronically sustain representational events' to be independent of a second question. The second question is whether the correlations illuminate diachronic causal relations involving representational events. I do not see how synchronic correlations could be of any ontological interest or integrity if they did not track diachronic causal relations at the level of representational events. So I think that Kobes's two worries are not independent.

Of course, as Kobes recognizes, the key issue regarding compositional materialism is not merely a generic correlation and coincidence, even between causal sequences of neural events and causal sequences involving propositional representational events. It is whether the correlated neural events explain their effects in a way that illuminates causation at the psychological level 'in the manner of familiar sciences of materially constituted things.' As I have long emphasized, composition is a *specific* theoretical relation. The issue is not only whether the two levels of causal structure can be correlated, but also—as Kobes rightly emphasizes—whether they can be correlated *in a manner familiar from sciences that make use of causal aspects of material components to illumine causal aspects of higher level composed kinds.*

Are there the relevant event-by-event, structure-preserving correlations? Do they ground explanations that use the causal powers of the candidate neural and chemical components to explain the causal behavior of the psychological propositional events in a compositional manner?

[6] Ned Block noted that the same point can be made about computations between different levels within physics.

Many philosophers will be inclined to ask, dismissively, how *could* answers to these questions not be affirmative. I believe that such philosophers allow ideology to turn them dogmatic about a very complex and not very deeply explored empirical matter.

Psycho-physics gives us no such correlations. There are promising correlations in pre-perceptual aspects of the perceptual system. For example, firings of banks of neurons in the retina correlate with registered information regarding the spatial distribution of proximal stimulation from light arrays. This registered information forms the first input into the psychological mechanisms that yield visual perception. There are even some nice correlations between aspects of genuine visual *perceptual* representation (such as edge representations) and firings of lines of neurons. Moreover, there are correlations between the timing of neural occurrences in areas of the brain and the timing of certain stages in the formation of visual perceptions. There is certainly the beginning of a massive scientific effort to correlate processes in perceptual systems with neural processes. Correlation and perhaps even composition present a natural paradigm for research in these areas. Still, correlations are currently quite generic, except at the periphery of psychological processes. More importantly, correlation is far from explaining all psychological causation in the compositional manner of the material sciences.

With regard to *propositional* psychological events, there is not even a serious beginning in establishing the correlations needed to support compositional materialism. Propositional thinking is known to have a variable correlation with brain processes, both over time in given individuals and across individuals. This situation is compatible with compositional materialism. But if there is little stability or uniformity in the types of physical events that underlie types of propositional thoughts, there will be limited scope for explanations of patterns of psychological causation in terms of (putatively composing) patterns of neural causation.

In the natural sciences, relative stability and uniformity in relations between kinds of components and kinds of things that they compose facilitates part–whole causal explanation. Of course, science can appeal to many levels of kind specifications that might be used to carry out compositional explanations of psychological states. And we are in the earliest stages of understanding the brain, not to say its relation to psychology. So it would be premature to take the lack of established correlations between neural events and propositional attitude events to show that compositional materialism is false. It seems possible, however, that the plasticity of the brain and the genetic and developmental differences among individuals may prevent compositional forms of explanation from being viable for psychological science. More crucially, neural occurrences may not match up well with propositional attitude occurrences, *even in individual cases*. There might turn out to be no clear correlations, even in individuals at specific times, between particular propositional occurrences (among the propositional occurrences at any given time) and particular neural occurrences (among the neural occurrences at the same time) that could provide the beginning of a compositional materialist account. If science does not provide correlations, there will remain no reason to believe in compositional materialism.

Kobes delineates one way in which a failure of correlation could emerge. As noted earlier, he holds that psychology could provide explanations that feature psychological events as causes of a given physical or psychological event, and could cite further psychological states and events as part of the causal *enabling* conditions for the transaction. At the same time, neural science could cite a complex of neural and chemical causes of the physical event, and a background of causal *enabling* neural and chemical events. There could be a general correlation between causally *relevant* events at the two levels, where causally relevant events include both causing and enabling events. But there could turn out to be no illuminating correlation of any subset of the *causing* neural events with the *causing* psychological events. Yet each theory, psychological and neuro-chemical, could provide a true and illuminating account of why certain effects, including physically specified effects, occur. It seems to Kobes, and to me, an open question whether the two types of causal explanation will line up so that causation at neural/chemical levels is correlated with causation at the psychological level, even in given contexts, in the way required by compositional materialism.

The burden on compositional materialism is heavy. It must correlate neural causes and their effects with psychological causes and their effects. And it must illuminate *psychological* causation, of both physical and psychological effects, in ways familiar from the material sciences. I will not try to formulate a precise assumption about physical causation entailed by compositional materialism. However, causation at the level of wholes must be a physical composite of causation at the level of material parts, for some natural division of the material parts. Of course, causation at the level of the parts can capitalize on physical *relations* among the parts. For the psychological causing event to be composed materially, *psychological* causation must depend on the causation of the material parts in one of the ways familiar from causation in the natural (material compositional) sciences. To *know* that such causation occurs, we must have explanations that take psychological causation to operate in such ways.

For example, can one explain the psychological causation of an occurrent thought that is the conclusion and causal effect of a piece of reasoning in such a way that the inference to that thought, including the inference's causally relevant *rational* aspects, are illuminated by the composite causation of the components of an antecedent chemical or neural event? As I elaborate shortly, it is hard to see how the rational aspects of psychological causation can be illuminatingly explained as a material composite of the causal operations of putative neural or chemical components of the inferential process, even taking into account the physical relations among those components in their causal operation.

Some philosophers claim that appeal to a language of thought is assumed in the psychology of propositional attitudes.[7] The idea is that there is a level of psychological activity that is 'purely syntactical.' Representational content is taken to be a further matter attributed at a different stage of explanation. It is frequently also assumed that the tokens of the syntactical, 'linguistic' items are brain events. (Here one often hears intoned the portentous but utterly misleading slogan that the brain is a syntactic engine.) It is concluded that a language of thought illuminates how occurrences of thoughts—inner 'linguistic' episodes—could be composed of brain processes. The language of thought picture is sometimes said to indicate that compositional materialism is already in place in psycho-physical science.

Such reasoning skates too fast at each turn. In the first place, except perhaps in psycho-linguistics, there is no autonomous account of syntactical processing in cognitive psychology. Theories of perception and perceptual belief, theories of natural inference and practical reasoning, do not attribute a syntax *except* as a structure *in* the representational contents of psychological states. There is no purely syntactical level of explanation in most of cognitive psychology. Even in the syntactical part of psycho-linguistics, syntax appears to be an abstraction from—and the structure of a capacity embedded in—capacities to understand meaningful sentences.

Thus the view that psychology contains a theory of an autonomous 'purely syntactical' level of processing—one that operates independently of a representational capacity—is misleading even as applied to syntactical aspects of psycho-linguistics. It is without *any* solid grounding as applied to perceptual or propositional attitude psychology, indeed the whole representational part of psychology. The idea that all of propositional thinking is, at some level, a processing of syntactical symbols that in themselves are neutral as regards representational content has no scientific basis. Propositional psychology is about thought. Thought has a structure. One can abstract that structure and study its properties. Some thinking—for example, certain deductive proofs—but probably not very much thinking, hinges purely on that structure. Even then, the structural elements are not content-neutral. In fact, the structural elements have representational content that is relevant to rationality-based explanations. The category of logical constant or predicate, for example, gets its content from roles in reasoning with representational contents. Even in purely deductive thought, the logical constants have representational content.[8] The structure has no psychological status apart from its association with representational thought.

In the second place, since syntactical elements in psychology are associated with attributions of contentful representational states, they have no specific

[7] The picture that I will criticize is substantially that of Jerry Fodor (1979).

[8] Hilbertean proof theory is an abstraction from reasoning that uses the representational contents of logical constants.

association, in current psychological theory, with brain states or brain events. Correlating brain states with syntactical states is just as much an open empirical problem as correlating brain states with representational states. In fact, these are different specifications of substantially the same problem, insofar as we have any clarity about what the syntactical states are (to wit, aspects of representational states). So appeal to a language of thought gets us *no* closer to material correlates that are supposed to constitute the compositional material of the syntactical states, or the representational states. It is just part of the materialist ideology, not part of scientific theory, that 'syntactical' states or events are instantiated by specific neural states or events.

Thus in our present state of knowledge, the language of thought hypothesis cannot make compositional materialism more plausible. The hypothesis is not even a plausible gloss on psychology unless standard presentations of it are severely qualified.

One important difference between representational psychology, on one hand, and neuro-physiology and chemistry, on the other, is brought out by anti-individualism. Representational kinds are partly individuated through patterns of causal relations that they bear to entities in an environment that lies well beyond them. Neural and chemical kinds are not individuated in ways that rely on these long-range patterns.

A further difference between representational psychology, particularly the psychology of propositional attitudes, and the neural and bio-chemical sciences is that the part–whole relation of material composition plays no evident role in psychological theorizing. Propositional psychological structures are compositional in a different way. They are broadly rational structures. They include predication and structures of deductive inference. These structures are notoriously not assimilable to physical structures. The bonds between elements in rational structures are not physical but propositional. Since the seventeenth century, it has been evident—in fact, virtually axiomatic in the natural sciences—that the physical world, as described by the natural sciences, does not have the form or content of a text. It is not made up of rational structures at all. It is not a direct expression of reason. The brute nature of physical relations—the difference, for example, as Kant put it, between resisting force and logical negation,[9] or between property inherence and predication—should, I think encourage puzzlement about how physical structures *per se* could compose instances of propositional attitudes.

There have been ways of trying to blunt this concern, insofar as it is ever raised. One might, for example, say that rational structures are relevant only to *norms* governing psychological transactions; material composition is relevant to the ontology and the causal transactions among psychological events that may or may not fulfill the norms. This line is not plausible. Psychological explanation takes the propositional structure of propositional attitudes to be fundamental to

[9] Immanuel Kant (1968: 175–6).

what they are. Psychological explanation gives the rational, propositional aspects of psychological states a causal role. It is not that we are infallibly guided by reason. Obviously we often fall short. Rather, both common sense and scientific explanation indicate that rational, propositional aspects of psychological states and events figure both in the ontological individuation and in the causal powers of psychological occurrences. Indeed, I think that any psychological science that did not acknowledge a role for rational elements in psychological identity and psychological causation could be reasonably counted inadequate.[10]

Let us reflect on examples involving the causal powers of propositional attitudes with rational structure. The thought occurrence that is the conclusion of an individual's deductive inference is caused by transitions involving premise thoughts that incorporate competence with the logical structures of the premises. The rational-structural aspects of the premise attitudes figure causally in drawing the conclusion. An individual's predicating a concept of a perceived particular, in a perceptual judgment, is part of the cause of the individual's practical reasoning about how to deal with the particular. These rational aspects of psychological causation—deductive inference and predication—are not construed as summations of material forces.

Similar examples support a constitutive role for rational structure in the ontology of propositional attitudes. What it is to be an occurrent thinking *since all humans are mortal, if that human is Socrates, then Socrates is mortal* depends constitutively on the propositional structure of the thought. Any event occurrence that lacked that structure would not be the same event.

I have two concerns about compositional materialism, beyond concern about lack of evidence for it. First, it is hard to see how material compositional structures could ground causation by propositional psychological states or events. Second, it is hard to see how material compositional structures are consistent with the nature of propositional psychological states or events. I shall elaborate the causal point first.

Rational, propositional structures are fundamental aspects of psychological causation by propositional states and events.[11] Rational, propositional structures do not appear to be identifiable with structures of material composition. Everything we know about causation by material composites indicates that such causation is not rationally structured. In fact, as noted, this point is a virtual axiom in the natural sciences (physics, chemistry, biology, including neuro-physiology). Moreover, scientific reduction of rational, propositional structures to material compositional structures has little prospect of success. For such a reduction to succeed, natural science would have to show that a *constitutive* aspect of causation

[10] For a discussion of this point, see the end of 'Mind–Body Causation and Explanatory Practice,' in Burge (2007).

[11] The psychological cause could have a psychological effect. Or it could have a physical effect, for example a physical activity motivated by propositional reasoning.

by propositional psychological states and events—their rational structure—is fully *explainable* in terms of the causal properties of brute material composition. There is not the slightest reason to think that such a reduction can succeed. These points derive as much from fundamental commitments of the natural sciences as from fundamental commitments of psychology.

Earlier, I did not attempt a precise formulation of the sort of physical causation produced by material composites. But any such causation must be a physical composite of causation by the material parts (on some natural division of those parts), where the parts operate through their physical relations to one another. Since nature is not a text, such causation does not have rational or propositional structure. And as noted, rational, propositional structures do not appear to be explainable in terms of brute material compositional structures. Thus causation that depends on rational, propositional structure appears not to be identifiable with or reducible to causation by material composites, or by material components of material composites operating through their physical relations to one another. So material composites, such as chemical and neural composites, appear not to exhaustively constitute causation by rational, propositional states or events. Causation by states or events that are material composites is purely causation of material composites. So rational, propositional states or events appear not to be material composites.

More simply: reason is a constitutive structural feature of causation by propositional psychological states and events. According to the natural sciences, reason is not a structural feature of material composites. The causation by material parts of material composites, operating in their physical relations to one another, must suffice to alone compose causation by material composites. It is hard to see how the causal powers and causal structure of material components could alone compose the causal powers and causal structure of causal transactions that hinge on the rational, propositional structures of propositional states and events. So it appears that rational, propositional, psychological causation is not the causation of a material composite. Propositional psychological states and events are material composites only if their causation is purely that of a material composite. So it appears that propositional psychological states and events are not material composites.

The second concern about compositional materialism is similar, but does not feature causation. Here it is: the physical structure of material composites consists in physical bonds among the parts. According to modern natural science, there is no place in the physical structure of material composites for rational, propositional bonds. The structure of propositional psychological states and events constitutively includes propositional, rational structure. So propositional states and events are not material composites.[12]

[12] There is a distant kinship between these arguments and the argument for the simplicity of the soul that Kant criticizes in the second Paralogism. Cf. Immanuel Kant, *Critique of Pure Reason*,

Both arguments depend on a structural contrast between material composites and propositional psychological states and events. Psychological causation hinges often, but then constitutively, on rational propositional structure. And the nature of propositional psychological states and events constitutively involves rational, propositional structure. Causation associated with material composites is, to all appearances *constitutively, not* causation that involves rational, propositional structure. And it is a principle of physical nature that physical structures of material composites are constitutively *not* rational, propositional structures. So it appears that psychological causation by propositional states and events is constitutively not causation by material composites. And it appears that propositional psychological states and events are constitutively not material composites.[13]

These arguments are, of course, very abstract.[14] Perhaps developments in empirical science will show how to overcome them. But the developments would have to be fundamental. At present, the arguments seem to me to support provisional rejection of compositional materialism—independently of doubt that derives from the absence of the correlations and of the explanations of psychological causation that compositional materialism requires.

What are we to say about demands for a *mechanism* for psychological causation? Such demands are often just question-begging insistence on physical mechanism, specified by the natural sciences. Postulating such mechanisms is warranted if and only if they enhance empirical explanation. We want to understand relations between respective explanations and subject matters in representational psychology and the biological and chemical sciences. It does no good to insist that empirical explanation in psychology conform to explanations in very different sciences. There may be no deeper way to explain *how* psychological events cause

A351–361. The argument that Kant criticizes does appeal to the propositional unity of thought, in effect predication, which is one of the rational, propositional structures that my arguments appeal to. But the argument Kant criticizes aims to establish the simplicity of thinkers, and by (alleged) extension, their exemption from dissolution (A356). My arguments are not for simplicity, only against *material* compositeness. And my arguments center not on the thinker but on psychological states and events. They also center on a basic feature of modern natural science—that natural physical relations do not include rational, propositional structures. I believe that my arguments are not subject to any of the objections that Kant raises.

[13] These arguments differ in two respects from the arguments, discussed early in this chapter, against the materialist token identity thesis. First, the earlier arguments depend essentially on anti-individualism and on the view that the identity of neural/chemical events in the brain does not depend on the sorts of long-range patterns of relations to the distal environment that the identities of representational states do. The arguments against compositional materialism do not depend on anti-individualism or on denying that the identities of neural/chemical events in the brain depend on the same long-range causal patterns that the identities of representational states do. Second, the earlier arguments do not center on the causal or constitutive roles of rational, propositional structures, whereas the later arguments do.

[14] It should be obvious that both of the arguments could be modified into additional arguments against materialist token *identity* theories. Such arguments would supplement the two arguments against such theories that I discussed near the beginning of this chapter.

psychological or physical events than to specify the law-like patterns by which they do so, and the neural or chemical patterns some of which seem necessary to those psychological patterns' occurring.

The most popular way of indicating how psychological causation could be composed of physical causation has been to appeal to computers. Computers are claimed to be physical machines that 'instantiate' reason. This claim does not solve the problem. Computers' processes express reason only insofar as we give them programs and interpret their processes in accord with those programs. There is nothing in the computer's physical processes *per se* that makes the computer reasonable, or explains whatever rational causation might occur in computers. Actual computers do not reason autonomously. They go through a sequence of states that were fashioned to express and amplify our rational states. In them, we simply mimic physical symbolization of our own reasoning, and amplify our reasoning by relying on the computers' processing of those symbols.

I leave open whether more sophisticated robots could reason autonomously. It would, however, not follow from the assumption that a robot reasons autonomously that rational causation in the robot is explained by material composition, even on the further assumption that there are known correlations between its physical states and its propositional states. One still needs to explain rational causation in terms of a composition of physical causal relations. That is what we have some reason to believe cannot be done. Supposing that a robot could reason autonomously does not entail supposing that its reasoning events are *composed* of material processes. The robot's reasoning events would depend for their existence on material processes, just as ours appear to. Its supposed rational propositional events would appear not to be materially composed any more than ours appear to be. Only a cartoonish view about what rejecting material composition amounts to—a view that would see the rejection as postulating immaterial soul stuff in the robot—would have to be embarrassed by a reasoning robot.[15]

The language of thought hypothesis is often conjoined with the computer analogy to try to support materialism. The idea is that reasoning in computers hinges causally on the shape, size, and physical configuration of symbols. In addition to the difficulties with this hypothesis that I catalogued earlier, there is a further one. The argument I gave regarding rational causation applies to any attempt to use the language of thought hypothesis to support compositional materialism. Psychological causation depends on the rational, hence representational, aspects of psychological states. Insofar as the language of thought hypothesis tries to account for psychological causation in terms of the shapes,

[15] I am abstracting from issues about consciousness. I think that robots that have the *sorts* of material bases that they commonly have would not and could not be conscious. I am doubtful about counting such beings autonomous reasoners, without serious qualifications. I claim that even if one *lays aside* issues about consciousness, supposition of autonomously reasoning robots does not threaten objections to compositional materialism about propositional reasoning.

sizes, and configurations of symbols, it fails to connect with the type of causation that is referred to by common sense and psychological explanation in science. Psychological causation that hinges on the rational, propositional properties of psychological states is not independent of representational content in the way that causal processes that hinge on the physical properties of symbols are. As I have indicated, there is no scientific basis for an explanation of psychological causation in terms of an autonomous syntax realized in neural or chemical entities. Even if there were, the account of causation provided by such explanation would fail to explain psychological causation that hinges on the rational aspects of the form and content of psychological events. The language of thought hypothesis is not only ungrounded in scientific explanation. It is irrelevant to explaining the central feature of psychological causation by propositional psychological states and events. Kobes cites my interest in Descartes' apparent view that psychological states and events are distinctive in that their being or nature is grounded in consciousness, activity, power, and point of view, not in substance, soul stuff, or composition. Let us reflect on that list.

With Kobes, I have bracketed consciousness. I think that consciousness is constitutively associated with our psychological being. Constitutive relations between consciousness and specific types of psychological states are complex. I will not take on these issues here.

Not all representational psychological states or events, even propositional ones, are active. But all propositional states and events are *constitutively associated* with activity. For all propositional psychological states and events are constitutively associated with inference—the exercise of the capacity to make use of propositional structure. Inference is activity.

Propositional psychological states and events seem to be constitutively associated with at least generic causal vulnerabilities and powers, including some active powers. Those states and events seem to be constitutively associated with points of view that mark them—representational contents. Unlike material entities, propositional psychological states and events are not identifiable through their material substance, stuffings, or compositions. This conception of the natures of propositional psychological states and events accords with what we now know about them. I see no good reason to believe that such psychological states and events are materially composed. I think it reasonable to think that they are not.

Kobes highlights elements in my dualism that are at odds with traditional dualisms: First, it is not a substance dualism, in the early-modern sense of 'substance.' I have no reason to believe, alas, that psychological events can exist *self-sufficiently*, or *independently* of physical material. Representational psychological events seem to depend for their existence on physical events and material that 'sustain' them. A lot of physical states do not depend on psychological states, but all psychological states seem to depend on physical states. (It may be that the particular underlying physical states could not be what they are if they did not sustain, underlie, or otherwise associate with psychological states.) Second, a

strengthening of this first point: although I am not committed to a belief in global supervenience of the psychological on the physical, I incline toward such a belief, pending better reasons against it than any that I know of. Third, the physical world is without gaps in physical causation, modulo quantum indeterminacy. It is *approximately* causally deterministic. Most or all of these points are uncongenial with most early-modern forms of dualism. My dualism is not only conjectural; it is modest and undramatic. It does not encourage belief that our souls can soar out of the material world.

On the other hand, it is not merely a dualism of concepts. Nor is it merely what is often called a property dualism. It is a dualism of occurrent events as well as states, kinds, and properties. For now, I leave open how to think of psychological *agents*. I am not convinced by Strawsonian claims that one *must* conceive of *every* such agent as having material properties, at least if material properties are not comprised of *relations* of dependence on matter.

Kobes speculates that my modest dualism might be counted an 'extremely weak' form of physicalism. I am inclined to resist this speculation. Psychological states and events depend on the physical. But I see no intellectual substance in counting them physical. They are unlike numbers and logical forms in that they have causal powers and vulnerabilities, and in that they occur in time. They are like numbers and logical forms in seeming to lack material composition, mass, physical force, and physical energy. In fact, propositional psychological states and events are what they are through their having logical forms. None of the primary attributes that we cite in theorizing about them—including logical forms—are cited as physical structures in the natural sciences. I see no clear sense in which propositional psychological states or events are physical. Supervenience is consistent with dualism. To be materialist (or physicalist), a view must claim that psychological entities are themselves material (or physical), not merely that they vary with or depend on material (or physical) entities.

So I do not think that counting my view a weak form of physicalism accounts for the diffidence in my position in recommending—in the passage that Kobes quotes—that philosophers be more open and relaxed about whether some form of materialism (or physicalism) is true.[16] I am no type of physicalist or materialist. But I recommend—not sloth or indifference, but—disinterested, open reflection on the issue.

On relaxation: I think that the question whether something like my form of dualism or some non-reductive form of materialism is true is not momentous. At least, it is not momentous for traditional reasons for caring intensely about

[16] From 'Mind–Body Causation and Explanatory Practice,' in Burge (2007: 360), 'It seems to me that philosophers should be more relaxed about whether or not some form of materialism is true. I think it a thoroughly open—and not very momentous—question whether there is any point in insisting that mental events are, in any clear sense, physical . . . What matters is that our mentalistic explanations work and that they do not conflict with our physicalistic explanations. But it serves no purpose to over-dramatize the conflict between different ontological approaches.'

whether dualism is true. Traditional issues of life after death, the existence of freedom and moral responsibility, and the explanatory powers of natural science within its own domain, do not seem to hinge on the answer to the question.

On openness: I think that we do not know enough about the relation between psychology and the natural sciences to take hard lines for or against materialism. I believe that at this point a modest dualism is clearly more reasonable than materialism. But I think that materialistically motivated empirical research is heuristically tenable. The key to any view in this area is openness to empirical exploration and philosophical reflection. Both positions should be developed in a dialectically open spirit.

A modest dualism, however, cleaves more closely to what we know. It does not make warrantless claims. It is more reasonable than compositional materialism, or other sorts of materialism or physicalism, partly through abnegation, and partly through appreciation of the deep differences between rational structures and physical structures. Psychological events have not been shown to have any attributes that are distinctive of physical events. Their primary attributes are not those cited in the natural sciences.

I confess to a more psychological influence on my dualism. I was educated in philosophy in a climate in which materialism had become smug ideological dogma. It dripped with the more supercilious aspects of blind religious orthodoxy. It was not just that herd instinct in philosophy was itself a danger signal. I thought that there was an implicit hypocrisy in the climate. What concerned me was that many philosophers saw any doubt about materialism as *ipso facto* irrational and unscientific. Many still do. Many philosophers exuded a certainty that was out of line with the speculativeness and lack of force in the grounds supporting their positions. Many still do. Such philosophers assume the mantle of science while contravening its letter and spirit. I hope that new generations of philosophers will do better.

12

Descartes' Revenge Part II: The Supervenience Argument Strikes Back

Neal Judisch

1.

For nearly three decades Jaegwon Kim has argued that antireductionist theories of mind are inconsistent with mental causation. Here I aim to show that if Kim's charge against antireductionism is correct his own reductionist view falls prey to the same plight. Specifically, I argue that (i) Kim's theory salvages mental causation only if mental properties are multiply realizable, physically reducible, and have instances that are causally efficacious in virtue of being *mental* property instances, but (ii) his theory is unable jointly to satisfy these conditions. In section 2 I present the Supervenience Argument designed by Kim to prove that psychophysical causation requires reductionism. In section 3 I highlight two principles upon which Kim's charge against nonreductive physicalism essentially relies, and in section 4 I show how a consistent application of these principles places demands upon theories of mental causation that Kim's reductive functionalism cannot meet. The result is that no theory of mind other than type physicalism is consistent with mental causation if Kim's case against antireductionism holds up. If, on the other hand, Kim's reductive functionalism does the trick, antireductive physicalism has nothing to fear from the Supervenience Argument.

2.

The Supervenience Argument incorporates three central assumptions. The first one specifies that the physical world is causally closed:

> *Closure*. If a physical event has a cause at t, then it has a physical cause at t. (Kim 2005: 15)

OXFORD
UNIVERSITY PRESS
Great Clarendon Street, Oxford OX2 6DP
Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in
Oxford New York
Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto
With offices in
Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam
Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries
Published in the United States
by Oxford University Press Inc., New York

First published 2010

British Library Cataloguing in Publication Data
Data available
Library of Congress Cataloging in Publication Data
Library of Congress Control Number 2010920518
Typeset by Laserwords Private Limited, Chennai, India
Printed in Great Britain
on acid-free paper by
CPI Antony Rowe, Chippenham, Wiltshire
ISBN 978-0-19-955618-2 (hbk.)
978-0-19-955619-9 (pbk.)
1 3 5 7 9 10 8 6 4 2

Contents