

CDD: 128.2

ATTITUDES IN ACTION. A CAUSAL ACCOUNT

LYNNE RUDDER BAKER

Department of Philosophy
University of Massachusetts
AMHERST, MA 01003
USA

lr baker@philos.umass.edu

Abstract: This article aims to vindicate the commonsensical view that what we think affects what we do. In order to show that mental properties like believing, desiring and intending are causally explanatory, I propose a nonreductive, materialistic account that identifies beliefs and desires by their content, and that shows how differences in the contents of beliefs and desires can make causal differences in what we do.

Key-words: mental causation; causal explanation; intentional properties; supervenience; counterfactuals.

It is a truism that what we think affects what we do. We all suppose that we can change the way that people act by changing their minds. Otherwise, people would not spend billions of dollars on education, propaganda, advertising, fund-raising drives, political campaigns, or market research. As obvious as this point is, there lurks a philosophical scruple that makes reflective people stand back and wonder how it is possible that what we think has any causal bearing on what we do. The problem is that deep-seated contemporary assumptions about the nature of reality seem to leave no room for our ordinary convictions such as “what we think affects what we do.” This is the

problem of mental causation: to show how, given plausible background assumptions, it is possible that what we think affects what we do.

Here is one way to see the problem: There are two steps. Step 1: If what we think affects what we do, then what we do has a cause that operates in virtue of its content. "She went to the store because she wanted to buy some milk and she believed that the store was open and had milk for sale." The cause of her going to the store was her belief/desire complex – essentially about buying milk.¹ This belief/desire complex causally explains her going to the store in virtue of the contents that they had – wanting to buy milk, and a believing that the store sold milk. So, Step 1 on the way to the problem of mental causation is this: If what we think affects what we do (in the relevant sense), then our beliefs and desires causally explain our actions in virtue of their content. Differences in content – in what beliefs and desires are about – must be able to make causal differences in the effects of the beliefs and desires.²

Step 2 on the road to the problem of mental causation is the claim that that differences in content of our beliefs and desires cannot make causal differences in the effects of the beliefs and desires. Many actions are physical events. It is commonly held that every physical event that has a cause at all has a complete physical cause. (This is the principle of the causal closure of the physical.) So, if actions are physical events, then actions have complete physical causes, and there is no causal work remaining for content properties – like wanting to buy milk – to do.

¹ For convenience, I shall label the putative mental cause of an action 'a belief/desire complex,' but I take no stand here on whether intentions are independent mental causes.

² This is why holding that beliefs are brain states does not solve the problem of mental causation. If beliefs and desires explained action in virtue of their neural properties, then the fact that they were beliefs and desires – that they had content – would be causally irrelevant. In that case, beliefs and desires would be epiphenomenal.

Even if beliefs and desires are particular brain states (something I doubt), then the brain states that are my beliefs and desires may cause my body to move toward the store. But the properties that actually caused my body to move toward the store would be neural properties that controlled the “motor” part of my brain, that in turn controlled my muscle contractions that moved my body toward the store. It appears that my going to the store has a complete physical cause, and is totally explainable in terms of the laws of motion – without any reference to beliefs or desires. The conclusion seems to be that my wanting to buy milk was causally irrelevant to my going to the store.

So, here we have a puzzle: How can our beliefs and desires causally explain our actions in virtue of their contents? This is one of the problems of mental causation: Given the facts that most actions involve physical motions, and physical motions have complete physical explanations, there seems to be no room for propositional attitudes like beliefs, desires and intentions in the causal or explanatory chain. So, beliefs and desires, identified by content, seem to be epiphenomenal, causally idle, and they make no difference to what we do. The problem of mental causation, then, is that we have a valid argument for the conclusion that what we think never affects what we do.

P₁: If what we think ever affects what we do, then differences in the contents of our beliefs and desires can make causal differences in what we do.

P₂: Differences in the contents of our beliefs and desires cannot make causal differences in what we do.

Therefore,

C: What we think never affects what we do.

One possible response to the problem of mental causation is to abandon the principle of the causal closure of the physical, and accept Cartesian dualism. In that case, our beliefs and desires, identified by their contents, could causally explain our actions even though they are not physical. A second possible response is to hypothesize that content properties are identical to, or supervene on, neurophysiological properties that account for what we do. In that case, the neurophysiological properties that are identical to the content properties could cause our actions without violating the principle of the causal closure of the physical. A third possible response is to abandon beliefs and desires altogether, and to accept eliminativism, the view that there are no propositional attitudes, like beliefs, desires and intentions. In that case, the springs of action have nothing to do with what we think.

None of these possible responses is defensible today: The difficulties with the first possible response, Cartesian dualism, are well-known: How *could* purely mental properties affect purely physical properties? The second possible response, the reduction of content properties to neurophysiological properties, is barred by convincing arguments for externalism. What we think depends not only on neurophysiological properties, but also on our social and physical environments. It is utterly implausible to think that properties like believing that there's a conference in João Pessoa are completely determined by what is in our heads, without regard for the fact that we live in a world that has gatherings like conferences. Neurophysiological theories do not mention such properties. The third possible response, eliminativism, is thoroughly paradoxical: If we really gave up the supposition that beliefs and desires account for action, we could have no defensible legal system, no educational system, no economics or other social sciences, no study of history. (Indeed, we would have no history – no battles, no kings and queens, no ideas, no inventions, etc.) We would be unable to make sense of the world. Not even the idea of making sense would make sense. So, we seem stuck.

Elsewhere, I have taken a different approach altogether and tried to *dissolve* the problem of mental causation rather than to solve it. I have argued that there really is no need to show how it is possible that what we think affects what we do. Our conviction that what we think affects what we do is more secure than *any* metaphysical argument against it. As I suggested above, our understanding each other in terms of beliefs and desires is indispensable to our making sense of each other and of the world that we live in. It is obvious that what we think affects what we do. Looked at this way, the task would be to show how something we know to be true could *seem* to be impossible. In a 1993 article, I tried to show that the problem arose because of dubious metaphysical assumptions (Baker (1993)). If we rejected the metaphysical assumptions that generated it, the problem would dissolve. Although I stand by what I said in that article, here I want to take a different, and more constructive, tack.

I propose to try to solve the problem of mental causation by setting out an account that shows how it is possible that what we think affects what we do. (In that case, P_2 is false: Differences in content of our beliefs and desires do make a causal difference in what we do.) My account has the following characteristics: First, the account is physicalistic in the sense that it posits no immaterial or supernatural substances or properties. It extends to properties the notion of material constitution that I developed for particular things (see Baker (2000)). So, I do not fall into Cartesian dualism. Second, the account construes intentional properties like believing, desiring, and intending as distinctive – that is, such properties are not identical with neural properties or with biochemical properties or with properties that are mentioned in theories of physics. So, I am not reductionistic. Third, the account shows how propositional attitudes (like beliefs, desires, and intentions) can be causally explanatory in virtue of their contents. So, I avoid the conclusion of epiphenomenalism or eliminativism. In short, I shall propose an

account that construes beliefs and desires identified by content in a physicalistic, nonreductive, and causally explanatory light.

Let me make a couple of preliminary comments. (1) I think that the use of the language of types and tokens to speak about the mental has led philosophers astray. Talk of "belief tokens" suggests – falsely, in my view – that beliefs are entities, located somewhere specific in the brain (perhaps in a "belief box"). So, I will talk instead of mental properties (like believing that *p*) and their instantiations (like *S*'s believing that *p* at time *t*). The bearers of mental properties are agents, whole persons. It may be that you can believe that Brazil is beautiful only if you have a certain kind of brain in a certain kind of state, but the believer is you – the whole person, situated in the world – not your brain. (Compare: You may be hungry in virtue of your stomach's being in a certain state, or perhaps your brain's being in a certain state, but the bearer of the property of being hungry is you – not your stomach or your brain.) Persons are embodied, and they are situated in the world. To consider persons in abstraction from their bodies or from their situations in the world is to distort. Rather than speaking of mental events, considered abstractly, and their putative relations to physical events, considered abstractly, I'll usually speak about the instantiation of mental properties that are properties of persons situated in the world.

In effect, to use the language of properties instead of events is to endorse Kim's fine-grained conception of events as property instantiations over Davidson's coarse-grained conception of events as spatio-temporal particulars. So, instead of saying 'mental event token,' I'll say 'instantiation of a mental property at a time,' where propositional attitudes like believing that *p* are paradigmatic mental properties and the bearers of such properties are people.³ The talk of instantiation of

³ Although the word 'instantiation' is unappealing, I use it because I think that 'instance' is confusing. Property instances are sometimes taken to be entities possessing a certain property. When I say "instantiation of the property

properties can be translated back into “event” language, but I think that it is clearer to use the “property” language.⁴ Mental events are not spatiotemporal particulars. When I do speak of mental events, I mean instantiation of a mental property. (When I speak of properties, I will usually mean property instantiations, as the context will make clear.)

(2) The second preliminary comment concerns the term ‘intentional.’ Intentional properties are an important class of properties that extends far beyond mental properties expressible by ‘that’-clauses. Intentional properties, as I use the term ‘intentional,’ are properties whose instantiation entails that beings with propositional attitudes exist. A property that could not be instantiated in a world without beings with beliefs, desires and intentions is intentional. For example, writing a check is an intentional property, because there would be no such thing as writing a check in a world lacking the social and economic conventions that presuppose that people have beliefs, desires and intention. Intentional properties stand in contrast to nonintentional properties – e.g., being a promise as opposed to an audible emission, being a signature as opposed to a mark on paper, being a dance step as opposed to a bodily motion. The audible emission, the mark on paper, the bodily motion could all exist or occur in a world lacking beings with propositional attitudes, but the promise, the signature, the dance step could not.

of being red,’ I do not mean a red thing, but rather the state of affairs of a thing’s being red.

⁴ Since I am not talking about Davidsonian events, it is clearer to speak of properties instead of events. I can make the points that I want to make in Davidson’s terms; but if I did, I would not speak of (Davidsonian) mental events, but rather of mental descriptions. Davidson speaks of descriptions, rather than properties, and “mental events” as Davidson construes them have nothing to do with the points that I want to make. When I say ‘event’ here, I mean property instantiation.

Indeed, many different kinds of things are intentional: events (e.g., a baseball game), objects (e.g., a passport), actions (e.g., voting), properties (e.g., being in debt), dispositions (e.g., being honest), activities (e.g., reading your mail), institutions (e.g., a national bank) – all these are intentional. Intentional language contains terms (e.g., ‘wants to buy milk,’ ‘was elected president,’ ‘paid her taxes’) whose application presupposes that there are beings with beliefs, desires, intentions. For purposes here, what is important is that a property is intentional if and only if its instantiation entails that there exist beings with beliefs, desires and intentions. So, actions – like buying a car, sending an email, or washing the dishes – are intentional events whose occurrence entails that there are beings with beliefs, desires and intentions.⁵

In general: What distinguishes action from mere motion is that actions are instantiations of intentional properties. No intentional property, no action. (So, not everything you do is an action. You snore; you withdraw your hand instinctively from the fire; you digest food – these are not actions.) Even though all actions require instantiation of intentional properties in the sense just discussed, some actions are not intentional in the colloquial sense – the sense of being accidental, or not on purpose. Some actions (say, my knocking over the vase) are unintentional in the colloquial sense (I didn’t intend to knock over the vase), but if knocking over the vase was an action of mine at all, there was *something* that I did intentionally in the colloquial sense (e.g., reach for a cup). If I did nothing intentionally in the colloquial sense (if, say, an epileptic seizure caused my arm to move), then there was no action of mine at all. Even an unintentional action (in the colloquial sense of not

⁵ I assume that the instantiation of content-properties, like believing that *p*, entails that there are beings with beliefs and desires. So believings that *p* are intentional properties. Indeed, intending that *p* is also an intentional property. When we say that an action was intentional, in the colloquial sense, we mean that the agent intended to do it. This is a narrower sense of ‘intentional’ than the one just discussed.

being done on purpose) requires the instantiation of an intentional property (in the broader sense of 'intentional' just discussed).⁶

Instead of focusing on mental properties (or the subset of intentional properties that are the attitudes), I propose that we focus on the whole class of intentional properties – properties whose instantiation presupposes that there are beings with beliefs and desires. I want to broaden the dimensions of the inquiry for two reasons: First, restricted focus on mental properties or mental events has led to a disembodied and de-contextualized construal of propositional attitudes and actions. The wider focus shifts attention to the world-with-us-in-it – which, to my way of thinking, is the place to start. The second reason to broaden the dimensions of the inquiry to intentional properties generally is that what links propositional attitudes and actions is that they are both instantiations of intentional properties. We are more likely to get an account of explanation of action by attitudes if we begin with a conception of how attitudes and actions are formally related: Propositional attitudes and actions are both instantiations of intentional properties. Take away intentional properties, and human action disappears altogether.

1. PROPERTY-CONSTITUTION

The key to my solution of the problem of mental causation is what I call 'property-constitution.'⁷ Property-constitution is analogous to the idea that I developed for understanding particular things in terms of what I have simply called 'constitution.' But instead of using constitution to apply

⁶ Although since *every* action requires instantiation of an intentional property, not every instantiation of an intentional property is an action. The properties of being a felon, or being married, are intentional, but they are not actions.

⁷ Derk Pereboom and Hilary Kornblith have an excellent defense of the general idea of constitution as important for philosophy of mind. See Pereboom and Kornblith (1991).

to objects (e.g., ‘ x constitutes y at t ’), here I’ll consider the conditions under which constitution applies to property instantiations (e.g., ‘ x ’s having F at t constitutes (at t) y ’s having G at t ’). For purposes here, I accept the notion of a layered universe, with different levels of reality. At the bottom level, if there is one⁸, are microphysical properties (like the property of having charge). Higher up are macrophysical nonintentional properties (like the property of being round). Still higher are intentional properties of the sorts just discussed. So, an instantiation of an intentional property (e.g., x ’s being a wedding ring at t) may be constituted by an instantiation of a nonintentional macrophysical property (e.g., x ’s being a piece of gold at t), which is constituted in turn by an instantiation of a nonintentional microphysical property (e.g., x ’s being a particular aggregate of gold atoms at t). The basic idea of property-constitution is that when certain properties are instantiated in certain kinds of circumstances, further properties are instantiated.

To get a feel for property-constitution of propositional attitudes and actions, consider an example. Suppose that a professor says to her class on the first day, “Raise your hand if you are taking this course for credit.” In these circumstances, your hand’s rising constitutes your raising it, and your raising your hand constitutes a declaration that you are taking this course for credit. The property of declaring that you are taking a course for credit is not the same property as your raising your hand since either property can be instantiated without the other.⁹ In other circumstances, your raising your hand may be a request to speak, not a declaration that you are taking a course for credit at all. Or in still other circumstances, your declaring that you are taking a course for credit may be accomplished by signing a piece of paper, not by raising

⁸ See Schaffer, J. (forthcoming) for an argument calling into question the assumption that there is a bottom level.

⁹ I take identity to be classical identity that conforms to versions of “Leibniz’s Law:” If A is identical to B , then A and B share *all* their properties – even modal ones. What is called ‘contingent identity’ is not really identity at all.

credit may be accomplished by signing a piece of paper, not by raising your hand. Neither the property of declaring that one is taking a course for credit nor the property of raising one's hand is identical to the property of one's hand's rising. Both raising one's hand and declaring that one is taking a course for credit are intentional properties, but a hand's rising is not an intentional property at all. Hands may rise in worlds without beliefs and desires. Since your hand's rising, your raising your hand, and your declaring that you are taking a course for credit are three distinct properties, their instantiations are distinct instantiations.¹⁰ So, property-constitution is not identity. (See Appendix A for a definition of 'property-constitution'.)

Many philosophers today are familiar with the idea of supervenience. Roughly, the idea of supervenience is that if properties in a set *B* supervene on properties in a set *A*, then two things that are just alike in their *A*-properties are just alike in their *B*-properties. In this case, the *B*-properties are supervenient, and the *A*-properties are their subvenient base. The thesis of global supervenience is that all properties supervene on microphysical properties in this sense: if two possible worlds are indistinguishable in all their microphysical properties, then they are indistinguishable in all their properties.

Property-constitution is a different relation from supervenience: property-constitution is context-dependent; supervenience is context-independent. For example, the nonintentional property of a person's hand's rising in one context constitutes one's greeting a friend; in another context, an instantiation of the same nonintentional property by the same person constitutes one's voting a certain way in a meeting. If the greeting of a friend supervened on the rising of a hand, then for *every* rising of a hand, there would be a greeting of a friend. Supervenience

¹⁰ This is seen in Kim's criterion for event (property-instantiation) identity: *x*'s being *F* at *t* is identical with *y*'s being *G* at *t'* if and only if *x* = *y*, *F* = *G*, and *t* = *t'*.

applies to *all* instantiations of a given property in that the supervenient property is always instantiated when the subvenient property is.¹¹ But there is no such correlation between constituting and constituted properties apart from circumstances. In different circumstances, the same property of declaring that one is taking a course for credit may have an instantiation that is constituted by signing a piece of paper instead of by raising one's hand; and conversely we have just seen that an instantiation of the property of raising one's hand may constitute something else besides a declaration that one is taking a course for credit.

The general idea behind supervenience is that the instantiation of certain properties determines or settles or "fixes" the instantiation of others. Fix the subvenient properties and the supervenient property will be instantiated. Alternatively, there is no change in a supervenient property without a change in base or subvenient properties. Property-constitution has neither of these features. Recall the professor who said to her class on the first day, "Raise your hand if are taking this course for credit." Suppose that she had said, "Raise your hand if are taking this course for credit" as before, but added, "Leave your hand up if you are a graduate student." After the first request, an uplifted hand constitutes declaration that one is taking the course for credit. After the second request, the uplifted hand constitutes a declaration that one is a graduate student. The same constituting property that was instantiated continuously – the uplifted hand – constitutes different higher level properties without any change in the constituting property of the uplifted hand. So, in contrast to cases of supervenience – where fixing the subvenient property thereby fixes the supervenient property – cases of constitution are different: fixing the constituting property does not fix the constituted property.

¹¹ In the case of 'multiply realized properties,' the converse does not hold. There is more than one subvenient property on which a multiply-realized supervenient property supervenes.

The change in the intentional property constituted by the uplifted hand (from declaring that one is taking the course for credit at one minute to declaring that one is a graduate student at another minute) is a matter of change of context when you said, "Leave your hand up if you are a graduate student." There was no underlying change in the uplifted hand. What makes it the case that the uplifted hand constitutes one property or the other simply does not come "from below."

The properties whose instantiations *constitute* declaring that one is taking a course for credit are instantiated locally, at the same place and time as the declaration. By contrast, the microproperties on which your declaring that you are taking a course for credit *supervenes* are not localized. The properties on which such a declaration supervenes have instantiations that guarantee the occurrence of a declaration that one is taking the course for credit. So, the subvenient properties of such a declaration must include properties on which the conventions of the property of taking a course for credit supervene.¹² The property of declaring that one is taking a course for credit is not just a discrete property that can be instantiated in isolation. Such a property cannot be instantiated at all except in a context of many complex conventions – those of schooling, matriculation, requirements for graduation, and so on. Nonintentional – and ultimately microphysical – properties on which these conventions and intentional properties supervene are instantiated all over space and time. By contrast, constituting properties are instantiated at the same time and place as the properties that they constitute.

Since constituting properties are instantiated locally, in some cases we know which properties constitute which. (E.g. the property of being a piece of plastic constitutes the property of being a drivers' license in my

¹² Those who doubt that the property of promising supervene on microproperties like spin and charge should doubt the truth of global supervenience; it would be mad to doubt that people sometimes make promises.

state.) But supervenience lacks this local character, and we have absolutely no knowledge of the relevant subvenient properties on which intentional properties allegedly supervene. Indeed, the subvenient bases for intentional properties are not only unknown, but almost surely unknowable, and maybe even infinite. It is noteworthy that our lack of knowledge of subvenient bases of intentional properties is no theoretical handicap in the social and psychological sciences. (Economists do not worry about nonintentional properties on which economic transactions supervene. Social psychologists do not worry about nonintentional conditions on which family interactions supervene.) So, it seems that, with respect to intentional properties, supervenience is not relevant to inquiry. If it were, we would be stymied in the absence of knowledge of subvenient bases for intentional properties.

Nevertheless, since supervenience has been at the forefront of recent philosophy, some philosophers may want to assimilate constitution to supervenience.¹³ For example, consider the intentional property of being a credit card. The instantiation of this property by my VISA card is constituted right now by an instantiation of the property of being a piece of plastic. But, of course, the property of being a credit card does not supervene on the property of being a piece of plastic. What makes this piece of plastic instantiate the property of being a credit card is that it is in (what we may call) 'credit-card-favorable circumstances' – economic circumstances of conventions of credit and debt and circumstances of having been issued by a credit-card company, among other things. A supervenience-theorist may insist, however, that the property of being a credit card (though not supervening on the property of being a piece of plastic) does supervene on {the property of being a piece of plastic + the property of being in credit-card-favorable circumstances}.

I do not think that it is theoretically useful to assimilate constitution to supervenience in this way, for the following reason:

¹³ This was suggested to me by Jonathan Schaffer and Brandt van der Gaast.

Unlike supervenience, constitution distinguishes the background conditions from the immediate condition of something's being a credit card. The background conditions are the credit-card-favorable circumstances that include, among other things, all the economic conditions that make it possible for something to be a credit card. Before there were human beings with their beliefs, desires and complicated conventions, it was impossible that there were credit cards. The property of being a credit card cannot be instantiated in the absence of these credit-card-favorable circumstances. These credit-card-favorable circumstances are in place before consideration of whether a particular piece of plastic constitutes a credit card. The credit-card-favorable circumstances may involve properties that were instantiated in the past when the conventions of debt and credit were being established, but are no longer instantiated at *t*. However, if a particular instantiation of the property of being a piece of plastic constitutes an instantiation of the property of being a credit card at *t*, then the properties of being a piece of plastic and of being a credit card are instantiated simultaneously. So, constitution makes a clear distinction between background conditions (e.g., the credit-card-favorable circumstances) and immediate conditions (e.g., being a piece of plastic of a certain sort). The constituting properties are only the immediate conditions of the constituted property.

By contrast, supervenience cannot distinguish between background and immediate conditions: every lower-level property required for something to be a credit card goes into the subvenient base, willy-nilly. No distinction can be made in terms of supervenience between the contribution made by the property of being a piece of plastic of a certain kind (or by the microphysical properties on which that property supervenes) and the contribution made by the credit-card-favorable circumstances (or by the microphysical properties on which the economic conventions allegedly supervene). But economic conventions are a different kind of condition for being a credit card from being a piece of plastic. By distinguishing background conditions from

constituting property, constitution honors that distinction. Supervenience provides no place for such a distinction. So, property-constitution should not be assimilated to supervenience.

By shifting attention from supervenience to property-constitution, we can now solve the problem of mental causation.

2. A CONSTITUTION VIEW OF MENTAL CAUSATION

Property instantiations have constitution relations to other property instantiations (as in one's hand's rising and the declaring that one is taking a course for credit), and property instantiations have causal relations to other property instantiations (as in the professor's request and your response). But constitution relations are not themselves causal relations. (Constitution relations are always "vertical" and causal relations are typically "horizontal.") Although I sometimes omit the term 'instantiation' in 'property instantiation' or 'instantiation of your belief/desire complex,' the relation of property-constitution applies to instantiations, not to (abstract) properties themselves.

Counterfactual conditionals – E.g., 'The window would not have broken if you had not thrown the rock at it' – play a large role in our understanding of causation. Although causation may not be analyzable in noncausal terms (see Anscombe (1981), the truth of a relevant counterfactual is a typical indication of causation. Which counterfactuals are the relevant ones depends on the causal context.¹⁴ The instantiation of an intentional property has different effects in different contexts. (Seeing on TV students burn an American flag may cause outrage or sympathy for the students – depending on one's other beliefs and desires.) In general,

¹⁴ Suppose that I were in a burning building, and my going down the stairs brought about my escape. The counterfactual "If I hadn't gone down the stairs, I would not have escaped," may not be true. It would not be true – although my doing down the stairs did cause my escape – because if I hadn't gone down the stairs, I would have gone down the fire ladder.

which counterfactuals are relevant as indicators of a causal transaction is likewise relative to a context. Without taking a stand on the nature of causation, I shall just assume that relevant counterfactuals are indicators of causation.

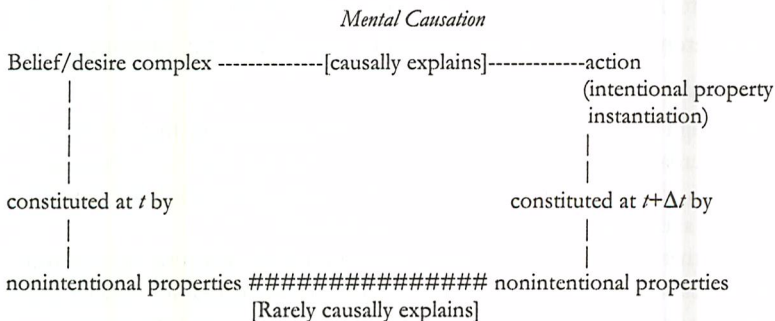
Suppose that we want to explain your promising to call Joe at noon, and that the putative explanation is that you wanted more time before you talked to Joe, and believed that you would be ready to talk to him by noon. The explanatory connection is between your belief/desire complex and your promise. The explanation is defeasible: the belief/desire complex would not have caused you to promise if, say, you had had a stroke or changed your mind at the last minute. By contrast, the properties that constitute your belief/desire complex do not causally explain your promising to call Joe at noon.

Although a full account of causal explanation is beyond the scope of this paper, there are two important conditions of adequacy on causal explanation – one metaphysical, the other epistemological.

An instantiation of property *F* at *t* causally explains an instantiation of property *G* at *t*+ Δt only if:

- (1) The instantiation of *F*, in virtue of being an instantiation of *F*, causes the instantiation of *G*.
- (2) It is possible that knowledge of the instantiation of *F* produces understanding of the instantiation of *G*.

Here is an explanatory schema for mental causation:



The causal connection between the belief/desire complex and the action is indicated by a true relevant counterfactual (which, in turn, may be buttressed by a well-confirmed theory). If you had not wanted more time before talking to Joe, and believed that you would be ready by noon, you would not have promised to call him at noon (unless you had some other reason). There need be no relevant counterfactual, or causal connection, between the properties that constituted your belief/desire complex and the properties that constituted the promise.¹⁵

The point here is a general one about the effects of intentional properties, not just about the effects of mental properties. Suppose that as a result of your promising to call Joe at noon, Joe had a late lunch, say, at 2:00. Suppose that your promising to call Joe at noon was constituted by microproperties M_1 and that Joe's having lunch at 2:00 was constituted by microproperties M_2 . Joe's having lunch at 2:00 is causally explained by your promising to call him at noon. But it by no means follows that M_1 causally explains M_2 . The assumption that M_1 must causally explain M_2 is an artifact of a reductive picture. If we focus on intentional properties that M_1 and M_2 constitute, then it is apparent that the causal relations between intentional properties do not depend on causal relations between nonintentional properties that constitute them.¹⁶ *Which* microproperties constituted your promising to call Joe at noon depends on how the promise was made (e.g., by an utterance or by a nod). But the *effect* of that promise – Joe's having lunch at 2:00 – is indifferent to *how* the promise was made (by an utterance or a nod) and thus indifferent to which microproperties constituted the promise. Your

¹⁵ Although Jaegwon Kim has shown that not all counterfactuals are causal – e.g., “If yesterday had not been Monday, today would not have been Tuesday” – the truth of a relevant counterfactual is clearly a *necessary* condition for causation. See “Causes and Counterfactuals” (1993), pp. 205-207.

¹⁶ Note that the action of promising does not inherit its causal powers from the properties that constitute it. Again, I reject Kim's “causal inheritance principle” and its analogues. See Jaegwon Kim's “Making Sense of Emergence.”

promise would have had the same effect no matter which micro-properties constituted it.

Moreover, an intentional property that has an intentional effect also *indirectly* has nonintentional effects. Say that instantiation of an intentional property I_1 (e.g., your promising to call Joe at noon) *indirectly causes* instantiation of a nonintentional property N_2 (e.g., Joe's body's being in a certain place in the middle of the afternoon) if and only if the instantiation of I_1 causes instantiation of another intentional property I_2 (e.g., Joe's having lunch at 2:00) that is constituted by an instantiation of N_2 . Since the effect of the promise is Joe's eating lunch at 2:00, and Joe's eating lunch at 2:00 is constituted in part by Joe's body's being at a certain place (the restaurant) at 2:00, the promise indirectly caused Joe's body to be in a certain place at 2:00. The relevant counterfactual is true: If you had not promised to call Joe at noon, his body would not have been present there in the middle of the afternoon.¹⁷ The instantiation of an intentional property indirectly causes instantiation of the nonintentional properties that constitute the intentional effect.¹⁸

By contrast, there may be no relevant counterfactuals between the nonintentional properties that happened to constitute your promise and the nonintentional properties that constitute the presence of Joe's body at the restaurant in the middle of the afternoon. To see this, suppose that your promise was constituted by a nod, which, in turn was constituted by an up-and-down motion of your head. It is false that if your head had not moved up and down in the circumstances, then Joe's body would not have been at the restaurant in the middle of the afternoon. The

¹⁷ This suggests that some kind of "downward causation" is a real phenomenon. I shall discuss this matter elsewhere.

¹⁸ The details of the causal mechanisms that connect your promising in the given circumstances to the position of Joe's body several hours later are an empirical matter, almost wholly unknown at this point. The philosophical point is that the connection between these property instantiations supports the relevant counterfactual.

The problem of mental causation arises, I think, because philosophers begin with a Cartesian bifurcation of events into mental and physical, then add levels to get a layered universe. Such philosophers try to adjust the Cartesian picture to make it conform to their physicalist commitments. The adjustments – such as a bottom-up construal of causal explanation – rest on metaphysical principles that are just assumed to be true.²⁸ This approach ends up precluding the possibility of explaining much of what we want to explain – not just actions, but instantiations of intentional properties generally.

My proposal is to avoid the Cartesian picture at the outset: Do not frame the problematic in terms of putative categories of the physical and mental, even combined with an idea of a layered universe. If we take our data from the world that greets us when we wake up in the morning, we will appreciate the importance of intentional properties and the unimportance of a physical/mental bifurcation. The idea of property-constitution allows us to see our world as one world, held together by constitution, rather than as an epiphenomenon of an unseen world of microphysics. If we look at the world this way, we can see how nonreductive, physicalistic mental causation is possible.²⁹

²⁸ For example, Kim's picture of bottom-up causal explanation relies squarely on his causal inheritance principle, what he calls 'the principle of downward causation,' and the principle of mereological supervenience. I hope to discuss Kim's views more explicitly elsewhere.

²⁹ Thanks to Gareth B. Matthews and Jonathan Schaffer for commenting on an earlier draft of this work.

APPENDIX A

A SCHEMA FOR PROPERTY-CONSTITUTION

Here is a schema for property-constitution, analogous to the schema for constitution for particulars that I gave in *Persons and Bodies* (Baker (2000)):

- (P-C) x 's having F at t constitutes at t y 's having G at $t =$
- (a) x 's having F at t is spatially coincident with y 's having G at t ;
 - (b) x 's having F at t is in G -favorable circumstances at t .
 - (c) It is necessary that: $\forall z[(z \text{ has } F \text{ at } t \ \& \ z \text{ is in } G\text{-favorable circumstances at } t) \rightarrow \exists u(u \text{ has } G \text{ at } t \ \& \ u \text{ is spatially coincident with } z \text{ at } t)]$;
 - (d) It is possible that: $(x \text{ has } F \text{ at } t \ \& \ \sim \exists w[w \text{ has } G \text{ at } t \ \& \ w \text{ is spatially coincident with } x \text{ at } t])$;
 - (e) If G is an immaterial property, then F is also an immaterial property.

– Clause (a) insures that if one property instantiation constitutes another, then both properties are instantiated at the same place at the same time.

– Clause (b) specifies that the constituting property instantiation (x 's having F at t) be in circumstances favorable to the constituted property instantiation (favorable to y 's having G at t). For example, for the property of being a credit card to be instantiated, a piece of plastic (or whatever is to constitute the instantiation of the property of being a credit card) must be in credit-card-favorable circumstances – circumstances that include economic conventions of debt and credit.

The *G*-favorable circumstances are those necessary but not sufficient for a property-instantiation *y*'s having *G* at *t*.³⁰

– Clause (c) guarantees that when an instantiation of the constituting property is in the right circumstances (the *G*-favorable circumstances), the constituted property will be instantiated.³¹

– Clause (d) guarantees that property-constitution is not identity. This is so, because, according to (d), it is possible that the constituting property is instantiated when the constituted property is not.³² In

³⁰ We can specify the *G*-favorable circumstances by open sentences; when the *F*-instantiation occurs in circumstances that contain whatever is required to make the open sentences true, then there is a *G*-instantiation.

³¹ Clause (c) also guarantees that property-constitution is asymmetric in cases in which not every property instantiation of *G* is constituted by an instantiation of *F*. (That is, (c) guarantees that property-constitution is asymmetric in cases of “multiple realization”). Necessarily, if the hand makes audible contact with the door in circumstances favorable to being a showing up on time for the interview, then there is a showing up on time for the interview. All that the circumstances needed for there to be a showing up on time was the hand's audible contact with the door. But the following is not necessarily the case: If the showing up on time is in circumstances favorable to the hand's making audible contact with the door, then there is an instance of the hand's making audible contact with the door. That's not the case, because your showing up on time could have been constituted by your ringing the bell instead. So, property-constitution is asymmetric when a property may be constituted by different kinds of “lower-level” property.

³² Clause (d) also guarantees that property-constitution is asymmetric in cases when the *G*-property may be constituted by only one kind of *F*-property. For example, every instantiation of the property of being water is constituted by an instantiation of the property of being an aggregate of H₂O molecules. According to (d), it is possible that there be an aggregate of H₂O molecules that do not constitute a quantity of water (they are not in water-favorable circumstances – they are scattered throughout the universe, say). But (d) is not satisfied in the converse case. Since it is not possible that there be a quantity of water without there being an aggregate of H₂O molecules, (d) would not be

addition, clause (d) distinguishes property-constitution from both strong supervenience and global supervenience.

– Clause (e) guarantees that immaterial properties cannot be constituted by material properties. If there are immaterial property instantiations (like being a Cartesian ego), they are not constituted by physical property instantiations. If there are any immaterial properties (like the property of being a Platonic soul), they do not have instantiations that are constituted by material or physical properties.

REFERENCES

- ANSCOMBE, G.E.M. (1981). "Causality and Determination". In: *Metaphysics and the Philosophy of Mind: Collected Philosophical Papers*, Volume II (Minneapolis, University of Minnesota Press), pp. 133-147.
- BAKER, L.R. (1993). "Metaphysics and Mental Causation". In: John Heil and Alfred Mele (eds.), *Mental Causation* (Oxford, Clarendon Press), pp. 75-96.
- . (2000). *Persons and Bodies: A Constitution View* (Cambridge, Cambridge University Press).
- GULICK, R. van (2001). "Reduction, Emergence and Other Recent Options on the Mind/Body Problem: A Philosophical Overview," *Journal of Consciousness Studies*, 8, pp. 1-34.
- KIM, J. (1988). "Explanatory Realism, Causal Realism, and Explanatory Exclusion". In: Peter A. French, Theodore E. Uehling, Jr., Howard K. Wettstein (eds.), *Midwest Studies in Philosophy XII: Realism and Antirealism* (Minneapolis, University of Minnesota).

satisfied for 'being a quantity of water constitutes being an aggregate of H₂O molecules.'

- . (1993). “Causes and Counterfactuals”. In: Ernest Sosa and Michael Tooley (eds.), *Causation* (Oxford, Oxford University Press).
- . (1993a). “Multiple Realization and the Metaphysics of Reduction”. In: *Supervenience and Mind* (Cambridge, Cambridge University Press).
- PEREBOOM, D. and KORNBLITH, H. (1991). “The Metaphysics of Irreducibility”, *Philosophical Studies*, 63, pp. 125-145.
- SCHAFFER, J. (forthcoming). “Evidence for Fundamentality?”.