LYNNE RUDDER BAKER

WHAT WE DO: A NONREDUCTIVE APPROACH TO HUMAN ACTION

“What is left over,” Wittgenstein once asked, “if I subtract the fact that my arm goes up from the fact that I raise my arm?”1 In this paper, I shall try to give a nonreductive and nondualistic answer to this question. First, I shall simply present a view of human action that allows a straightforward answer to Wittgenstein’s question, and then show how this view of human action fits into a general conception of what a human person or agent is. Next, I shall show how this view can handle what Jaegwon Kim has called ‘the problem of explanatory exclusion.’ Finally, I shall offer reasons why a nonreductive approach to action is needed.

Before turning to theory, however, let me make a pretheoretical survey of the terrain that I want to explore. ‘Action’ is another name for what agents do. An action is either something done intentionally (in the sense that the agent means to do it, or does not do it by accident), or something done by doing something else intentionally. For example, Jack’s alerting the prowler is an action if he unintentionally alerted him by (intentionally) turning on the light. It makes sense, concerning any action, to ask the agent, “Why did you do that?” If the agent in fact did it, he answers appropriately either by giving a reason for doing it, or, as in the case of alerting the prowler, by saying that he had not meant to do that, but rather something else (like turning on the light). An agent does something intentionally only if she can recognize it as something that she has done when it is brought to her attention. Much of what is done intentionally — e.g., your opening the door in order to walk through it — is beneath notice; but if brought to your attention, you would immediately recognize that you had done it. To keep things as simple as possible, I shall ignore what we do unintentionally and focus exclusively on paradigmatic intentional actions and their attendant bodily motions. Now on to ontology.


1 Wittgenstein (1958): I, 621. Wittgenstein himself did not take this question seriously. Although I agree with Wittgenstein that it is misguided to construe action as bodily-motion-plus-some-other-element, I think that the question has merit as long as we don’t look in the wrong place for an answer.

J. Bransen and S. E. Cuypers (eds.), Human Action, Deliberation, and Causation, 249–270.
ONTOGONY AND ACTION

What are the candidates for the relation between raising one’s arm and one’s arm’s going up? A popular answer is that the relation is identity: there is a single event with two descriptions. I want to offer an alternative, according to which a raising of one’s arm is not the same event as one’s arm’s going up, but neither are they two independent events. The relation between a raising of one’s arm and one’s arm’s rising is what I shall call ‘constitution.’ To see how one’s arm’s rising on a certain occasion could constitute a raising of one’s arm on that occasion, consider an analogy with material objects. On a certain view of statues, the relation between Michelangelo’s David and the piece of marble that it is made of is neither identity nor independence. On the one hand, the relation is not identity, because (omitting a lot of argument here) the piece of marble could have existed in a world in which there was no art-world, no statues, and hence no David. If a can exist without b, then a and b are not identical. On the other hand, David is not independent of the piece of marble that constitutes it either: David — that very statue — could not have existed apart from that piece of marble (or one quite like it).

Similarly, a raising of one’s arm on a particular occasion is constituted by one’s arm’s going up on that occasion. As in the case of the relation between the statue and the piece of marble, constitution is not identity: One’s arm could have risen without one’s raising it. Also, as in the statue case, a raising of one’s arm is not independent of one’s arm’s going up; for one could not have raised one’s arm without its going up.

Raising one’s arm is an example of what is known as a basic action — something that an agent (a) does without doing anything else, and (b) can avoid doing without doing anything else. ((b) excludes, e.g., snoring or hiccupping in the usual unintentional way, and it excludes cases of one’s raising one’s arm by picking it up with the other hand.) Not all actions are basic actions: raising one’s hand is a basic action. But voting for the propositions are either basic actions or not. What I want to make use of is a detailed account exactly like this without much deviation.

(C) x constitutes a basic action if
   (i) x and y conflict
   (ii) There are no Preconditions
       (a) Fxt
       (b) It is not associated with anything
         &
   (c) It is intentional

The basic idea is that certain properties cannot be properties constituting a basic action.

2 See my (1997b).
3 I am not committed to Kripke’s thesis of the necessity of origin.
4 Since reference to doing is needed in my account of a basic action (“without doing anything else”), there is no defining ‘do’ (in the relevant sense) in more fundamental terms. A more traditional way of defining ‘basic action’ is to take ‘action’ (rather than ‘do’) as primitive, and define a ‘basic action’ as “an action performed without performing any other action.” I think that these accounts of basic action are extensionally equivalent.

7 I aim to use the idea of “unity and composition,” or “parts” view of identity, to support the theory that an act of unity is not a basic action.
WHAT WE DO

other hand.) Not all actions are basic actions. For example, although raising one’s hand is a basic action, voting for the school budget is not a basic action. But raising one’s hand in certain circumstances constitutes voting for the school budget. All actions involving bodily motions are either basic actions or are constituted by basic actions. Since I want to make use of the notion of constitution, for which I have a detailed account elsewhere, I shall give only a sketchy definition here, without much development.

Definition of ‘constitution’

(C) x constitutes y at t =df

(i) x and y coincide spatially at t, and

(ii) There are essential properties of different kinds, being an F and being a G, and circumstances D such that:

(a) Fxt & Gyt & x is in D at t; and

(b) It is necessary that: ∀z[(Fzt & z is in D at t) → ∃u(Gut

u and z coincide spatially at t)]

(c) It is possible that: (x exists at t & ~∃w[Gwt & w is spatially coincident with x at t]

The basic idea behind constitution is this: When certain things with certain properties occur in certain circumstances, new things with new properties come into existence. For example, pieces of paper consti-

---

5 So, if Henry administered the poison at t, but the victim did not die until t’, Henry’s action occurred at t. The victim’s death (at t’) was a consequence of the action (at t). After the victim died at t’, the action at t acquires a new property: After t’, the poisoning at t is fatal.

6 See “Unity Without Identity,” (typescript 1997). Constitution is a relation that is irreflexive and asymmetrical. Hence, on this understanding of ‘constitution,’ constitution is not identity. Since I am stipulating the definition of ‘constitution,’ anyone who thinks that the target relation is identity should deny that the target relation is ‘constitution’ in my sense.

7 I aim to use the idea of constitution to do the theoretical work that others use the idea of “contingent identity” to do. Since I do not believe that “contingent identity” or “identity at a time” is identity at all (and I reject the “temporal parts” view of ordinary objects), and since I hope to preserve a degree of autonomy for “higher-level” phenomena that are constituted by “lower-level” phe-
tute dollar bills; DNA molecules constitute genes; stones constitute monuments. Constitution applies not only to material objects, but to events as well: chemical reactions constitute volcanic eruptions; utterances constitute promises; and, in general, bodily motions constitute actions. 9 However, (C) is almost empty until relevant essential properties and circumstances are specified for various kinds of things related by constitution. Here I shall first focus on the relation between bodily motions and actions as a species of constitution: Actions are typically constituted by bodily motions without being identical to them. 10

Let me illustrate how an action (like raising one’s arm) is constituted by a bodily motion (one’s arm’s rising). Suppose that Smith raises her arm at t and let us check to see whether, on the definition of ‘constitution,’ Smith’s arm’s rising at t constitutes Smith’s raising her arm at t.

(i):

Smith’s arm’s rising at t and Smith’s raising her arm at t coincide spatially; so (i) is satisfied.

---

nomena, I think that the idea of constitution is superior to the idea of “contingent identity.” (I hope that this is not too obscure. These are controversial ideas that I am trying to work out. I have made a start in Baker (1995b), Baker (1997a).

8 So, constitution relations are contingent relations between things and not to be confused with supervenience relations, which are necessary relations between properties (or families of properties). Someone may object like this: “Constitution may be converted into supervenience by enlarging the supervenience base to include not only the property of being F but also the property of being in circumstances, D. In that case, it would follow from (C) that: necessarily, if x has the properties of being F and of being in D, then there is a y that has the property of being G. That is, the property of being G would supervene on the properties of being F and of being in D, and we don’t need any separate reference to constitution.” While I grant that this move is logically permissible, I do not think it at all helpful. If we are concerned with explanatory properties, we are interested in causal properties. One has causal properties in virtue of being in certain circumstances; being in certain circumstances is not itself a causal property.

9 Note that some examples of constitution are wholly nonintentional, and that constitution does not split off the intentional from the nonintentional.

10 Mental acts and some things that one does intentionally (e.g., refraining from something) seem not to be constituted by bodily motions.

(ii)(a): When Smith rises her arm, she is performing a bodily act, not just a movement. So some natural language sentences, like “Smith will rise her arm,” are not underdetermined by (C) alone. While (C) is a necessary condition for Smith’s actions, it is not a sufficient one. Smith’s arm rising constitutes an action, but she may not be constituting it.

11 It does not mean that Smith has to try to raise her arm in order to do it. She typically try to do it without trying, and try it — doing it - without the idea of trying it. But Smith may have been poise, and her arm may have never been raised (or not at all) if she had not raised it in order to do it, without doing it. If she had not had the idea of trying it, Smith may have never raised her arm and not have raised it in order to do it.

12 Suppose that Smith has been poise, and her arm may never have been raised (or not at all) if she had not raised it in order to do it, without doing it. If she had not had the idea of trying it, Smith may have never raised her arm and not have raised it in order to do it. If she had not had the idea of trying it, Smith’s raising her arm would be a bodily motion, not a raised arm.

13 More precisely, if Smith’s raising her arm is a bodily motion, then the raised arm constitutes Smith’s raising her arm.
(ii)(a): What are the circumstances in which an arm's rising constitutes its being raised? Since raising one's arm in the usual way (without using pulleys, etc.) is a basic action, if Smith raised her arm in the usual way, then she did so without intentionally doing anything else, and she could have prevented the arm's rising without doing anything else. If Smith's arm would have gone up even if she had tried to prevent its rising, then, although her arm would have risen, she would not have raised it. So, Smith's arm's rising constitutes Smith's raising it when the circumstances are such that Smith could have prevented the arm's rising without doing anything else.

Smith's arm's rising at t has the property of being an arm's rising, and Smith's raising her arm at t has the property of being a raising of an arm. Being an arm's rising and being a raising of an arm are causal properties of different kinds. An arm's rising in Smith's situation (she is not underwater) disturbs air molecules; a raising of an arm in Smith's situation (she is at a meeting) causes the chairman to call on her. So, (ii)(a) is satisfied.

11 It does not follow that in order to raise one's arm in the usual way, one has to try to raise it. When I comply with your request to raise my arm, I do not try to raise it — suppose, e.g., that I've had a stroke, or been hypnotised.) The idea of trying figures in the account as a kind of defeasibility condition: I raise my arm at t (in the intended sense) only if: my arm would not have gone up if, without doing anything else, I had tried to prevent its rising.

12 Suppose that Smith wanted to raise her arm, and without Smith's intentionally doing anything else, Smith's arm went up; further suppose that someone had been poised above Smith to pull her arm up if Smith's arm had not gone up. Smith may have thought that she raised her arm, but on this account, she would not have raised her arm since she could not have prevented her arm's rising without doing anything else. I could avoid this consequence only by complicating the account in ways that would not advance the goals of this paper.

13 More precisely, the property of being an arm's rising is an unborrowed causal property of Smith's arm's rising, and a borrowed causal property of Smith's raising her arm; while the property of being a raising of an arm is unborrowed property of Smith's raising her arm, and a borrowed causal property of Smith's arm's rising. This refinement does not affect the satisfaction of (C) by the example, however. For a definition of 'borrowing,' see footnote 21.
(ii)(b): Necessarily, if anything has the property of being an arm’s rising at t in those circumstances — circumstances in which the possessor of the arm could have prevented the arm’s rising without doing anything else — then there is something that has the property of being a raising of an arm at t and is spatially coincident with the arm’s rising at t, and (ii)(b) is satisfied.

(ii)(c): It is possible that Smith’s arm’s rising at t occurs without there being anything that is a raising of an arm at t. (Suppose that, in circumstances other than those that provided the instance of D, someone jerked Smith’s arm up at t.) So, (ii)(c) presents no difficulty.

Since all the clauses of (C) are satisfied in this case, Smith’s arm’s going up at t constitutes Smith’s raising her arm at t.

Now armed with the manifestly non-ad-hoc relation of constitution, let us return to Wittgenstein’s question: “What is left over if I subtract the fact that my arm goes up from the fact that I raise my arm?” Although Wittgenstein emphasized asymmetries between the grammatical first-person and the grammatical third-person in certain contexts, the relation between my arm’s going up and my raising it, I think, is the same relation as that between Smith’s arm’s going up and Smith’s raising it. If so, then we can read off the answer to Wittgenstein’s question from the account of constitution: What is left over if I subtract the fact that my arm goes up from the fact that I raise my arm is this counterfactual: I could have prevented my arm’s rising without doing anything else.\(^\text{14}\) That is, my raising my arm at t consists of (1) my arm’s rising at t and (2) my being able to have prevented my arm’s rising at t without doing anything else.

But how are we to understand this counterfactual — I could have prevented my arm’s rising without doing anything else? From the vast literature on ‘I could have done otherwise,’ it is obvious that there is controversy about how to interpret the counterfactual. For purposes here, I shall assume a broadly compatibilist position: To say, “I could have prevented my arm’s rising without doing anything else” is to say, “I would have prevented my arm’s rising if my attitudes had been different in this way.”

Which of the factors that would not have led to the general arm’s rising would have led to Smith’s arm rising in this particular case? The fact that Smith’s arm was raised my arm was raised. Since this was everything that mattered, the factor would have risen: If I had been able to raise my arm, I would have believed it to be rising.

So, the factor that would not have led to the constitution of Smith’s raising her arm at t — the factor for raising my arm at t — is Smith’s raising her arm at t.

This approach, as I hope to have shown, renders Smith’s vote a constitutive fact about Smith’s raising her arm at t. The vote constituted Smith’s raising her arm at t, and cases are like this one. Constituting z at t is a causal property for both ‘x’ and ‘y’.

I have given a simple example in which

\(^{14}\) Similarly, if I subtract the fact that b followed a from the fact that b was caused by a, what is left over is also a counterfactual.
different in such-and-such ways." So, here we have another schema: Which of my attitudes are such that, had they been different, my arm would not have risen? And different in what ways? There are no general answers to such questions. In any particular case, the answer depends on my reasons for raising my arm at that time. Suppose that I raised my arm at t in order to vote for the school budget. Then, holding everything constant except my attitudes and whatever my attitudes depend on, if my attitudes had been different, my arm would not have risen: If I had not wanted the school budget to pass, my arm would not have risen at t; or even if I still wanted the school budget to pass, but I had believed wrongly that the chair asked for "no" votes at t, my arm would not have risen at t.

So, the expression "I could have prevented my arm’s rising at t without doing anything else" is just a stand-in for relevant counterfactuals about what I would have done if I had had different attitudes. And which attitudes are the relevant attitudes — i.e., the attitudes that were such that, had they been different, my arm would not have risen at t — depend on my reasons for raising my arm at t. And my reasons for raising my arm at t may be discerned from the action that my raising my arm at t constitutes if I did what I intended to do.

This appeal to reasons suggests — as I shall now attempt to show — that not only do bodily motions constitute actions, but also actions constitute further actions: Smith’s raising her arm at t constitutes Smith’s voting for the school budget at t. Although, according to (C), constitution is nontransitive, there are many cases in which x’s constituting y and y’s constituting z guarantees that x constitutes z. And these cases are important for action. x’s constituting y at t and y’s constituting z at t guarantees that x constitutes z at t just in case: the same causal property of y and the same set of circumstances D satisfy (C(ii) for both ‘x constitutes y at t’ and for ‘y constitutes z at t’.

I have an example to show how this applies to actions. But since the example is tedious, and the point intuitive, I shall skip the example.\[15\]

\[15\] Here is the example: Suppose again that Smith’s arm’s rising at t constitutes Smith’s raising her arm at t, and that Smith’s raising her arm at t constitutes Smith’s voting for the school budget at t. Then, Smith’s arm’s rising at t constitutes Smith’s voting for the school budget at t just in case there is a single circumstance and a single causal property that meet two conditions. Consider a complex circumstance that includes not only Smith’s attitudes (e.g., her desire for the school budget to pass and her belief that the chair is calling for votes in
The point is this: If $x, y, z$ range over actions and bodily motions, then if $x$ constitutes $y$ at $t$ and $y$ constitutes $z$ at $t$, then $x$ constitutes $z$ at $t$. For example, if Smith certified that Jones had completed the requirements for a doctoral degree by signing a form, and she signed the form by moving her hand in a certain way, then: (a) her hand’s moving in that way constituted at $t$ her moving her hand in that way; and (b) her moving her hand in that way constituted at $t$ her signing her name; and (c) her signing her name constituted at $t$ her certifying that Jones had completed the requirements for a doctoral degree. Hence, Smith’s moving her hand in a certain way (in the specified circumstances) constituted at $t$ her certifying that Jones had completed the requirements for a doctoral degree; and, still further, Smith’s hand’s moving in that way constituted at $t$ her certifying that Jones had completed the requirements for a doctoral degree.

The moral is two-fold. First, we can treat ordinary descriptions of spatially coincident “nested actions” that employ the ‘by-location’ in terms of constitution relations. Second, since the ‘by-location’ conforms to transitivity, we may treat spatially coincident actions related by constitution as transitive. If Smith does $A$ by doing $B$ at $t$, and does $B$ by doing $C$ at $t$, then we may conclude that Smith does $A$ by doing $C$ at $t$. Favor of the school budget and so on), but also the conventions of voting and of budgeting. Since Smith’s arm’s rising at $t$ and Smith’s raising her arm at $t$ and Smith’s voting for the school budget at $t$ all occur within this complex circumstance, this complex circumstance satisfies (ii) in (C) for both ‘Smith’s arm’s rising at $t$ constitutes Smith’s raising her arm at $t’ and for ‘Smith’s raising her arm at $t$ constitutes Smith’s voting for the school budget at $t’. Second, Smith’s raising her arm at $t$ must have a causal property in the complex circumstance that satisfies (ii) in (C) both for ‘Smith’s arm’s rising at $t$ constitutes Smith’s raising her arm at $t’ and for ‘Smith’s raising her arm at $t$ constitutes Smith’s voting for the school budget at $t’. Suppose that a causal property of Smith’s arm’s rising at $t$ that satisfies (ii) in (C) is the property responsible for the disturbance of the air molecules, and that a causal property of Smith’s voting for the school budget at $t$ is the property of contributing to a faculty pay increase. Now it is easy to find a relevant causal property of Smith’s raising her arm at $t$. For example, the property responsible for the vote-counter’s putting a certain mark in a certain place is a causal property of Smith’s raising her arm at $t$. In the complex circumstance, $D$, Smith’s raising her arm at $t$ has the property of being responsible for the vote-counter’s putting a certain mark in a certain place, and this property satisfies (ii) in (C) both for ‘Smith’s arm’s rising at $t$ constitutes Smith’s raising her arm at $t’ and for ‘Smith’s raising her arm at $t$ constitutes Smith’s voting for the school budget at $t’.

[16 Or another way to allow, say, a special concern for human activities, 17 This is why the basic action of constituting Smith’s voting for the school budget at $t$ constitutes Smith’s voting for the school budget at $t’.
C at \( t \). Putting this in terms of constitution, for any spatially coincident actions, \( x, y, z \), performed by a single agent at time \( t \), if \( x \) constitutes \( y \) at \( t \) and \( y \) constitutes \( z \) at \( t \), we may infer without further ado that \( x \) constitutes \( z \) at \( t \). This is not a full analysis of the ‘by-location,’ however. For the ‘by-location’ is used to refer to noncoincident actions as well. If Jones started a riot by shouting ‘Fire,’ the relation between the riot’s starting and Jones’s shout is causal, not constitutional. Nevertheless, the constitution view of action does give an account of those uses of the ‘by-location’ when it describes the relation between spatially coincident actions.

Let me conclude the discussion of this constitution view of action with two comments. Both concern the contrast between bodily motions (like one’s arm’s moving) and actions. The first comment is that each performance of a basic action of a certain kind entails an occurrence of a bodily motion of a certain kind.\(^{16}\) For example, raising one’s arm entails that one’s arm rises. For every basic action-type \( A \), there is a type \( B \) of bodily motion \( B \), such that, necessarily, for any token of type \( A \), there is an occurrence of a token of type \( B \). That is, there is an entailment between types of basic action and types of bodily motion. By contrast, there is no entailment between types of action generally and types of bodily motion. Rather, for every nonbasic action of a certain type (e.g., voting for the school budget), there is a bodily motion of some type or other, such that the bodily motion constitutes the action. Actions of a single “higher-level” type may be constituted by bodily motions of indefinitely many types, but actions of a single basic-action type are constituted by bodily motions only of a single type.\(^{17}\) So, the first comment is that kinds of basic action are intimately connected to kinds of bodily motions, but that kinds of action generally and kinds of bodily motions are not.

The second comment is that the actions of interest — of moral, social, legal, political interest — are almost always nonbasic actions;

\[^{16}\text{Or an instance of a particular bodily property. This qualification is needed to allow, say, sitting still to be a basic action. Throughout this paper, however, my concern is with the connection between bodily motions and those paradigmatic human actions connected to them.}\]

\[^{17}\text{This fact is connected to facts of constitution: unlike “higher-level” actions, basic actions are directly constituted by bodily motions. Say that \( a \) directly constitutes \( b \) if and only if: there is no \( x \) such that \( a \) constitutes \( x \) and \( x \) constitutes \( b \).}\]
and the particular bodily motions that ultimately constitute the actions of interest are usually of no interest whatever. Indeed, we cannot begin to produce a catalogue of bodily motions that could constitute a vote for the school budget; the types of bodily motions (and basic actions) that could constitute a vote for the school budget are limited only by the imagination of whoever is taking the vote. Since there is not even a definite disjunction of basic actions and bodily motions available for constituting a vote for the school budget, focus on basic actions and bodily motions — types or tokens — is misguided if one’s interest is in the school budget. So, although “higher-level” actions (e.g., promising to meet you for lunch, going to a concert, taking out a loan) are finally constituted by bodily motions, knowing that the promise was made by my mouth’s moving in certain ways is almost never illuminating. What is of greater interest are the circumstances in which my mouth’s moving in certain ways is the creation of an obligation, and it is to these circumstances that the constitution view of action calls attention.

Now, this constitution view of action — according to which Smith’s vote at t is constituted by her raising her arm at t, and Smith’s raising her arm at t is constituted by her arm’s rising at t — accords nicely with what I take to be a promising conception of human persons or agents. It is uncontroversial that there are no actions (in the relevant sense) without agents; and there are no bodily motions without bodies. So, one might ask: what is the relation between agents and human bodies? Perhaps unsurprisingly, the answer is again analogous to the answer about statues and pieces of marble.

As I see it, the relation between a person and a human body is constitution and hence not identity — just as the relation between a statue and a certain piece of marble is constitution and not identity. On the one hand, a piece of marble that in fact is a statue could exist without being a statue (e.g., suppose that the piece of marble had been shaped by the weather on an uninhabited planet), and a body could exist without being a person (e.g., suppose that what would be a corpse in the actual world spontaneously coalesced in a world without human life). So, assuming classical identity as a relation that holds necessarily between each thing and itself, a statue is not identical to a piece of marble, nor is a person identical to her body.\(^{18}\)

On the other hand, a person is essentially an instantiated entity. Just as a representation requires that the represented object exist, a person or representation requires that the first-person exist, and apart from the first-person, the representation is essentially empty. The relation of instantiation is essentially empty.

At least superficially, classical notions of identity, on the other hand, are not necessarily instantiated. Persons is a relation of one person to another person, and there is something significant at the point where we cross the relation.

If everything that constitutes A’s arm constitutes A and constitutes something constituting A’s arm (and similarly for bodies and agents), why postulate a different relation for constituting?\(^{19}\)

Second, in contrast to constituting, the relation of properties of a person to a property of an object is a relation across limits of kind. Thus, the relation of the properties of a person (e.g., the properties of a person’s brain) and the properties of an object (e.g., the properties of a piece of marble) constitutes a new kind of relation.

\(^{18}\) I realize that there are vast and complicated issues embedded in this argument, but I cannot address them here.

\(^{19}\) The point about the brain having a relation of representing to a property, the intentional object, may not be well known. Intentionality is a relation to some property, and the relation may be called a representation by someone (who will present the property that physical properties constitute as the

On the other hand, neither is a person separate from, or independent of, her body: a person is essentially embodied, just as a statue is essentially embodied. The person is not her body-plus-some-other-entity. When a human body (via the brain and the environment) acquires the capability to “support” intentional states rich enough to make first-person reference, a new entity — a person — comes into existence. When a piece of marble is shaped with a certain intention or presented in certain circumstances (in the context of the art-world, perhaps), a new entity — a statue — comes into existence. I make first-person reference to a body when I think of that body as mine — apart from any name or description of it. I am constituted by my body; my actions are constituted by my body’s motions.

At least two features recommend constitution as a basic metaphysical notion. First, the idea of constitution affords an integrated view of the natural order without any bifurcation. The constitution view of persons is no more person/body dualism than it is statue/piece-of-clay dualism. A statue cannot exist without being constituted by a piece of something; nor can a person exist without being constituted by a body. If everything that exists is either a fundamental physical particle or is constituted by something that is constituted by something ... that is constituted by something ... that is constituted by fundamental physical particles, then in a very weak sense, “physicalism” is true. We need postulate no immaterial souls or supernatural beings to describe and explain natural phenomena.

Second, what is crucial about constitution, and what distinguishes constitution from identity is this: if x constitutes y, then y has causal properties that x would not have had if x had not constituted y. For example, if a building constitutes a school, then the school has causal properties (e.g., it relieves its owners from having to pay property taxes on it) that the building would not have had if the building had not constituted a school, or something else with the property of being tax-

---

19 The person remains in existence at least until death, which occurs when the brain has deteriorated to the point of never being able to support first-person intentional states again. The body of a patient in an irreversible coma may or may not continue to constitute a person. If the brain still has the physical capacity to support first-person intentional states, the body still constitutes the person (whether the person ever has any more such states or not); if the brain lacks that physical capacity, then the person has died, and the body no longer constitutes a person — whether the medical staff or anyone else realizes it or not.
exempt. Here’s one way to put it: the fact that the building constitutes a school confers on the building causal properties that the building would not have had otherwise.\(^{20}\) The building borrows its tax-exempt status from the school that it constitutes.\(^{21}\) (Of course, borrowing is a two-way street: the school borrows its size from the building that constitutes it.) The idea of borrowing applies naturally to actions as well: the arm’s rising at \(t\) borrows the property of being a vote for the school budget from the action that the arm’s rising constitutes. If \(x\) constitutes \(y\), then \(y\) has causal properties that it does not derive from \(x\). Hence, if \(x\) constitutes \(y\), \(y\) is irreducible to \(x\).

So, first naturalism and second anti-reductionism — two powerful reasons to continue to try to work out the details of constitution.\(^{22}\) In sum, not only does the constitution view of action offer an answer to Wittgenstein’s famous question, but also it fits into a larger metaphysical picture of persons. Let me now try to put it to work to solve the problem of explanatory exclusion pressed by Jaegwon Kim on any nonreductive metaphysics.

**EXPLANATION AND ACTION**

In several articles, Jaegwon Kim has developed what he calls the problem of explanatory exclusion. The problem stems from a principle, according to which, roughly, “[n]o event may be given more than one complete and independent explanation.”\(^{23}\) The problem that this prin-

---

\(^{20}\) ‘Otherwise’ should be read charitably. If the building had been (i.e., had constituted) a church, then that fact would have conferred on the building causal properties that the building would not have had otherwise.

\(^{21}\) ‘Borrowing’ is a technical term. First, define a new predicate, ‘\(x\) has \(H\) at \(t\) independently of \(x\)’s constitution relations’ to mean this: (a) \(Hxt\) and (b) for any \(y:\) (i) if \(x\) constitutes \(y\) at \(t\), then \(x\) could have had \(H\) at \(t\) (in the given background) even if \(x\) had constituted nothing at \(t\); and (ii) if \(y\) constitutes \(x\) at \(t\), then \(x\) could still have had \(H\) at \(t\) (in the given background) even if whatever constituted \(x\) at \(t\) is such that: it could not have had \(H\) at \(t\) without constituting something at \(t\). Then, \(x\) borrows property \(H\) at \(t\) from \(y\) if and only if: (a) \(x\) has constitution relations to \(y\) at \(t\); & (b) \(x\) has \(H\), but \(x\) does not have \(H\) independently of \(x\)’s constitution relations; & (c) \(y\) has \(H\) independently of \(y\)’s constitution relations.

\(^{22}\) My thinking about constitution was stimulated by Pereboom and Kornblith (1991).

\(^{23}\) Kim (1989), p. 239.
WHAT WE DO

Pinciple generates rests on the plausible assumption that every physical event at \( t \) that has a cause has a complete physical cause at \( t \). This thesis — Kim calls it ‘the causal closure of the physical’ — is that we never need leave the physical domain in order to explain a physical event. Now, given the principle of explanatory exclusion and the principle of the causal closure of the physical, it follows that any cause of a physical event is a physical cause. Add to this the view that a complete causal explanation gives conditions nomologically sufficient for the occurrence of an event, and it follows that any complete causal explanation of a physical event is (or is reducible to) a physical explanation of that event.

The problem of explanatory exclusion may be illustrated as follows: Assume that Smith’s arm’s rising at \( t \) is a physical event. According to the principle of explanatory exclusion, there can be no more than one complete and independent explanation of Smith’s arm’s rising at \( t \). According to the principle of the causal closure of the physical, that one complete and independent explanation is a physical explanation. In that case, there is no room, so to speak, for Smith’s wanting to vote for the school budget at \( t \) — as an irreducible intentional phenomenon — to be part of the cause of her arm’s rising at \( t \). Thus, irreducible intentional explanations (in terms of beliefs and desires, for example) are excluded: there can be no such explanations.

Note that the problem is not that neurophysiological explanations drive out belief-desire explanations. For on the causal-closure principle, Smith’s arm’s rising is no more explainable by an independent neurophysiological (or other macrophysical) explanation than it is by an intentional explanation. For the neurophysiological domain is not causally closed: it is not the case that every neurophysiological event at \( t \) has a complete neurophysiological cause at \( t \) since neurophysiological processes may be interrupted by molecular or even quantum phenomena. Moreover, no neurophysiological state by itself is nomologically sufficient for one’s arm’s rising; for a neurophysiological state that in normal circumstances is sufficient for one’s arm’s rising would have no effect if, say, one were in a straitjacket. The only domain that is even a candidate for being causally closed is that of fundamental physics. So, Kim’s problem of explanatory exclusion does not drive a wedge between neurophysiological and intentional explanations, but between microphysical explanations and macro-explanations of any sort — whether intentional or neurophysiological.
My solution to the problem of explanatory exclusion has several parts. First, I shall define 'physical property' in a way congenial to the nonreductivist. Second, I shall show how intentional explanations of actions — actions like voting for the school budget — do not fall prey to the problem of explanatory exclusion. Finally, and most controversially, I shall show how intentional explanations of bodily motions that constitute actions — bodily motions like an arm's rising, when that motion constitutes a vote for the school budget — escape the problem of explanatory exclusion.

A physical property, as I shall construe it, is either (a) a microphysical property of fundamental particles or (b) a property of things wholly constituted by fundamental particles. I shall assume that a nonreductive materialist holds that all contingently exemplified properties (including intentional properties) are physical properties in this sense, where a contingently exemplified property is one that might not have been exemplified. Similarly, let us define 'physical explanation' as an explanation that entails only that physical properties are exemplified. Given the intuitive plausibility of these definitions, a reductionist who wanted to reject them should have a reason other than the obvious one that the definitions do not support reductionism.

So, the nonreductive materialist, with whom I am aligning myself here, holds that all (correct) explanations are physical explanations. This construal of physical properties (and of physical explanation) shifts the ground of the problem of explanatory exclusion. Whereas, on Kim's view, causal closure pertains only to the microphysical, on my construal of the physical, all domains are physical domains, and the principle of the causal closure of the physical is rendered toothless. So, given a nonreductive construal of the physical, a nonreductivist has no fear of the logical and nonphysical exclusion of physical explanations of physical properties.

Now, it might be de-facto nonreductive, a threat of explanatory exclusion that we could run afoul of. As a nonreductivist, I have taken an explanation of bodily motions by bodily motions to be an explanation that involves an indexical fact that an indexical fact that I would not be directly relevant to Smith's argument. The rising of an arm's vote for the school budget without an arm's rising would not support Smith's argument.

(1) Smith's case for the arm's rising being a neuropathic symptom of belief-depends on terms of appearance. Now, voting for the school budget is a logical explanation of an arm's rising, and Smith's explanation has different terms of appearance than the arguments for the school budget.

24 Elsewhere, I have argued that Kim's views on causation and causal explanation have unacceptable consequences. See Baker (1993). Also, I have argued against Kim's "realist" view of causation in Ch. 4 of Baker (1995a).
25 I am confining attention here to the natural order. If there is a supernatural order as well, it is beyond the scope of this investigation. But I assume that the view here is not contravened by whatever supernatural order there is.
26 Although burden-of-proof arguments are generally frustrating, the reductionist has challenged the nonreductivist to come up with a coherent nonreductive view compatible with materialism. And that is what I am attempting to do. So, my definition of 'physical property' — which, I think, is both intuitively plausible and compatible with a suitably relaxed materialism — begs no questions against the reductionist who has laid down the challenge.

27 Against this, see Smith's (1997).
28 This is where the metaphysical context is critical, as any belief or action explanation may be indelicate. See 1.
no fear from the causal-closure principle. There is no danger that a nonphysical explanation will compete with a putatively complete physical explanation, because, on this nonreductive construal of ‘physical property,’ all explanations are physical explanations.\footnote{Again, I am confining attention to the natural order.}

Now even if the principle of the causal closure of the physical can be de-fanged by a nonreductionist construal of the physical, there still is a threat from the principle of explanatory exclusion. For even if all explanations are physical explanations, a single explanandum would run afoul of the principle of explanatory exclusion if it had two complete independent explanations of any sort. To vindicate the nonreductionist, I have a two-part response — one part to show that intentional explanations of actions are not excluded by nonintentional explanations of bodily motions, the other part to show that nonintentional explanations of bodily motions that constitute actions are not complete explanations of such bodily motions anyway. (1) First, I shall argue that an intentional explanation of Smith’s voting for the school budget would not be excluded by a complete neurophysiological explanation of Smith’s arm’s rising — if there were one — when Smith’s arm’s rising constitutes Smith’s voting for the school budget. (2) Then, I shall argue that, when Smith’s arm’s rising constitutes Smith’s voting for the school budget, the complete correct explanation of Smith’s arm’s rising (a bodily motion) is itself an intentional explanation.

(1) So, first, consider a neurophysiological explanation of Smith’s arm’s rising when the rising constituted a vote. It may appear that such a neurophysiological explanation would exclude any independent belief-desire explanation of Smith’s voting for the school budget at \( t \) in terms of, for example, her wanting to raise teachers’ salaries.\footnote{This example is similar to Malcolm’s famous case, in which a neurophysiological explanation of the man’s climbing the ladder would seem to exclude any belief-desire explanation of the man’s climbing the ladder to retrieve his hat. See Malcolm (1968).} But the appearance would mislead. For a belief-desire explanation of Smith’s voting for the school budget would not compete with a neurophysiological explanation of Smith’s arm’s rising. The two explanations would have different \textit{explananda}. Since, on the constitution view of action, Smith’s arm’s rising is not the same event as Smith’s voting for the school budget at \( t \), ‘Smith’s arm’s rising at \( t \)’ is not the same \textit{explanandum} as ‘Smith’s voting for the school budget at \( t \).’ In this case, an
explanation of the arm’s rising would not be an explanation of the vote. Hence, a complete neurophysiological explanation of the arm’s rising would not compete with an intentional explanation of Smith’s voting for the school budget.

Nor is an intentional explanation of Smith’s voting for the school budget at t replaceable by a nonintentional explanation of Smith’s arm’s rising at t — even though her arm’s rising at t constituted her voting for the school budget at t. For a nonintentional explanation of an arm’s rising would not distinguish between a vote for the school budget and a vote against it. (Indeed, a nonintentional explanation of Smith’s arm’s rising at t would not distinguish between Smith’s voting for (or against) the school budget and her signaling that she wants to speak.) The same type of arm motion — in the same meeting in which Smith wanted to vote for the school budget — that constituted a vote for the school budget would have constituted a vote against the budget just ten seconds later. One could know the nonintentional causes of Smith’s bodily motion without knowing why (or event knowing that) Smith voted for the school budget. So, since an intentional explanation of Smith’s voting for the school budget at t cannot be replaced by a nonintentional explanation of Smith’s arm’s rising at t, there is no danger that an intentional explanation of an action will be excluded by a nonintentional explanation of a bodily motion.

(2) The argument just given still has not disposed of the problem of explanatory exclusion altogether. For there remain cases for which the strategy of splitting the explanandum is not available. What I’ve done so far is to say that all explanations are physical explanations and that

29 As I argued in Baker (1995a), intentional explanations of intentional phenomena are in general not replaceable by what I called ‘physical-counterpart’ explanations of the constituents of the intentional phenomena. Intentional explanations explain different things from physical-counterpart explanations of intentional phenomena. For example, suppose that a certain investment was constituted by a certain computer operation, which was constituted in turn by certain motions of microphysical particles; an explanation of an ensuing bankruptcy in terms of the bad investment, I argued, could not be replaced by an explanation in terms of the motions of the microphysical particles that constituted the bad investment. The motions of the microphysical particles do not provide a better explanation of the bankruptcy, or even an explanation of the bankruptcy at all. For, explanations must support counterfactuals, and in a world without economic practices, the same microphysical explanation would suffice to explain the microphysical motions without there being any bankruptcy.

30 For, explanations must support counterfactuals, which is a further stipulation I have (even though in a neurophysiological world) in what I call the Principle of Explanatory Sufficiency (1995a).
different explanations of actions and bodily motions are not in competition, because actions are distinct from bodily motions. But, as we have seen, kinds of basic actions (like Smith's raising her arm) entail particular kinds of bodily motions (like Smith's arm's rising). In a case in which Smith votes for the school budget by raising her arm (and thus her arm rises), we still seem to have two competing explanations of the bodily motion of Smith's arm's rising. A nonintentional explanation (in terms of, e.g., brain states) of Smith's arm's rising at t would seem to exclude an independent intentional explanation of Smith's arm's rising at t (in terms of, e.g., her desire to vote for the school budget).

Even here, I think, the nonreductivist has a response. First, ask: what is the complete cause of Smith's arm's rising at t (when her arm's rising at t constitutes her voting for the school budget at t)? Initially, one might think (as Malcolm seemed to think) that a neurophysiological cause is the complete cause of the bodily motion, and that the neurophysiological explanation drives out any independent belief-desire explanation. But, as we have seen, neurophysiological states are no more nomologically sufficient for an arm's rising than are belief-desire states. If neurophysiological states were nomologically sufficient for an arm's rising, then in any possible world with the same laws as ours and in which an agent was in the relevant neurophysiological state, her arm would rise. But, of course, a possible world with our laws of nature in which the agent's arms were securely tied down would not be a world in which the agent's arm rises. Nomological sufficiency is hard to come by. Indeed, it is so hard to come by that I doubt that we actually know the nomologically sufficient conditions for many events. For (almost?) any kind of macro-event, the laws governing occurrence of events of that kind are all hedged laws, with ceteris paribus conditions that cannot be spelled out. But if we can state no nomologically sufficient conditions for the occurrence of an action or bodily motion, and if complete explanations provide nomologically sufficient conditions for the occurrence of the explanandum-event, then no one can actually show that a nonreductivist violates the principle of explanatory exclusion. Until someone can actually produce

---

Footnote 30: Fodor holds that, when the ceteris paribus conditions are actually fulfilled (even if we cannot specify them), the antecedents of hedged laws are nomologically sufficient for their consequents. I discuss this move critically in Baker (1995a), pp. 113-115.
a violation of a principle that a nonreductivist accepts (or ought to accept), the nonreductivist need not worry.

No doubt many will consider this reply too quick; so I shall try to deepen it. Suppose that Smith’s arm’s rising at \( t \) constituted no action, but was a result of stimulation by electrodes attached to Smith’s head in a laboratory; now compare this event with Smith’s arm’s rising at \( t \), when Smith voted for the school budget. In both cases, Smith’s arm rises; in both cases, if certain counterfactual circumstances had obtained, Smith’s arm would not have risen. (For example, if there had been a power failure (in the electrode case), or if Smith had not understood the chair’s call for votes in favor of the school budget (in the voting case), her arm would not have risen). But the two causal processes that resulted in Smith’s arm’s rising would have been interrupted in the two cases by different circumstances.\(^{31}\) So, since explanations support relevant counterfactuals (to the effect that the \textit{explanandum}-event would not have occurred if certain conditions had been different), and since the relevant conditions in which the \textit{explanandum}-event would have failed to occur are different in the two cases, Smith’s arm’s rising should have different explanations in the two cases. In particular, the complete explanation of Smith’s arm’s rising when the rising was caused by electrodes attached to Smith’s brain is not a complete explanation of Smith’s arm’s rising when the rising constitutes a vote for the school budget. The explanation of Smith’s arm’s rising when the rising constitutes a raising of the arm must support counterfactuals about Smith’s attitudes; but Smith’s attitudes are irrelevant in the explanation of Smith’s arm’s rising in the electrode case. So, different occurrences of one’s arm’s rising have different kinds of complete explanations in different situations.

Moreover, if we say that any explanation that entails that someone has certain beliefs and desires is a belief-desire explanation, then the complete explanation of Smith’s arm’s rising at \( t \) when that rising constitutes Smith’s raising her arm at \( t \) is a belief-desire explanation. To see this, consider the following principle of explanation (one that I hope is unproblematic): If something that did occur would not have occurred in the absence of some condition \( C \), then a complete expla-

\(^{31}\) This is not to say that there is no overlap in the circumstances that would interrupt one or the other of the two processes. For example, Smith’s arm would not have risen in either of the cases if she had had a heart attack just before it would have gone up.
nation of that occurrence will entail that the condition C obtained. So, if something that did occur would not have occurred in the absence of certain beliefs and desires, then a complete explanation of that occurrence will entail that someone had certain beliefs and desires. Now recall our answer to Wittgenstein’s question: If Smith’s arm rising at t constituted Smith’s raising her arm at t, then Smith could have prevented her arm rising at t without doing anything else. And we understood the modal locution ‘... could have prevented ...’ to be a stand-in for a group of counterfactuals about Smith’s attitudes: If Smith’s attitudes had been different — if she had not wanted to vote for the school budget at t, for example — her arm would not have risen at t, and so on. This is so, because the complete causes of bodily motions that constitute actions include states in virtue of which certain intentional counterfactuals are true; and the relevant intentional counterfactuals are the ones in virtue of which the agent has the attitudes that she has. In this case, complete explanations of bodily motions that constitute actions are guaranteed to entail that someone has certain beliefs and desires, and hence to be belief-desire explanations. Therefore, bodily motions can have belief-desire explanations, without violating the principle of explanatory exclusion. Let me lay this out:

1. If Smith’s arm rising at t constitutes Smith’s raising her arm at t, then Smith could have prevented the rising without doing anything else. [Constitution view of action]

2. If Smith could have prevented the rising without doing anything else, then the rising would not have occurred in the absence of certain beliefs and desires. [Explanation of ‘could have prevented’]

3. If the rising would not have occurred in the absence of certain beliefs and desires, then a complete explanation of the rising would entail that someone had certain beliefs and desires — i.e., the complete explanation of the rising would be a belief-desire explanation. [Principle of explanation]

32 This would be so even if intentional explanations were reducible to nonintentional explanations.
If Smith’s arm’s rising at t constitutes Smith’s raising her arm at t, then the complete explanation of Smith’s arm’s rising at t would be a belief-desire explanation. (1-3)

The only way to deny the propriety of a belief-desire explanation of Smith’s arm’s rising at t is to deny that Smith’s arm’s rising at t constitutes Smith’s raising her arm at t. For if Smith’s arm’s rising at t constitutes Smith’s raising her arm at t, then it just falls out of the analysis that Smith’s arm’s rising at t has a belief-desire explanation.

Let me summarize how I have addressed the problem of explanatory exclusion. First, I pointed out that the principle of the causal closure of the physical poses no threat to a nonreductivist, who can construe ‘physical properties’ in a nonreductive way. Then, I argued that the strategy of ‘splitting the explanandum’ insulated intentional explanations of actions from threat by nonintentional explanations of bodily motions. Finally, after arguing that different occurrences of a single type of bodily motion have different kinds of complete causes, I tried to show that the complete explanations of bodily motions that constitute actions are themselves intentional explanations.

IN DEFENSE OF NONREDUCTIVISM

Why does anybody need a nonreductive approach to human action — or to anything else, for that matter? I have two answers, one methodological and one substantive. The methodological answer is this: Only nonreductionism respects the integrity of things that people care about. The motivation for reductionism is to eliminate all but fundamental physical entities and their properties from one’s ontology. On a reductive view, properties that we pretheoretically take to be properties peculiar to persons — properties of being generous or fair-minded or overbearing — are not genuine properties in their own right; rather, a completed physics would be required to reveal what such putative properties really are. Given reductionism, what befalls people, how they fare — as well as what they do — dissolves in a cloud of particles. If you assume, as I do, that what is ultimately real is something that somebody could possibly care about, it is worth trying to develop a nonreductive view of things that matter. So, there is a methodological reason to shun wholesale reductionism.

But there is also a substantive reason to shun wholesale reductionism: One dominant form of reductionism is “the belief that many...
cruel aspects of a whole including its existence and nature are dependent on those of its parts.”33 I think that this belief gets things wrong. Many important properties of things cannot be understood in terms of properties of their parts. The value of an oil painting has nothing to do with the value of the materials, or even of the materials-plus-labor. The effects that the dean has on the philosophy department have nothing to do with any physical properties of the dean’s body. The causal properties that you have in virtue of owning a piece of land do not derive from any properties of your body. Causal properties such as these are no less important for being relational and intentional. Intentional and relational properties are the ones that matter: being employed, being happily married, being able to pay your bills, having friends, and so on. Many causally important properties are not explainable in terms of properties of the constituents of their bearers, and this fact should be recognized by an adequate metaphysics.

CONCLUSION

I have tried to develop a view of human action that fits into a comprehensive metaphysical framework that is both materialistic and nonreductive. Several things recommend this picture: First, it offers a congenial answer to Wittgenstein’s famous question: “What is left over if I subtract the fact that my arm goes up from the fact that I raise my arm?” Second, it allows escape from the problem of explanatory exclusion. Third, it satisfies the methodological and substantive objections to reductionism. For these reasons, I think that a nonreductive approach to human action is worthy of further study.34

Department of Philosophy
University of Massachusetts at Amherst

33 Kim (1984), p. 54. Kim is simply reporting this belief here, not endorsing it.
34 An earlier version of this paper was read in Holland at the Conference on Human Action and Causality, sponsored by the Institute for Research in Philosophy, Utrecht, April 24-26, 1996. My thanks to Michiel ter Hark of Groningen University for insightful comments.
REFERENCES


In his article in the Philosophy of Mind, there is a section where he developed a theory of action, which is interesting because it is based on the idea that actions are not just events in the world, but are also mental states. Fregens, on the other hand, defines the concept of action as a mental event that is accomplished through the use of language and symbols. He argues that the concept of action is fundamental to the understanding of philosophy and is closely related to the nature of human beings.

In the paper by Baker, he explores the relationship between mental causation and the nature of actions. He suggests that mental states are necessary for the occurrence of actions, and that mental states must be causally effective in order to produce actions. To further support his argument, he cites several works that have explored this topic in more depth.

J. Brennan, 1995
© 1999