22

Substance Dualism: A Non-Cartesian Approach

E. J. Lowe

Substance dualism in the philosophy of mind is, naturally enough, commonly thought of on a Cartesian model, according to which it is a dualism of two radically different kinds of substance, one (the 'body') purely material and the other (the 'mind') wholly immaterial in nature. This view is subject to many familiar difficulties. However, the almost universal rejection of Cartesian substance dualism has blinded many philosophers to the possibility of formulating other and more plausible versions of substance dualism. Non-Cartesian substance dualism (NCSD), as it may most perspicuously be called, is a dualism not of *minds* and bodies, but of *persons*—or, more generally, of *subjects of experience*—and their 'organized' bodies. This is an ontological distinction that is chiefly motivated not by some fanciful notion that there could be disembodied persons—although NCSD does not rule out that possibility—but by much more solid considerations which require us, for instance, to distinguish between the *identity-conditions* of persons and their bodies. Much of the intuitive appeal of Cartesian dualism is retained and explained by NCSD, without any of the former's counterintuitive features and metaphysical difficulties. NCSD is, however, still a *non-materialist* position, because it is incompatible even with very weak forms of non-reductive physicalism. In what follows, I shall begin, in section 1, by explaining and justifying NCSD's distinctive ontology of persons, before moving on, in section 2, to present and argue for its novel anti-physicalist account of the metaphysics of mental causation.

1. NCSD'S ONTOLOGY OF PERSONS

1.1. Non-Cartesian Substance Dualism Defined

Dualism in the philosophy of mind is customarily divided into two chief kinds: *substance* dualism and *property* dualism, the former maintaining the distinctness

of mental and physical *substances* and the latter maintaining the distinctness of mental and physical *properties*. But what are we supposed to understand by a mental or physical 'substance' in this context? I shall take it that by a *substance*, here, we should simply mean an *individual object*, or *bearer of properties*. I shall further take it that by a *mental* substance we should mean a bearer of mental or psychological properties, and that by a *physical* substance we should mean a bearer of physical properties. Thus, whereas the property dualist holds that mental and physical properties are distinct, the substance dualist *additionally* holds that certain *bearers* of those properties are distinct—the implication being that substance dualism entails property dualism but not *vice versa*. I assume, incidentally, that both kinds of dualism entail dualism with regard to mental and physical *states* and *events*, since I take these to consist in the exemplification of properties by objects at times.[1] All this being so, what is needed at this point is a defensible account of the two key concepts of a *mental* property and a *physical* property. These, it seems clear, are distinct *concepts*, although whether the *properties* of which they are concepts are themselves distinct is, of course, one of the main issues under dispute. However, it is one thing to say that these concepts are distinct and quite another to provide an account of that conceptual distinction that would satisfy everyone. In fact, it has proved remarkably difficult to produce an uncontentious characterization of either concept.[2] Fortunately, it is much easier to provide *paradigm examples* of mental and physical properties that almost all parties to the debate will be happy to accept as such. For instance, *pain* and *desire* are universally recognized as being mental properties, while *mass* and *velocity* are universally recognized as being physical properties. In what follows, therefore, I shall take it for granted that the conceptual distinction now at issue is a genuine one and that for practical purposes it can be captured by appeal to such paradigm examples.

Now, substance dualists contend that certain bearers of mental properties, such as pain and desire, are distinct from—that is, are not to be *identified* with—certain bearers of physical properties, such as mass and velocity. What *are* these 'bearers,' though? The bearers of mental properties may be called, quite generally, *subjects of experience*—understanding 'experience' here in a broad sense, to include not just sensory and perceptual experience, but also introspective and cognitive states or, in other words, 'inner' awareness and thoughts.[3] *Human persons*—we ourselves—provide prime examples of subjects of experience, but no doubt we should also include examples drawn from the 'higher' reaches of the non-human animal domain. As for the bearers of *physical* properties, for the purposes of the present discussion I shall mostly be referring to *bodies*, or

[1] See Kim (1980). [2] See Crane and Mellor (1990).
[3] See further Lowe (1996: chapter 1).

parts of bodies—on the understanding that what we are talking about here are not mere lumps or masses of matter, but *organized* bodies and their parts, the paradigm examples being the human body and its organic parts, such as the brain and the neurons and other kinds of cell making up the brain and central nervous system. In these terms, then, the substance dualist may be construed as holding that a person is not to be identified with his or her body, nor with any part of it, such as the brain. On this view, a *person*—*not* the person's body or brain—feels pain and has desires, even if it is true to say that a person feels pain or has desires *only because* his or her body or brain is in a certain physical state. The physical state in question—a certain pattern of excitation in nerve cells, say—is not to be *identified* with the pain or desire consequently experienced by the person, according to the substance dualist. It is at this point that I want to introduce a key distinction between two different types of substance dualism. An implication of what I have said so far concerning substance dualism might seem to be that, according to it, a bearer of mental properties—a subject of experience—*only* bears mental properties, whereas a bearer of physical properties, such as a human body or brain, *only* bears physical properties. This was indeed the view of the most famous substance dualist of all, René Descartes, for whom the human self or ego is an *immaterial* substance.[4] However, even if I am distinct from—not to be identified with—my body or any part of it, as Descartes held, it does not automatically follow that I can have only mental, not physical properties. And, indeed, there is a modern form of substance dualism—which may be called, aptly enough, *non-Cartesian substance dualism*—which differs from Cartesian substance dualism precisely over this point. According to NCSD, it is *I*, and *not* my body or any part of it, who am the bearer of mental properties, just as Descartes maintained. However, unlike Descartes, the advocate of NCSD does not make the further claim that I am not the bearer of any physical properties whatsoever. This sort of substance dualist may maintain that I possess certain physical properties *in virtue of* possessing a body that possesses those properties: that, for instance, I have a certain *shape* and *size* for this reason, and that for this reason I have a certain *velocity* when my body moves.[5] It doesn't follow that such a substance dualist should allow that *every* physical property possessed by my body is also possessed by me, however, for the possession of some of these properties may entail that the thing possessing them is a *body*—and the advocate of NCSD wants to deny, of course, that I am a body. One such property, for instance, would appear to be the property of being *wholly composed of bodily parts*, which is possessed by my body but presumably not by me.

[4] See Descartes (1985a) and, for prominent modern sympathizers, Swinburne (1986) and Foster (1991).

[5] Compare Lowe (1996: chapter 2), and also Baker (2000).

1.2. The Inadequacy of Neo-Cartesian Arguments for Substance Dualism

Setting aside, for the time being, the distinction between Cartesian and non-Cartesian substance dualism, what sorts of arguments can be advanced in favour of such dualism, and how good are they? Some of the best-known arguments have been inherited from Descartes himself and hence their contemporary versions may be described as 'neo-Cartesian.' Two neo-Cartesian arguments in particular are worthy of consideration: the argument from the conceivability of disembodiment and the argument from the indivisibility of the self. For brevity's sake, I shall call them the *conceivability argument* and the *indivisibility argument* respectively.

The conceivability argument has both a strong and weak version, the difference in strength being a difference in the strength of their premises—that is to say, the premises of the strong version of the argument entail those of the weak version, but not *vice versa*. That being so, one might suppose that the weak version is to be preferred, because it assumes less. The *weak* version may be reconstructed as follows.

(1) It is clearly and distinctly conceivable that I should exist without possessing a body.

(2) What is clearly and distinctly conceivable is possible. Hence,

(3) It is possible that I should exist without possessing a body.

(4) If it is possible that I should exist without possessing a body, then I must be distinct from my body. Therefore,

(5) I am distinct from my body.

The *strong* version of the argument replaces premise (1) by

(1*) It is clearly and distinctly conceivable that I should exist without any body whatever existing,

which clearly entails (1).[6] The historical source of (1*) is, of course, Cartesian doubt about the very existence of the physical world in its entirety—a doubt which at least appears to be coherent and therefore to describe a possible state of affairs. As I say, one might suppose the weak version of the argument to be preferable to the strong version because it assumes less. However, it could be contended that (1) is only plausible, or at least is most plausible, in the context of (1*), on the grounds that it is difficult to conceive of oneself as existing in a disembodied state save under the hypothesis that the existence of the entire physical world is an illusion.

Whether we consider the strong or the weak version of the conceivability argument, it presents certain difficulties. Particularly controversial is premise

[6] Compare Meixner (2004: chapter 3).

(2), that what is clearly and distinctly conceivable is possible.[7] Let us grant the truth of premise (1*), that it is clearly and distinctly conceivable that I should exist without any body whatever existing, basing this claim on the coherence of Cartesian doubt about the existence of the physical world. The content of such doubt is something like this: Perhaps, for all that I know, the entire physical world as it seems to be presented to me in perception is non-existent and that perception is wholly illusory. This is a doubt about the nature of the *actual* world, amounting to a surmise that the actual world contains no physical objects although it does contain me and my mental states. I am inclined to think that the surmise is at least a coherent, or logically consistent one. But the question is whether this is enough to establish that there is a possible world in which I and my mental states exist but no physical objects exist. Of course, if the surmise is *correct*, then the actual world is just such a world. But we are not given that the surmise is correct, only that it is *coherent*. To this it may be replied that it suffices that the surmise *could* be correct—it doesn't have actually to *be* correct. But the trouble, I think, is that *we simply don't know* whether or not it could be correct, because there may, for all we know, be some reason why it *couldn't* be correct—a reason that we haven't yet thought of. For instance, it might be that there simply couldn't be a world containing no physical objects, whether or not it also contained me and my mental states.

We might sum up this response to the conceivability argument by saying that the trouble with premise (2) is that it illicitly conflates 'real' or *metaphysical* possibility with mere *epistemic* possibility. That is to say, (2) together with either (1) or (1*) does not serve to ground the truth of (3), that it is *possible* that I should exist without possessing a body, in the requisite sense of 'possible.' The most that can be established by these means is that I *might* actually exist without possessing a body, in an epistemic sense of 'might.' This is the sense of 'might' in which we can say, for instance, that there *might* be an even number greater than 2 which is not the sum of two prime numbers, because we don't know whether or not Goldbach's conjecture is true. But in the *metaphysical* sense of 'necessary', it is either *necessarily the case* or else *necessarily not the case* that every even number greater than 2 is the sum of two prime numbers, so the matter is not in this sense a contingent one. Likewise, then, we cannot assume that it is a contingent matter whether or not I possess a body just because it is true that, in the epistemic sense, I *might or might not* possess a body.

Let me pass on now, rather briefly, to the *indivisibility argument*. This may be reconstructed as follows.

(6) I contain no parts into which I am divisible.
(7) My body is composed of parts.

Therefore, (5) I am distinct from my body.

[7] For well-informed discussion of this issue, see Gendler and Hawthorne (2002).

I take it that (7) is uncontentiously true. Premise (6), however, may appear to be straightforwardly question-begging, since it simply denies that I possess a property that (7) uncontentiously attributes to my body—namely, the property of being a composite entity—and hence, it may be said, already presumes the truth of the conclusion, (5), that I and my body are distinct. Certainly, if the indivisibility argument is to acquire any persuasive force, an independent reason needs to be advanced in support of premise (6). My own view, I should at once declare, is that premise (6) is indeed *true*, but that the most plausible argument for its truth requires (5) as a premise, so that (6) cannot without circularity be appealed to in an argument for the truth of (5). If (5) is to be successfully argued for, then, we need to look elsewhere than to the indivisibility argument. I shall suggest an alternative shortly. What we can conclude so far, however, is that neither of the two neo-Cartesian arguments for substance dualism that we have just examined is particularly compelling.

1.3. An argument for the simplicity of persons

Now I need to explain my chief reason for thinking (6) to be true, that is, for holding *myself*—and, by the same token, any other person—to be a *simple* or *non-composite* entity. This is that I consider the following argument—and note that its first premise includes (5) as a conjunct—to be not just valid but sound.[8]

(8) I am not identical with my body nor with any part of it.

(9) If I am composed of parts, then all of those parts must be parts of my body.

(10) Anything that is wholly composed of parts of my body must either itself be a part of my body or else be identical with my body as whole. Hence,

(11) I am a simple entity, not composed of any parts.

(11), notice, is just another way of expressing (6). The crucial premise here is, of course, (8), to which I shall return shortly. As for premise (9), this should be uncontentious in the context of a debate between substance dualism and its physicalist opponents, since those opponents will naturally agree with (9), holding as they do that I am identical with my body or some part of it, such as my brain. Premise (10) seems equally uncontentious—but more of that in a moment. I should acknowledge, however, that not *all* substance dualists will be happy to assert premise (9). Some, for instance, adopt the following view of the self: they hold that I am distinct from—not identical with—my body, but am *composed* of it and another, immaterial entity, my *soul*. On this view, I am a *body–soul composite*.[9] Such a composite is still a 'substance'—that is, an

[8] For a fuller account, see Lowe (2001).

[9] For discussion and criticism, see Olson (2001) and Kim (2001).

individual object or property-bearer—but one which, in violation of premise (9), contains both parts of my body *and something else*, my soul, as parts. Indeed, Descartes himself sometimes writes as if he endorses this view. I can only say that I find it implausible and unattractive myself.

Another kind of substance dualist will reject premise (10), holding that I am wholly composed of parts of my body and yet am not identical with any part of it nor with my body as a whole. This kind of substance dualist sees the relation between me and my body as being analogous to that between a bronze statue and the lump of bronze of which it is made. On this view, I am *constituted by*, but not identical with, my body.[10] And, indeed, the example of the bronze statue may be seen as posing a threat to a generalized version of (10). For doesn't it show that it simply *isn't true* that anything that is wholly composed of parts of an object *O* must either itself be a part of *O* or else be identical with *O* as a whole? For *the bronze statue*, it may be said, is wholly composed of parts of *the lump of bronze* and yet is neither itself a part of the lump of bronze nor identical with the lump of bronze as a whole. However, here a great deal turns on the question of how, precisely, we are to understand the assertion that the bronze statue is 'wholly composed of parts of the lump of bronze.' If the assertion is taken to mean that we can *decompose* the statue into parts all of which, without remainder, are parts of the lump of bronze, then it is certainly true. For we can decompose the statue into bronze particles, all of which are parts of the lump. But if, instead, the assertion is taken to mean that *all of the parts of the statue are also parts of the lump of bronze*—which is, *mutatis mutandis*, the interpretation that I was assuming in proposing premise (10)—then it is far from evident that it is true. For example: the *head* of the statue—assuming it to be a statue of a man—is a part of the statue and yet is not, plausibly, a part of the lump of bronze.[11]

However, is it not open to the constitution theorist to *agree*, now, with premise (10), interpreted in the manner I intend and instead reject premise (9), although not for the same reason that this was rejected by the proponent of the body–soul composite theory? Cannot the constitution theorist say that, just as the statue has parts, such as its head, that are not parts of the lump of bronze, so *I* have parts that are not parts of my body—but not because I have any *immaterial* part or parts, any more than the statue has? In principle, I agree, the constitution theorist *could* say this. However, I simply don't see what these 'additional' parts could at all plausibly *be*. The reason why the statue has parts that are not parts of the lump of bronze is that it has parts—such as its head—that are, like the statue, *constituted by*, but not identical with, a portion of bronze. If, analogously, I were to have parts that are not parts of my body, they would have to be parts that are *constituted by*, but not identical with, parts of my body—just as, according to the constitution theorist, *I* am constituted by my body as a whole. But there are, surely, *no* such parts of me—no parts of me that are related to parts of my

[10] See especially Baker (2000). [11] See further Lowe (2001).

body in the way that I am related to my body as a whole. As a self or subject of experience, I do not, for example, have lesser or subordinate selves or subjects as parts of me, each of them associated with different parts of my body—as though I were a kind of collective or *corporate* self, on the model of a company or club.[12] At least, it certainly doesn't seem that way to *me*!

1.4. Identity-Conditions and the Replacement Argument

Now I need to return to unfinished business—the search for a plausible argument in favour of the main claim of substance dualism, that *I am not identical with my body nor with any part of it*. This was premise (8) of my argument for the simplicity of the self. We have seen that neither the conceivability argument nor the indivisibility argument is satisfactory for the present purpose. I believe, however, that a much more compelling consideration in favour of (8) is this:

(12) My identity-conditions differ from those of my body or any part of it.

Entities possessing different identity-conditions cannot be identified with one another, on pain of contradiction.[13] But what *are* 'identity-conditions'? Speaking quite generally, the identity-conditions of entities of a kind K are the conditions whose satisfaction is necessary and sufficient for an entity x of kind K and an entity y of kind K to be identical, that is, for them to be *one and the same K*. Thus, for example, the identity-conditions of *sets* are these: a set x and a set y are *one and the same set* if and only if x and y have exactly the same members. In the case of things that *persist through time*, their identity-conditions will also provide their *persistence*-conditions, since a thing persists through time just in case *that same thing* exists at every succeeding moment during some interval of time. Now, there are, of course, notorious difficulties attaching to the question of personal identity, and particularly to the question of what conditions are necessary and sufficient for the identity of the self over time. However, even without being able to settle this question, we may well be in a position to determine that the identity-conditions of the self, whatever they may be, are *different* from those of the body or any part of it, such as the brain.

Here is one sort of consideration that seems quite compelling in this respect. We know already that parts of the human body can be replaced by artificial substitutes that serve the same function more or less equally well, as far as the person possessing that body is concerned. For example, a 'bionic' arm can replace a natural arm and serve the person who owns it pretty much as well as the original did. And, indeed, it seems perfectly possible *in principle* that *every* part of a person's biological body should, bit by bit, be replaced in this fashion, with nerve cells gradually being replaced by, say, electronic circuits mimicking their

[12] For more on the latter notion, see Scruton (1989).
[13] See further Lowe (1989: chapter 4).

natural function.[14] If such a procedure were carried out completely, as it seems it could be, the *person* whose biological body had been replaced by an entirely artificial one would, very plausibly, *survive* the procedure and so *still exist* at the end of it. And yet, clearly, neither that biological body nor any part of it would have survived and still exist. If correct, this shows that the persistence-conditions of human persons are different from those of their biological bodies and their various parts, such as their biological brains, and hence that such persons—*we ourselves*—are *not identical with* those bodies nor with any of their parts. In short, it establishes the truth of (8), the main claim of substance dualism.

This argument—which may aptly be called the *replacement argument*—can be set out rather more formally as follows. Its first premise is:

(13) I could survive the replacement of every part of my body by a part of a different kind,

where by a part of a different 'kind' I mean one that is alien to the kind of thing that my body is, in the way that a bionic arm, say, is alien to the kind of thing that a biological human body is. Equally, of course, a *biological* arm would be, in this sense, 'alien' to a wholly bionic body, in the case of a person with such a body. The second premise is:

(14) My body could not survive the replacement of every part of it by a part of a different kind,

the reason for this being that such a process would leave me with *a body of a different kind*, and an object cannot undergo a change with respect to the kind of thing that it is—not, at least, with respect to the *highest* kind to which it belongs.[15] And here I take it that biological organisms and bionic artefacts, for example, are things which clearly do *not* belong to the same highest kind. Now, (13) and (14) together entail (12)—that I and my body (or any part of it) have different identity-conditions—and thereby entail (8), that I am not identical with my body (or any part of it). Of course, (8) may be inferred *directly* from (13) and (14) by an application of Leibniz's Law: but it is nonetheless important to notice that they entail (12), which itself entails (8), because this renders more perspicuous the relevant difference between persons and their bodies that precludes their identification with one another, namely, the difference in their *identity-conditions*.

Notice, however, that the foregoing argument for substance dualism, while it serves the purposes of *non-Cartesian* substance dualism well enough, is not sufficient to establish the truth of *Cartesian* substance dualism, since the latter

[14] Compare Lowe (1989: 120) and Baker (2000: 122–3).

[15] This claim is central to the sort of 'individuative essentialism' that is defended by David Wiggins in Wiggins (2001: chapter 4), with which I am broadly in agreement: see Lowe (1989). I don't mean to imply, however, that Wiggins himself would have sympathy for the replacement argument.

maintains that the self possesses *only* mental properties, not any physical ones. For the replacement argument doesn't show that the self could survive in a completely *disembodied* state and hence doesn't show that the self might exist even in circumstances in which no physical properties whatever, such as shape or mass, could possibly be attributed to it. The conceivability argument *does* purport to show this, of course, but has been found wanting in persuasive force. As for the indivisibility argument, it, like the replacement argument, cannot be used specifically in support of *Cartesian* dualism, even setting aside the other difficulties that attach to it—for its conclusion is only that I am *distinct* from my body, not that I lack, or could lack, any physical properties whatever.

1.5. The Unity Argument

Although, as we have seen, the indivisibility argument is unsatisfactory, there is another argument that is in some ways reminiscent of it but which, I think, deserves considerably more respect. I also think that it is even more compelling than the replacement argument, since it does not depend upon speculations which at present, it might be said, belong only to the realm of science fiction. I call this the *unity* argument—the unity in question being the unity of *the self* as the unique subject of all and only its own experiences. The first premise of the unity argument is:

(15) I am the subject of all and only my own mental states.

which is surely a self-evident truth. The second premise is:

(16) Neither my body as a whole nor any part of it could be the subject of all and only my own mental states.

The conclusion is, once again, (8) I am not identical with my body nor with any part of it.

Of course, (16) is the crucial premise, so let us see how it might be defended. First, then, observe that my body *as a whole* does not need to exist in order for me to have *every one* of the mental states that I do in fact have. If, for instance, I were to lack the tip of one of my little fingers, I might as a consequence lack *some* of the mental states that I do in fact have, but surely not *all* of them. I might *perhaps* lack a certain mildly painful sensation in the finger tip—a sensation that I do in fact have—but many of my other mental states could surely be exactly the same as they actually are, such as the thoughts that I am in fact having in composing this essay. Indeed, I *could* still even have that sensation 'in my finger tip', because the phenomenon of 'phantom' pain is a well-attested one. However, I venture to affirm that no entity can qualify as the *subject* of certain mental states if those mental states could exist in the absence of that entity. After all, *I* certainly

qualify as the subject of *my* mental states, as (15) asserts, but for that very reason *those* mental states could not exist in my absence. Mental states must always *have* a subject—some being whose mental states they are—and the mental states that in fact belong to one subject could not have belonged to another, let alone to no subject at all.[16] But, as we have just seen, very many and quite possibly *all* of my own mental states could exist even if my body as a whole were not to exist—that is to say, even if certain parts that my body actually possesses were not to exist. This, I suggest, indicates that my body *as a whole* cannot qualify as the subject of all and only my own mental states and so cannot be identified with *me*. Now, many physicalists may agree with my reasoning so far, but draw the conclusion that, rather than being identical with my body *as a whole*, I am identical with some *part* of it, the most obvious candidate being *my brain*. However, it is easy to see that the foregoing reasoning can now just be repeated, replacing 'my body as a whole' by 'my brain as a whole' throughout. For it seems clear that, although I may well need to *have* a brain in order to have mental states, neither my brain as a whole nor any distinguished part of it is such that *it* in its entirety needs to exist in order for me to have *every one* of the mental states that I do in fact have. Indeed, even if every one of my mental states depends in this fashion upon *some* part of my brain, it by no means follows, of course, that there is some part of my brain upon which every one of my mental states thus depends. (To suppose that this does follow would be to commit a so-called 'quantifier-shift' fallacy.) And yet I, being the subject of all and only my own mental states, am such that every one of those mental states *does* depend upon *me*. Hence, we may conclude, neither my brain as a whole nor any part of it can qualify as the subject of all and only *my* mental states and so be identical with *me*. Putting together the two stages of this train of reasoning, we may thus infer that (16) is true and from that and (15) infer the truth of (8), the main claim of substance dualism.

I should perhaps stress that it is important to appreciate, when considering the foregoing argument, that I am by no means denying that there may be some part of my brain that is such that, *were it to be completely destroyed*, all of my mental states would thereby cease to be. After all, I am happy to concede that this may very well be true of my brain as a whole—that if *it* were to be completely destroyed, all of my mental states would thereby cease to be. All that I am denying, in effect, is that there is any part of my brain that is such that, were *any part of it*—such as one particular neuron—to be destroyed, all of my mental states would thereby cease to be. That is to say, neither my brain *as a whole*, nor any distinguished part of it *as a whole*, is something with which I can be identified—any more than I can be identified with *my body* as a whole—because no such entity is such that all and only *my* mental states can be taken to depend on it, in the way that they clearly do depend on *me*.

[16] See further Strawson (1959: chapter 3), and Lowe (1996: chapter 2).

However, here it may be objected that the foregoing defence of premise (16) depends upon an illicit assumption, namely, that if my body as a whole were to lack a certain part, such as the tip of one of my little fingers, then *it*—my body as a whole—would not exist. This assumption, it may be said, is unwarranted because it presupposes, questionably, that every part of my body is an *essential* part of it, without which it could not exist. As it stands, this may be a fair objection—although it should be acknowledged there are some philosophers who *do* hold that every part of a composite object is essential to it.[17] However, I think that the reasoning in favour of premise (16) can in fact be formulated slightly differently, so as to make it independent of the truth of this assumption. The initial insight still seems to be perfectly correct—that, as I put it, my body as a whole does not need to exist in order for me to have every one of the mental states that I do in fact have. Thus, to repeat, the thoughts that I am having in composing this essay plausibly do not depend upon my body including as a part the tip of one of my little fingers. Call these thoughts T. Consider, then, that object which consists of my body as a whole *minus* that finger tip. Call this object O and call my body as a whole B. (It should be conceded here that there are some philosophers who would deny that any such object as O exists[18]—but that is, to say the least, a controversial claim.) Suppose, now, that it is proposed that I am identical with B, and hence that B is the subject of the thoughts T. Then we can ask: on what grounds can B be regarded as the subject of T *in preference to* O, given that T do not depend upon B's including the part—the finger tip—that O does not include? Isn't the material difference between B and O simply *irrelevant* to the case that can be made in favour of either of them qualifying as the subject of T? But in that case, we must either say that *both* B *and* O are subjects of T, or else that *neither* of them are. We cannot say the former, however, because B and O are numerically distinct objects, whereas the thoughts T have just *one* subject—*myself*. We may conclude, hence, that neither B nor O is a subject of T and thus that I, who *am* the subject of T, am identical with neither of them. This sort of reasoning can then be repeated, as before, with respect to any specific *part* of B, such as my brain.

However we exactly formulate the defence of premise (16), the basic point of the unity argument, as I call it, is that my mental states do not all depend on my body as a whole or on any part of it in the unified way in which they all depend upon me as their subject. This point, it seems to me, is a good one. Indeed, between them, the unity argument and the replacement argument provide, I think, fairly compelling grounds for belief in the truth of non-Cartesian substance dualism.

[17] See, for example, Chisholm (1976: chapter 3).

[18] See, for example, van Inwagen (1981).

2. NCSD AND THE METAPHYSICS OF MENTAL CAUSATION

2.1. The Causal Closure Argument Against Interactive Dualism

In this second half of my essay, I want to explore certain *causal* considerations that inevitably arise in the debate between dualism and its opponents. For dualism—whether we are talking about substance dualism or property dualism—is traditionally divisible into interactionist, epiphenomenalist, and parallelist varities. Perhaps the most powerful argument against *interactive* dualism is the so-called *causal closure argument*.[19] By *interactive dualism* I mean the doctrine that mental events or states are not only *distinct* from physical events or states, but are also included amongst the *causes and effects* of physical events or states. Of course, the causal closure argument can have no force against either *epiphenomenalist* or *non-interactive parallelist* dualism, but since even the first and more credible of these positions has relatively few modern advocates, I shall not consider them here.[20] In any case, even those who do support them would presumably concede that they would prefer to endorse interactive dualism if they thought that it could meet the physicalist's objections, so let us concentrate on seeing how those objections can indeed be met, focusing on the causal closure argument.

The key premise of the causal closure argument against interactive dualism is the principle of the causal closure of the physical domain. This principle has received a number of different formulations—some of which are really too weak for the physicalist's purposes[21]—but the relatively strong version of the principle that I shall chiefly consider here is this:[22]

(17) No chain of event-causation can lead backwards from a purely physical effect to antecedent causes some of which are non-physical in character.

It may be objected on behalf of interactive dualism that (17) is simply question-begging, because it rules out by *fiat* the possibility of there being non-physical *mental* causes of some physical effects. However, as we shall see, (17) does not in fact rule out this possibility. Dialectically, it is in the dualist's interests to concede to the physicalist a version of the causal closure principle that is *as strong as possible*—provided that it still falls short of entailing the falsehood of interactive dualism—because if the causal closure argument in its strongest

[19] For further background, see Lowe (2000a: chapter 2).
[20] But see, for example, Robinson (2004).
[21] See further Lowe (2000b).
[22] Compare Kim (1993a).

non-question-begging form can be convincingly defeated, the physicalist will be left with no effective reply. Weaker versions of the causal closure principle can, of course, be countered by interactive dualists relatively easily, but tend to be countered by them in implausible ways which leave the physicalist with a telling response.

To illustrate the latter point, consider the following widely advocated version of the causal closure principle:

(18) Every physical effect of a mental cause has a sufficient physical cause.

An interactive dualist may accommodate (18) by, for example, espousing the doctrine of *interactive parallelism*, which maintains that there is a *one-to-one correlation* between the mental and physical causes of any physical event that has a mental cause, such that *both* the mental *and* the physical causes of any such event are *sufficient* causes of it.[23] (By a *sufficient* cause of a given event, I mean an event or conjunction of events that *causally necessitates* the event in question.) The physicalist may object that this doctrine has the highly implausible implication that every physical effect of a mental cause is *causally overdetermined* by that mental cause and the physical cause that is, suppposedly, one-to-one correlated with it. To this the interactive parallelist may reply that such causal overdetermination is no mere accident but, rather, the upshot of psychophysical laws, so that the fact that it occurs is a matter of nomic or natural necessity. However, it may nonetheless appear surprising to the impartial bystander that psychophysical laws of this character should be thought to govern the causal interactions of mind and body, when so many other possibilities are compatible with interactive dualism. The *non*-interactive parallelist has, it seems, much better reason to suppose that there is a one-to-one correlation between the *apparent* mental causes of physical events and their *actual* physical causes, because (traditionally, at least) they see this as being the upshot of a divinely instituted *pre-established harmony* between the mental and physical domains. Equally, the *physicalist* has a perfectly good reason to suppose that there is a one-to-one correlation between the mental and physical causes of physical events, because they *identify* those causes, and identity is a one-to-one relation *par excellence*. But the interactive parallelist, it seems, must simply regard it as a *brute fact* that psychophysical laws sustain such a one-to-one correlation—a fact that is all the more remarkable because so many other arrangements are consistent with the truth of interactive dualism. Neutral parties to the debate could be forgiven for suspecting that the interactive parallelist postulates the one-to-one correlation of mental and physical causes simply in mimicry of the physicalist's position, with a view to denying the physicalist recourse to any empirical evidence of a causal character that could discriminate between the two positions. For wherever the physicalist claims to find evidence of *one and the same cause* of a certain

[23] For an exposition and defence, see Meixner (2004: chapter 8).

physical effect—a single cause that is both mental and physical—the interactive parallelist will be able to reply that we in fact have *two distinct but correlated causes*, one of them mental and the other physical. However, this is a dangerous game for a dualist to play, because the physicalist can very plausibly urge that their *identity* theory provides a much more economical explanation of the one-to-one correlation of mental and physical causes that both they and the interactive parallelist believe to obtain.

Let us now consider a version of the causal closure argument against interactive dualism that appeals to the very strong formulation of the causal closure principle embodied in premise (17)—that no chain of event-causation can lead backwards from a purely physical effect to antecedent causes some of which are non-physical in character. Two additional premises are needed. First,

(19) Some purely physical effects have mental causes,

which the interactive dualist accepts as true, of course. Second,

(20) Any cause of a purely physical effect must belong to a chain of event-causation that leads backwards from that effect.

These three premises entail the conclusion,

(21) All of the mental causes of purely physical effects are themselves physical in character,

which contradicts the defining thesis of interactive dualism. My defence of interactive dualism will rest upon a challenge to premise (20). Moreover, it will endorse a version of interactive dualism that combines it with the sort of non-Cartesian substance dualism defended earlier.

2.2. Two Different Perspectives on the Causal Explanation of Voluntary Action

In order to keep matters relatively simple and to confine my discussion to manageable proportions, I shall concentrate, in what follows, on issues concerning voluntary and deliberative human action, where it is most obviously pressing that some coherent story needs to be told as to how mental and neurophysiological causes interrelate with one another. So let us focus on a specific case of such an action, such as an agent's deliberate (that is, premeditated and entirely voluntary) raising of an arm, for whatever reason (for instance, in order to catch a lecturer's attention with a view to asking a question). Now, what seems relatively uncontroversial, on the purely neurophysiological side of the causal story involved in such a case, is that if we were to trace the purely *bodily* causes of the relevant peripheral bodily event—in this case, the upward movement of the agent's arm on the given occasion—backwards in time indefinitely far, we would find that those causes *ramify*, like the branches of a tree, into a complex

maze of antecedent events in the agent's nervous system and brain—many of the neural events in the agent's brain being widely distributed across fairly large areas of the motor cortex and having no single focus anywhere, with the causal chains to which they belong possessing, moreover, no distinct *beginnings*.[24] And yet, intuitively, the agent's mental act of *decision* or *choice* to move the arm would seem, from an introspective point of view, to be a *singular* and *unitary* occurrence that somehow *initiated* his or her action of raising the arm. The immediate question, then, is how, if at all, we can reconcile these two apparent facts. It seems impossible to *identify* the agent's act of choice with any individual neural event, nor even with any combination of individual neural events, because it and they seem to have such different causal features or profiles. The act of choice seems to be unitary and to have, all by itself, an 'initiating' role, whereas the neural events seem to be thoroughly *disunified* and merely to contribute in different ways to a host of different ongoing causal chains, many of which lead independently of one another to the eventual arm-movement.

I believe that NCSD can enable us to see how *both* of these causal perspectives on deliberative physical action can be correct, without one being reducible to the other and without there existing any sort of rivalry between the two. First of all, the act of choice is attributable to the *person* whereas the neural events are attributable to parts of the person's *body*: and a person and his or her body are, according to this conception of ourselves, *distinct* things, even if they are not *separable* things. Moreover, the act of choice *causally explains* the bodily movement—the upward movement of the arm—in a different way from the way in which the neural events explain it. The neural events explain why the arm moved *in the particular way* that it did—at such-and-such a speed and in such-and-such a direction at a certain precise time. By contrast, the act of choice explains why a movement *of that general kind*—in this case, a rising of the agent's arm—occurred around about the time that it did. It did so because shortly beforehand the agent decided to raise that arm. The decision certainly did not determine the precise speed, direction, and timing of the arm's movement, only *that* a movement of that general sort would occur around about then. The difference between the two kinds of causal explanation reveals itself clearly, I suggest, when one contemplates their respective *counterfactual* implications. If the agent had not decided to raise his or her arm, there wouldn't have been an arm-movement of that kind *at all*—the arm would either have remained at rest or, if the agent had decided to make another movement instead, it would have moved in a quite different way. It doesn't seem, however, that one can isolate any neural event, or any set of neural events, whose non-occurrence would have had *exactly the same consequences* as the non-occurrence of the agent's decision. Rather, the most that one can say is that if this or that neural event, or set of neural events, had not occurred, the arm-movement might have proceeded in a

[24] See, e.g., Deecke, Scheid, and Kornhuber (1969) and Popper, and Eccles (1977: 282–94).

somewhat different manner—more jerkily, perhaps, or more quickly—*not* that the arm would have remained at rest, or would instead have moved in a quite different kind of way.

2.3. A Counterfactual Argument Against Psycho-Neural Causal Identity

This last point is an extremely important one and requires further elucidation. It is now standard practice amongst philosophers of logic and language to interpret counterfactual conditionals in terms of possible worlds, very roughly as follows.[25] A counterfactual of the form 'If it were the case that *p*, then it would be the case that *q*' is said to be true if and only if, in the *closest* possible world in which *p* is the case, *q* is also the case—where the 'closest' possible world in question is the one in which *p* is the case but otherwise *differs minimally* from the actual world. Now, suppose that a physicalist in the philosophy of mind were to propose that the agent's decision, *D*, to raise his or her arm on a given occasion—the agent's mental act of choice—is identical with a certain neural event, *N*, which is correctly identifiable as being a *cause* of the subsequent bodily event, *B*, of the arm's rising. Here I must stress that *D*, *N*, and *B* are, each of them, supposed to be *particular events*, each occurring at a particular moment of time, with *B* occurring at least an appreciable fraction of a second later than *D* and *N*, since our decisions to act do not take effect immediately—and the physicalist must suppose, of course, that *D* and *N* occur at the *same* time, since they hold them to be identical. And let me add, too, that I do not wish to get embroiled here in the debates concerning Benjamin Libet's celebrated but highly controversial experiments on the precise timing of volitions,[26] as this would sidetrack me from my present concerns. Let us concede, consequently, that the following counterfactual is true: 'If *N* had not occurred, then *B* would not have occurred.' All that I am presupposing here is that if *N* was indeed a cause of *B*, then the foregoing counterfactual is true. The physicalist cannot, I think, have any quarrel with me on this account. I am not taking any advantage, then, of the various reasons that have been advanced for doubting, at least in some cases, whether causal statements entail the corresponding counterfactuals.[27] What I am now interested in focusing on is the following question: what sort of event *would* have occurred, instead of *B*, if *N* had not occurred? In other words: in the closest possible world in which *N* does not occur, what sort of event occurs instead of *B*? My contention is that what occurs in this world is an event *of the same sort as B*, differing from *B* only very slightly. The reason for this is as follows.

[25] See, especially, Lewis (1973b), although I do not replicate every detail of his account, but only those that are germane to the issues now under discussion.

[26] Libet (1985). Note, in any case, that Libet's experiments were not concerned with *premeditated* actions, but only with 'spontaneous' ones.

[27] For discussion of this, see Lowe (2002a: chapter 10).

It seems evident, from what we know about the neural causes of an event such as *B*, that *N* must be an *immensely complex* neural event: it must be, in fact, the sum (or 'fusion') of a very large number of individual neural events, each of them consisting in some particular neuron's firing in a particular way. Recall, here, that *N* must be supposed to occur an appreciable amount of time *before* *B*, at a time at which the neural antecedents of *B* are many and quite widely distributed across the agent's cerebral cortex. It would be utterly implausible for the physicalist to maintain, for example, that the agent's decision *D* is identical with the firing of just a *single* neuron, or even of a small number of neurons. If *D* is identical with any neural event *at all*, it can surely only be identical with an extremely complex one, consisting in the firing of many neurons distributed over quite a large region of the agent's cerebral cortex. However, it seems indisputable that if *N* is, thus, the sum of a very large number of individual neural events, then the *closest* world in which *N* itself does not occur is a world in which *another* highly complex neural event, *N**, occurs, differing *only very slightly* from *N* in respect of the individual neural events of which it is the sum. In other words, *N** will consist of *almost exactly the same* individual neural events as *N*, plus or minus one or two. Any possible world in which a neural event occurs that differs from *N* in *more* than this minimal way simply will not qualify as the *closest* possible world in which *N* does not occur. This is evidently what the standard semantics for counterfactuals requires us to say in this case. But, given what we know about the functioning of the brain and nervous system, it seems clear that, in the possible world in which *N** occurs, it causes a bodily event *very similar* to *B*, because such a small difference between *N* and *N** in respect of the individual neural events of which they are respectively the sums cannot be expected to make a very big difference between their bodily effects. There is, we know, a good deal of redundancy in the functioning of neural systems, so that the failure to fire of one or two motor neurons, or the abnormal firing of one or two others, will typically make at most only a minimal difference with regard to the peripheral bodily behaviour that ensues. Thus, the answer to the question posed earlier—what sort of bodily event would have occurred instead of *B*, if *N* had not occurred?—is this: a bodily event *very similar to B*. In other words, if *N* had not occurred, *the agent's arm would still have risen in almost exactly the same way as it actually did*.

Now, I hope, we can see the importance of this conclusion. For, if we ask what sort of bodily event would have occurred instead of *B* if *the agent's decision*, *D*, to raise his or her arm had not occurred, then we plausibly get a very different answer. Very plausibly, if *D* had not occurred—if the agent had not made the very act of choice that he or she did to raise the arm—then the arm *would not have risen at all*. It is, I suggest, quite incredible to suppose that if the agent had not made *that* very decision, *D*, then he or she would have made another decision virtually indistinguishable from *D*—in other words, *another* decision to raise the arm in the same, or virtually the same, way. On the contrary, if the

agent had not made *that* decision, then he or she would either have made a quite different decision or else no decision at all. Either way—assuming that there is nothing defective in the agent's nervous system—the arm *would not* have risen almost exactly as it did.

I suppose that a convinced physicalist might try to challenge the claim that I have just made and contend that, indeed, if *D* had not occurred, then *another* decision to raise the arm in virtually the same way would have occurred instead, giving rise to a slightly different bodily event of the same kind.[28] But, on the face of it, this would appear to be a purely *ad hoc* maneuver designed solely to save the envisaged physicalist's position. One serious problem with it is that contentful mental acts such as decisions are, very plausibly, *individuated* at least partly by their contents—and yet their contents surely cannot be as fine-grained as the physicalist's conjectured contention would appear to demand. How, exactly, would the *content* of the decision that, supposedly, would have occurred if *D* had not occurred, have differed from the content of *D*? If the putative difference in their contents is to match the *very slight* difference between the bodily events that are supposed to ensue from them, then a degree of *fine-grainedness* must be attributed to those contents that, it seems to me, is utterly implausible from a psychological point of view. For instance, we must suppose that *D* is a decision to raise the agent's arm along a quite specific trajectory *T*, whereas if *D* had not occurred then the agent would instead have decided to raise his or her arm along the very slightly different trajectory *T**, where the spatiotemporal differences between *T* and *T** are of the same order of magnitude as the very slight differences between the *actual* arm-movement *B* and the arm-movement that would have occurred if neural event *N** had occurred instead of neural event *N*. But the contents of our decisions to act are surely *never* as fine-grained as this—not, at least, if our conscious introspective awareness of those contents is to be relied upon. And to propose that they always have much finer-grained contents that are *inaccessible* to consciousness seems a desperate recourse on the part of the physicalist. When, for instance, I decide to raise my arm in a lecture in order to ask a question, I may indeed decide to raise it *quickly* and *vertically*, but never—surely—along a quite specific trajectory at a quite specific speed. Quite apart from anything else, I simply don't possess sufficient *voluntary control* over my limb-movements to be able to decide to execute them with such precision.

If all of this reasoning is correct, then it follows unavoidably that the decision *D* cannot be identical with the neural event *N* with which the physicalist proposes to identify it, for the counterfactual implications of the non-occurrence of these two events are quite different. If *D* had not occurred, the agent's arm would not have risen at all, but if *N* had not occurred, it would have risen almost exactly as it did. One fundamental reason for this—according to the conception of human persons that I favour as an advocate of NCSD—is that a mental act of choice

[28] I am grateful to José Bermúdez for pressing this line of response.

or decision is, in a strong sense, a *singular* and *unitary* event, unlike a highly complex sum or fusion of independent neural events, such as *N*. *N** differs from *N* only in excluding one or two of the individual neural events composing *N* or including one or two others. That is why *N* and *N** can be so similar and thus have such similar effects. But *D*, I suggest, cannot intelligibly be thought of, in like manner, as being *composed* of myriads of little events—and that is at least partly why, in the closest possible world in which *D* itself does not occur, there does *not* occur another decision *D** which differs from *D* as little as *N** differs from *N*. Note that this strong *unity* of our mental acts, whereby they resist decomposition into lesser parts, nicely parallels our own strong unity as 'simple' substances, revealed by the unity argument of section 1.5 above.

I should perhaps add that, although I do not have space enough to demonstrate this in detail here, the foregoing line of argument sustains not only the conclusion that the mental and neural causes of voluntary bodily movements must be numerically *distinct*, but also the stronger conclusion that those mental causes cannot even be taken to be 'realized by' any of those neural causes—where 'realization' is taken to be a relation distinct from identity itself, in virtue of which realized events or states inherit their causal features entirely from those of the events or states that realize them.

2.4. Intentional Causation Versus Physical Causation

So far, I have tried to explain why the mental and neural causes of voluntary bodily movements must be distinct, consistently with allowing, as I do, that such movements have *both* mental *and* neural causes. Now I want to say a little more about the respects in which mental causation is distinctively different from bodily or physical causation. Most importantly, then, mental causation is *intentional* causation—it is the causation of an *intended* effect *of a certain kind*. Bodily causation is not like this. All physical causation is 'blind,' in the sense that physical causes are not 'directed towards' their effects in the way that mental causes are. *Both* sorts of causation need to be invoked, I believe, in order to give a full explanation of human action and NCSD's conception of human persons seems best equipped to accommodate this fact. The very *logic* of intentional causation differs, I venture to say, from the *logic* of bodily causation. Intentional causation is *fact* causation, while bodily causation is *event* causation.[29] That is to say, a choice or decision to move one's body in a certain way is causally responsible for the *fact* that a bodily movement *of a certain kind* occurs, whereas a neural event, or set of neural events, is causally responsible for a *particular* bodily movement, which is a particular *event*. The decision, unlike the neural event, doesn't causally explain why that *particular* bodily movement occurs, not least because one cannot *intend* to bring about what one cannot *voluntarily*

[29] For more on this distinction, see Bennett (1988) and also Lowe (2002a: chapter 9).

control—for, as I pointed our earlier, one cannot voluntarily control the *precise* bodily movement that occurs when one decides, say, to raise one's arm.

As I have just implied, the two species of causal explanation, mental and physical, are both required and are mutually complementary, for the following reason. Merely to know why a *particular* event of a certain kind occurred is not necessarily yet to know why an event of *that* kind occurred, as opposed to an event of some other kind. Intentional causation can provide the latter type of explanation in cases in which bodily causation cannot. More specifically: an event, such as a particular bodily movement, which may appear to be merely *coincidental* from a purely *physiological* point of view—inasmuch as it is the upshot of a host of independent neural events preceding it—will by no means appear to be merely coincidental from an *intentional* point of view, since it was an event *of a kind* that the agent intended to produce.[30]

Notice, here, that the aforementioned fact—that a mental decision, *D*, to perform a certain kind of bodily movement, cannot be said to cause the *particular* bodily event, *B*, of that kind whose occurrence renders that decision successful—is already implied by the argument that I developed a little while ago in section 2.3. For, given that *D* is *not identical* with the actual neural cause, *N*, of *B*, the closest possible world in which *N* does not occur *is still a world in which D occurs*—but in that world a slightly different bodily movement, B^*, ensues, being caused there by a slightly different neural cause, N^*. (Clearly, if *D* is not identical with *N*, then there is no reason to suppose that the closest world in which *N* does not occur is also one in which *D* does not occur, for a world in which *both* of these events do not occur evidently differs more from the actual world than a world in which just *one* of them does not occur, other things being equal.) However, this means that the occurrence of *D* is causally compatible with the occurrence of two *numerically different* bodily movements of the same kind, *B* and B^*, and hence does not causally determine *which* of these occurs, but only that *some* bodily movement of their kind occurs.

2.5. Reasons, Causes, and Freedom of Action

Much more can and should be said on these matters, but since I have discussed many of them extensively elsewhere,[31] I shall rest content with the foregoing remarks for present purposes. Here, however, it may be asked: *But what about the causes of an agent's acts of decision or choice?* Are *these* bodily or mental, or both? My own opinion is that an act of decision or choice is *free*, in the 'libertarian' sense—that is to say, it is *uncaused*.[32] This is not to say that decisions are simply *inexplicable*, only that they demand explanations of a non-causal sort. Decisions are explicable in terms of *reasons*, not causes. That is to say, if we want to know

[30] See further Lowe (1999).
[31] See again, in particular, Lowe (1999).
[32] See further Lowe (2003a).

why an agent *decided* to act as they did, we need to inquire into *the reasons in the light of which* they chose so to act.[33] Since decisions are, according to NCSD's conception of the mind, attributable to the person and not to the person's body or any part of it, there is no implication here that any *bodily* event is uncaused. It's not that I want to exclude altogether the idea of causal explanation in terms of mental states in favour of purely rational explanation in the psychological sphere—as my earlier arguments make manifest. However, I do want to help to reinstate the idea that reason-giving explanation is *not* a species of causal explanation and that it is one form of explanation that is distinctive of the psychological sphere.

But now it may be wondered: how is it really possible for mental acts of decision to explain anything in the physical domain, if that domain is *causally closed*, as many contemporary philosophers of mind—and just about all physicalists—assume? This takes us back to the earlier concerns of section 2.1 above. As we observed there, much turns on precisely how the putative causal closure of the physical domain is to be defined, for this is no simple matter.[34] According to one popular view,[35] the thesis of physical causal closure amounts to the claim that no chain of event-causation can lead backwards from a purely physical effect to antecedent causes some of which are *non*-physical in character. This was premise (17) of the version of the causal closure argument presented in section 2.1. But intentional causation according to NCSD's conception of human persons, as I have tried to characterize it earlier, does not violate the thesis of physical causal closure just stated, since it does not postulate that mental acts of decision or choice are events *mediating between bodily events* in chains of causation leading to purely physical effects: it does not postulate that there are 'gaps' in chains of physical causation that are 'filled' by mental events. Thus, NCSD's model of mental causation is consistent with premise (17) of the causal closure argument and avoids the conclusion of that argument by repudiating, instead, premise (20).

As we have seen, according to NCSD's conception of human persons, a decision can explain the fact that a bodily movement *of a certain kind* occurred on a given occasion, but not the *particular* movement that occurred. Even so, it may be protested that if physical causation is *deterministic*, then there is really no scope for intentional causation on the model that I am defending to explain anything physical, because the relevant counterfactuals will all simply be *false*. It will be *false*, for instance, to say that if the agent had not decided to raise his or her arm, then a rising of the agent's arm would not have occurred: rather, precisely the same bodily movement *would* still have occurred, caused by precisely the same physical events that actually did cause it—for if physical determinism is true, there was never any real possibility that those physical events should not

[33] Compare Dancy (2000). [34] See Lowe (2000).
[35] Endorsed, for example, by Kim (1993a).

have occurred, nor that they should have had different effects. Maybe so. But, in view of the developments in quantum physics during the twentieth century, we now know that physical causation is *not* in fact deterministic, so the objection is an idle one and can safely be ignored. The model of intentional causation that I am proposing may nonetheless still seem puzzling to many philosophers, but if so then I suggest that this will be because they are still in the grip of an unduly simple conception of what causation involves—one which admits only of the causation of one event by one or more antecedent events belonging to one or more chains of causation which stretch back indefinitely far in time. Since this seems to be the only sort of causation that is recognized by the physical sciences, intentional causation on NCSD's model is bound to be *invisible* from the perspective of such a science.[36] To a physicalist, this invisibility will seem like a reason to dismiss NCSD's conception of intentional causation as spurious, because 'non-scientific.' I hope that to more open-minded philosophers it will seem more like a reason to perceive no genuine conflict between explanation in the physical and biological sciences and another, more humanistic way of explaining our intentional actions, by reference to our choices or decisions and the reasons for which we make them.

[36] Compare Lowe (2003b).