

TWO PROBLEMS IN SPINOZA'S THEORY OF MIND

James Van Cleve

Forthcoming in *Oxford Studies in Philosophy of Mind*

My aim in what follows is to expound and (if possible) resolve two problems in Spinoza's theory of mind. The first problem is how Spinoza can accept a key premise in Descartes's argument for dualism—that thought and extension are separately conceivable, “one without the help of the other”—without accepting Descartes's conclusion that no substance is both thinking and extended. Resolving this problem will require us to consider a crucial ambiguity in the notion of conceiving one thing without another, the credentials of Descartes's principle that each substance has a principal attribute, and the prospect of neutral monism as a theory of the mind. The second problem is how Spinoza can maintain that each mental event is identical with some physical event (and conversely) while denying that there is ever any causal interaction between mental and physical events. If one mental event causes another and the first is identical with some physical event, must that physical event not cause the mental event? Resolving this problem will require looking into the reach of opacity in Spinoza's philosophy and considering whether there are ever any legitimate exceptions to Leibniz's Law.

0. Background

I begin with a taxonomy of theories on the mind-body relation and a brief discussion of where Spinoza fits in. The scheme was originally laid out by C.D. Broad (1925, 607-11) and subsequently elaborated by David Chalmers (2003).

Broad defines a *differentiating attribute* (or *attribute* for short) as a property that is both (i) exemplified in the universe and (ii) fundamental, that is, not reducible to or derivable from any other type of property. We then get a main divide into four types of theory, depending on whether mentality and materiality are both attributes, mentality an attribute and materiality not, materiality an attribute and mentality not, or neither of them an attribute. *Dualism* is the view that mentality and materiality are both attributes. It subdivides into dualism of compatibles, according to which mentality and materiality may belong to the same substance (as with Spinoza) and dualism of incompatibles, according to which they cannot (as with Descartes). *Mentalism* is the view that mentality is an attribute and materiality not. If materiality is not an attribute because it is not exemplified in the universe at all, we get eliminative mentalism (as with Berkeley). If materiality is not an attribute because it is reducible to mentality, we get reductive mentalism (as in Leibniz and other phenomenologists). *Materialism* is the view that materiality is an attribute and mentality not. If mentality is not an attribute because it is not exemplified in the universe at all, we get eliminative materialism (as perhaps in Rorty 1965 or Churchland 1988; Broad could think of no advocate in his day). If mentality is not an attribute because it is somehow reducible to materiality, we get reductive materialism, with subdivisions depending on the material basis (behavior? brain states?) and the mode of reduction (analytic definition? identity? supervenience?). *Neutralism* is the view that neither materiality nor mentality is an attribute. The most interesting variety for our purposes is *neutral monism*, the view that materiality and mentality are both exemplified in the universe, but neither is fundamental, each deriving from some third type of thing or property that is neither mental nor physical. Views fitting this description were held in Broad's day by William James, Bertrand Russell, and Samuel Alexander; they are getting a second look in our own day under the rubric

‘Russellian monism’ (Alter and Nagasawa 2015). Finally, in the Chalmers elaboration of Broad, we get further subdivisions among any of the noneliminative categories by asking whether causal relations between the mental and the physical run in one direction, the other, both, or neither.

I list now the main tenets of Spinoza’s theory of mind as it relates to this scheme. The properties he calls ‘attributes’ are supremely general essential properties, making his usage consonant with Broad’s of ‘differentiating attribute’.

Dualism of Attributes. Thought (Spinoza’s all-purpose word for mentality) and extension (the core feature of materiality) are distinct and mutually irreducible attributes. As Jonathan Bennett puts it, no logically true or conceptually necessary conditional connects the mental and the physical (1984, 41, 47, and 136-37). Moreover, both attributes are exemplified in the universe. This makes Spinoza a dualist in Broad’s sense. Textual basis: 2p1, 2p2, 1p10.¹

Monism of Substance. For Spinoza, there is only one substance in the universe, so obviously he cannot follow Descartes in holding that mind and body are distinct *substances*. His dualism of attributes must be what Broad calls a dualism of compatibles. Textual basis: 1p14.

Denial of Interaction. Contrary to what Descartes taught, there are no causal relations between the mental and the physical. This is a consequence of attribute dualism plus Spinoza’s causal rationalism, his view that the connection between causes and their effects is broadly logical, a species of entailment. Textual basis: 2p6, 3p2, 5 preface. When I speak henceforth of interaction, I mean trans-attribute causation in either direction.

¹ I use the following scheme for referring to propositions in Spinoza’s *Ethics*: the initial numeral gives part number; ‘p’ stands for ‘proposition’, ‘d’ for definition, ‘s’ for scholium, and ‘d’ after a proposition number for demonstration. Thus ‘1p10’ means proposition 10 of Part 1 of the *Ethics*, and ‘1p10s’ means the scholium to that proposition. My translations are either those of Shirley or Curley—Shirley’s unless indicated by a ‘C’.

Psychophysical Isomorphism. This is asserted in the famous 2p7, “The order and connection of ideas is the same as the order and connection of things.” ‘Ideas’ and ‘things’ cover respectively (i) representations and their objects, (ii) minds and bodies, and (iii) mental events and physical events, potentially giving us three different doctrines.² We may for present purposes ignore the first, concerning representations and their objects.³

To arrive at a more exact formulation of Spinozistic isomorphism, let us begin with the following “Principle of Simultaneous Isomorphism,” adapted from Brandt and Kim 1967 (521):

For every mental property M, there is a physical property P such that for any individual x and time t, x has M at t iff x has P at t.

For short, every mental state has a physical correlate. Spinoza means to assert not only this, but its converse or mirror image, which goes beyond Brandt and Kim: for every physical property P, there is a mental property M such that for any individual x and time t, x has P at t iff x has M at t—every physical state has a mental correlate. Given that everything in the universe has some physical properties, that makes Spinoza a *panpsychist*—someone who holds that everything in the universe has mental properties.⁴

Spinoza's isomorphism involves one more ingredient: whenever there is a causal chain linking items in one of the realms (mental or physical), a parallel causal chain links the corresponding items in the other realm (Bennett, 127).

² Bennett calls attention to Spinoza's “conceptual minimalism”—his penchant for trying to state all the doctrines of his philosophy using just a small number of basic concepts (1984, 38-41). That leads to the packing of two or more doctrines into the same verbal clothing in 2p7. See Bennett, 127-31, for more on the content and basis of Spinoza's isomorphism, a.k.a. parallelism.

³ Spinoza sometimes identifies the human mind with the idea of the human body, as in 2p12 and 13 (and as discussed in Wilson 1999). That would collapse together idea-object isomorphism and mind-body isomorphism. On the other hand, perhaps he has an autonomous doctrine of isomorphism between ideas and objects. I need not get into that issue here; I will be concerned only with the isomorphisms between minds and bodies and between mental events and physical events.

⁴ See Bennett, 135-37, for an interesting reason Spinoza might have had for affirming panpsychism—that if physical configurations at one level of complexity were accompanied by mentality and others not, there could be no explanation for this, contrary to the demands of the Principle of Sufficient Reason.

What would explain such a remarkable isomorphism? There are three main possibilities: (1) divine intervention or pre-ordination (as in Malebranche and Leibniz respectively); (2) causal laws connecting the mental and the physical (as in Descartes); and (3) *identity* between corresponding items in the two realms. Spinoza rejects the first two and opts for the third. And given the double duty done by the term ‘idea’ (as well as by the term ‘mode’), his identity thesis is two pronged: minds are identical with bodies and mental events with physical events, as set forth in the next two italicized entries.

Identity of Mind and Body. Just as the thinking substance and the extended substance are one substance comprehended under two attributes, Spinoza tells us, “so also a mode of extension and the idea of that mode [read: the mind associated with that mode] are one and the same thing, but expressed in two ways” (2p7s, C). For further textual basis, see 2p21s, and 3p2s.⁵

Given only the identity of mind and body, it would be tempting to compare Spinoza's view to that of Strawson or Chalmers: a person is a unitary subject of two different sorts of events, mental and physical. But Spinoza goes further, asserting as well the

Identity of Mental Events and Physical Events. Here see 3p2s: a mental decision and the corresponding physical state are *the same thing*, considered under two different attributes. This identity explains the isomorphism that would otherwise be a vast brute fact.⁶

The reader doubtless has questions about how all these tenets can be true together and how some of them can be true at all. I hope to address some of these questions in what follows.

Part One

1. De Vries's Question

⁵ 3p2s says that 2p7s “tells us that mind and body are one and the same thing, conceived now under the attribute of Thought, now under the attribute of Extension.”

⁶ See Bennett, 139-43 and 149-50.

According to 1p10 of the *Ethics*, “Each attribute of one substance must be conceived through itself.” The main import of “conceived through itself” in Spinoza is negative; it means *not* conceived through something else (though, of course, still conceived). The scholium adds this:

From this it is clear that although two attributes be conceived as really distinct, that is, one without the help of the other, still we cannot deduce therefrom that they constitute two entities, or two different substances It is by no means absurd to ascribe more than one attribute to one substance. (1p10s)

Here Spinoza is asserting the first of the tenets above, his dualism of attributes. He is also affirming that his dualism is a dualism of compatibles—there is nothing in it to keep him from affirming monism about substance. He does not mention Descartes, but he is quietly, yet clearly, taking issue with his predecessor: we do not need two different sorts of substances to be the bearers of the distinct attributes of thought and extension.

Writing to Spinoza in February of 1663, Spinoza’s disciple Simon De Vries asked, in effect, how things can be that way.⁷ Spinoza admits the key premise in Descartes’s argument for substance dualism: that thinking may be conceived of apart from extension or any other bodily trait. How, then, can Spinoza resist the Cartesian conclusion that thinking substances and extended substances are distinct? Why is it not correct to conclude, as De Vries puts it, “that, where there are two different attributes, there are two different substances”? (Curley, 193).

Alan Donagan offers an answer on Spinoza’s behalf (1973). He points out that anyone who accepts the idea of God as the *ens realissimum* (as Descartes and Spinoza both do) must also accept the possibility that thought and extension characterize the same substance.⁸ An *ens realissimum* is an infinite being that is maximally real, brimming over with every attribute it can

⁷ DeVries was responding to an early version of the *Ethics* in which the passage calling forth his question occurred as a scholium to prop. 8 rather than, as in the published version, to prop. 10. He might have raised his question equally well to the published version.

⁸ There is a hint of this reply in Letter 9, Spinoza’s reply to De Vries of March, 1663, in Curley, 195. Of course, Descartes might not agree that an *ens realissimum* must be extended; he might claim the opposite.

have compossibly with its other attributes. In my view, this point may possibly serve as an argument *ad hominem* against Descartes—for he, too, believed that God was the most real being—but it does nothing to answer De Vries’s question. Granted that Spinoza has reason to affirm that the same substance may be thinking and extended, how can he do so consistently with his acceptance of the Cartesian premise? For all Donagan has shown, Spinoza might simply have an inconsistent philosophy. The general point is that if you are asked “How can you accept P and reject Q, given that P implies Q?,” it is not enough to give whatever independent reasons you have for rejecting Q.⁹

It could be, of course, that Spinoza is simply inconsistent, having reason to accept that a single substance can have both thought and extension as well as reason to reject it. Before acquiescing in that verdict, however, I would like to explore three alternative answers to De Vries’s question. According to the first, Spinoza does not really accept the starting point of Descartes’s argument. According to the second and the third, he does accept the starting point, but the starting point by itself does not entail the conclusion; it does so only in conjunction with further premises that Spinoza is free to reject.

2. What is Separate Conceivability?

What exactly *is* the starting point of Descartes’s argument for the conclusion that mind and body are distinct substances—that each of us is a purely thinking substance that lacks extension and all traits of body?

Here are a few relevant lines from paragraph 9 of Meditation 6:

⁹ Donagan also offers another solution drawing on the Spinozistic notion of *expression*, which he says has the following formal property: “two attributes might on the one hand be really distinct, and on the other constitute or express the same essence” (180). But I think this solution responds to a problem about the attributes that is different from the De Vries problem Donagan started with. It is a response to the following argument: a substance with two attributes has two essences; a substance can have only one essence; therefore, no substance has two attributes. Donagan has Spinoza denying the first premise. As far as I can see, however, nothing in Donagan shows what premise Spinoza would deny in the reconstruction of the Cartesian argument for substance dualism Donagan gives on pp. 166-67.

I know that everything which I clearly and distinctly understand is capable of being created by God so as to correspond exactly with my understanding of it. Hence, the fact that I can clearly and distinctly understand one thing apart from another is enough to make me certain that the two things are distinct, since they are capable of being separated, at least by God. . . . I may have . . . a body that is very closely joined to me. But, nevertheless, on the one hand I have a clear and distinct idea of myself, in so far as I am simply a thinking, non-extended thing; and on the other hand I have a distinct idea of body, in so far as this is simply an extended, non-thinking thing. And accordingly, it is certain that I am really distinct from my body, and can exist without it. (CSM 2, 54)¹⁰

Let me use ‘conceive’ as short for ‘clearly and distinctly understand’ and other similar phrases in Descartes. Then Descartes is saying that if he can conceive of one thing without another, the one can exist without the other and is therefore distinct from it.

But what is meant by “conceiving A without B”? As I have pointed out in earlier work, the phrase is ambiguous (Van Cleve 1983). It can mean either

(1) I can conceive of A without conceiving of B. In symbols: $C(A) \ \& \ \sim C(B)$.

or

(2) I can conceive of the combination A-without-B. In symbols: $C(A \ \& \ \sim B)$.

In his *Objections to Descartes*, Arnauld took Descartes to be arguing from a premise of form (1), perhaps ‘I can conceive that I exist without conceiving that my body exists.’ He then objected that from this premise, it does not follow that I could exist even though my body did not. He offered this counterexample: someone might clearly and distinctly perceive that a given triangle is a right triangle (perhaps because he sees that it is inscribed in a semicircle) without clearly and distinctly perceiving that its hypotenuse is equal to the sum of the squares of its other

¹⁰ I have deliberately omitted from the middle of the paragraph a sentence that commentators have found ambiguous—between ‘I do not know that anything pertains to my essence but thinking’ and ‘I know that nothing pertains to my essence but thinking.’ The first may be regarded as too weak for Descartes’s purposes and the second as begging the question. See Dicker 1993, 188-93, for a good discussion of the problem. Descartes’s argument as I present it here need not make use of either version of the premise I have elided.

two sides; but this surely would not show that a right triangle might fail to obey the Pythagorean theorem. (CSM 2, 141-142).

To this Descartes made three replies, of which the third is the most illuminating. Here it is:

True, that triangle may indeed be apprehended although there is no thought of the ratio prevailing between the squares on the base and sides; but we can never think that this ratio must be denied. It is quite otherwise in the case of the mind where, not only do we understand that it exists apart from the body, but also that all the attributes of body may be denied of it. (Descartes 1911, 2, 102; the CSM translation is at 2, 159.)

Descartes is saying that Arnauld has misconstrued his starting point. It is actually of form (2), not form (1). Arnauld's counterexample is a counterexample to a possibility conclusion drawn from (1), but Descartes's argument is an argument to a possibility conclusion drawn from (2). To that, Descartes says, there is no similar counterexample.¹¹

Elaborating, (1) might be true of someone who is ignorant of the Pythagorean Theorem or who is not currently giving it any thought, but it would not follow that a triangle could be a right triangle without being Pythagorean. To draw that conclusion, we would need (2), but we do not have it; no one clearly conceives of any triangle that it is right yet not Pythagorean. But we do have (2) in the case of mind and body, Descartes says; I can clearly conceive of my mind's existing without any body's existing.

Several other passages reinforce my interpretation. In a letter to Mesland (May 2, 1644), Descartes distinguishes abstraction (representing the soul without representing the body) from exclusion (representing the soul as something that can exist without the body). He says his argument relies on exclusion, not abstraction (CSM 3, 236, cited by Rozemond in n. 24 to ch. 1). The notion of exclusion reappears the same year in Principle I.60 of the *Principles of Philosophy*, where Descartes says we can understand ourselves as thinking while "excluding" (not merely failing to include) any corporeal traits (CSM 1, 213). In a letter to Gibieuf (Jan. 19,

¹¹ I have explained all this at greater length in Van Cleve 1983.

1642, CSM 3, 201-202; cited by Rozemond, 16) he says “although one can think of one without paying attention to the other, one cannot deny it of the other when one thinks of both.” For example, I can think of shape without thinking of extension, but I cannot conceive of shape while denying extension of it. In my notation, we can have $C(A) \ \& \ \sim C(B)$ when A is shape and B is extension, but not $C(A \ \& \ \sim B)$. In the mind-body case, by contrast, we *do* have $C(A \ \& \ \sim B)$.¹²

Pattern (1), $C(A) \ \& \ \sim C(B)$, might be called the double-occurrence pattern, since a cognitive verb occurs in it twice; pattern (2), $C(A \ \& \ \sim B)$, might be called the single-occurrence pattern, since the cognitive verb occurs only once. Descartes tells us in the passages I have noted that in the proper understanding of his argument, it proceeds from the single-occurrence pattern. So the following question becomes crucial: when Spinoza says that thought and extension are conceivable “one without the help of the other,” does he mean that we can conceive of thought without conceiving of extension (double occurrence) or that we can conceive of thought-without-extension (single occurrence)? If he affirms only the former, he does not accept the starting point of Descartes’s argument after all, contrary to what is assumed by De Vries and Donagan.

To answer our question, we need to pay close attention to the various passages in which Spinoza says that one thing is or is not conceivable, knowable, or understandable without the help of another—Spinoza tends to use these cognitive verbs interchangeably. I survey a few of these passages below. Preview: we will find instances of both the single-occurrence pattern and the double-occurrence pattern; Spinoza may not have been attuned to the distinction.

¹² These are all confirming passages I have discovered since writing my 1983 paper. There is yet another in the Sixth Replies, where Descartes says that there are cases in which we can understand one thing without thinking of the other, yet when we think of both, not see how one can exist without the other. “For example, we can understand the immeasurable greatness of God even though we do not attend to his justice; but if we attend to both, it is quite self-contradictory to suppose that he is immeasurably great and yet not just” (CSM 2, 298-299).

In 1d3, Spinoza defines *substance* as “that which is in itself and is conceived through itself.” He unpacks the second clause, “conceived through itself,” by saying, “that is, the conception of which does not require the conception of another thing from which it has to be formed.” That is an instance of the double-occurrence pattern.

In the paraphrase of 1d3 at 1p8s2, Spinoza says substance is “that the knowledge of which does not require the knowledge of any other thing,” whereas modifications are “those things whose concept is formed from the concept of the thing in which they are.” There again we have the double-occurrence pattern.¹³

In the important axiom 1a4, Spinoza says “the knowledge of an effect depends on, and involves, the knowledge of the cause.” That is unequivocally the double-occurrence pattern. The same is true of the next axiom, 1a5: “Things which have nothing in common with each other cannot be understood through each other; that is, the conception of the one does not involve the conception of the other.”

I turn now to a number of passages that are either explicitly of the single-occurrence pattern or are best construed that way.

According to 1a7, “If a thing can be conceived as not existing, its essence does not involve existence.” This formulation of the axiom is noteworthy for what it does *not* say. It does not say, “if we can conceive of x without conceiving of x as existing, its essence does not involve existence.” If it did, it would be open to a counterexample like Arnauld’s to Descartes, such as this one: a child can conceive of the number five without conceiving of it as being prime, but it

¹³ The translations in this paragraph are Curley’s. Shirley translates the substance clause the same way as Curley, but for the modifications clause he has “that which is in another thing, and whose conception is formed from the thing in which they are”—an instance of the single-occurrence pattern. I suspect, though, that Shirley was just being elliptical in his translation.

does not follow that five is not essentially prime. The axiom is much more plausible if read in the single-occurrence way, which is also how it is written.

In 2d2, Spinoza defines the *essence* of a thing (in part) as “that without which the thing can neither be nor be conceived.” Once again, the formulation is noteworthy for what it does not say; it does not say that the essence of a thing is that without *conceiving* which the thing cannot be conceived; the verb ‘conceive’ occurs once, not twice. Moreover, there is a liability Spinoza would incur if the definition were put forth in the double-occurrence way, ‘for any attribute A of a substance S, we cannot conceive of S without conceiving of A.’ In that case, there could be no unknown attributes of substance; yet Spinoza is generally interpreted as holding that substance has infinitely many attributes beyond the two that we are acquainted with, thought and extension. So this consideration favors a reading of ‘X cannot be conceived without Y’ as $\sim C(X \ \& \ \sim Y)$ and correlatively of ‘X can be conceived of without Y’ as $C(X \ \& \ \sim Y)$, which is the single-occurrence pattern.

Finally, let us revisit 1a4, which says the knowledge of an effect depends and involves the knowledge of the cause. I cited this above as an explicit case of the double-occurrence pattern, and so it is. Yet so taken it is immensely implausible, since it implies a kind of omniscience on our part with respect to causes: we cannot have knowledge of anything without having knowledge of its cause, and therefore, given Spinoza’s determinism, we cannot have knowledge of anything without having knowledge of its entire infinite causal ancestry. Suppose we took 1a4 instead as single: if c causes e, then if we consider e and c together, we cannot conceive (or see the possibility of) e happening without c. That is still implausible by my lights, but not *as* implausible; it is simply part and parcel of Spinoza’s generally acknowledged causal rationalism,

according to which causal relations hold with metaphysical necessity. Letting plausibility trump grammar, we would then have one more entry on the single-occurrence side.

On the other hand, if 1a4 is to play its intended role in the proof of Spinoza's isomorphism, we probably need the double-occurrence pattern! See Bennett, section 31, especially p. 129.

So where do we come down? As far as I have been able to discover, there is no systematic explanation of why Spinoza sometimes uses the single-occurrence pattern and sometimes the double. I suspect that he was simply not sensitive to the difference. Which pattern *ought* he to have favored, all things considered? I do not know the answer to that question, either. Some aspects of his philosophy are better served by one and others by the other, as we saw in the case of 1a4.

Spinoza is not alone among early modern philosophers in his disregard of the single vs. double distinction. Hume's Conceivability Principle, used throughout Book I of the *Treatise of Human Nature*, says that things separable in the imagination are also separable in reality. What is separation in the imagination—imagining one thing without imagining the other, or imagining the complex state of affairs in which one exists and the other does not? Hume passes back and forth indifferently between the two, and he sometime uses one when the other would better suit his purposes. For instance, to show that a given event E need not have a certain conjectured cause C, he says we can conceive of E without conceiving of C (T 1.3.3.3), but he should really say that we can conceive of (E's occurring without C's occurring).¹⁴ I find it striking that both Spinoza, the great rationalist, and Hume, the great empiricist, make extensive appeal to whether one thing is conceivable apart from another, and both are systematically insensitive to a crucial ambiguity in what such conceivability might mean.

¹⁴ For further instances of Hume's conflation, see Van Cleve, N.D.

I have been unable to reach a verdict on whether Spinoza accepts the starting point of Descartes's argument for dualism when he says thought and extension are conceivable "one without the help of the other." In the next three sections, I consider whether even on the assumption that he does accept the proper starting point, he can still resist Descartes's conclusion.

3. Descartes's Argument for the "Real Distinction" Between Mind and Body

To answer the question just posed, we must get clearer on the structure of Descartes's argument.

What premises does it involve in addition to the one we have been considering?

Here is how he puts his argument for the "real distinction" between mind and body in response to the objectors who asked him to present his arguments in geometrical fashion:

God can bring about whatever we clearly perceive in a way exactly corresponding to our perception of it (preceding Corollary). But we clearly perceive the mind, that is, a thinking substance, apart from the body, that is, apart from an extended substance (Second Postulate). And conversely we can clearly perceive the body apart from the mind (as everyone readily admits). Therefore the mind can, at least through the power of God, exist without the body; and similarly the body can exist apart from the mind.

Now if one substance can exist apart from another the two are really distinct (Def. X). But the mind and the body are substances (Defs. V, VI, and VII) which can exist apart from each other (as has just been proved). Therefore there is a real distinction between the mind and the body. (CSM 2, 119-20)

On the basis of this passage, I reconstruct the argument as follows:

1. It is conceivable that (I exist & my body does not).
2. Whatever is conceivable is possible.
3. Therefore, it is possible that (I exist & my body does not).
4. If it is possible that (I exist & my body does not exist), then $I \neq$ my body
5. Therefore, $I \neq$ my body.

Premise 1 is Descartes's starting point, put in the requisite single-occurrence form.¹⁵ Premise 2 is the conceivability-to-possibility principle endorsed by Descartes, Spinoza, and Hume. In the passages I have quoted from Meditation 6 and again just above, Descartes puts it more picturesquely as "what I can conceive, God can achieve." Step 3 follows from premises 1 and 2. Premise 4 is supported in the Second Replies by Descartes's definition X, "Two substances are said to be really distinct when each of them can exist apart from the other," but there are other good ways to arrive at it. By Leibniz's Law, the possibility of my existing without my body entails the possibility of my not being identical with my body. By the necessity of identity, the mere possibility of my not being identical with my body entails that in actuality, $I \neq$ my body.¹⁶ Finally step 5 follows from 3 and 4.¹⁷

So the argument is perfectly valid. But it is valid only because in my presentation of it, I have taken a provisional stand on another interpretive issue: is the conceivability Descartes affirms at the outset the conceivability of a *singular* proposition or a *general* one? In the presentation above, it is the conceivability of something singular—*I* could exist even if my body did not.

Singular versus general is another dimension of variation alongside single occurrence versus double occurrence in how to take Descartes's starting point. We have settled on single occurrence. But when it is said that I can conceive of $(A \ \& \ \sim B)$, are the instances of A and B singular propositions, as in

¹⁵ If desired, we could start instead from 'it is conceivable that I think and no body exists', then use premise 2 and the *cogito* (= my thinking entails my existing) to reach step 3. I used a reconstruction starting out this way in Van Cleve 1983.

¹⁶ Alternatively, since the best basis for the necessity of identity appeals to Leibniz's Law instantiated to modal properties, we could derive 4 from Leibniz's Law directly by instantiating to the modal property 'being possibly such as to exist without my body existing.'

¹⁷ As is often noted, there are strong parallels between Descartes's argument for substance dualism and the arguments of contemporary property dualists like Kripke and Chalmers. Substitute 'pain' and 'fiber firing' (or any other type of brain state) for mind and body in Descartes's argument and you get an argument for property dualism very close Kripke's (1972, 144-55). Reverse the direction of conceivability by making it the conceivability of physical states without mental states (which Descartes says you can equally well do) and you get an argument very close to Chalmers's zombie argument (2003, 3.2).

I conceive of (my existing without any body's existing)

I find it conceivable that (I think & I am unextended)

Or are they general propositions, as in

I conceive that Mind could exist apart from Body

I conceive of something's thinking without being extended. In symbols: $C\exists x(Tx \ \& \ \sim Ex)$.

Descartes presents his argument sometimes in singular terms, sometimes in general; he is as insensitive to this distinction as Spinoza is to single versus double. In the Second and Sixth Meditations and at the beginning of his reply to Arnauld, he asserts that *he* could exist even though nothing extended did. But in the reply to the Second Objections and throughout most of his reply to Arnauld, he only asserts that “Mind” (or “the mind”) could exist apart from “Body” (or “the body”).¹⁸

Despite Descartes's indifference to the distinction, it is of critical importance. When presented in singular terms, as in the argument four paragraphs back, it is valid. When presented in general terms, it is not. Suppose, for example, that his starting point were ‘It is conceivable that some things are thinking, but not extended.’ It would follow that the attributes of thought and extension are not necessarily co-instantiated and not necessarily identical. But it would *not* follow that *I* who think am not also extended; still less would it follow that *nothing* can be both thinking and extended. We would fall well short of Descartes's conclusion. (See Van Cleve 1983, 40, for elaboration.)

So why not simply adopt the singular starting point? There are two reasons that stand in the way, for Descartes if not for others. First, it is sometimes held that a condition of entertaining a singular proposition is that one be acquainted with its subject, and Descartes affirmed that he was

¹⁸ Latin has no articles. Some translators supply them and some don't.

not acquainted with himself.¹⁹ He inferred his existence in the *cogito* from the need for a subject of any attributes of which he was aware.²⁰ Second, it is a tenet of Descartes's philosophy that extension is an essential trait of bodies. If unbeknownst to him he were a body (a possibility very much alive in the Second Meditation), then even if thinking as such did not require extension, *his* thinking would require him to be extended. In that case, it is hard to see how he could find it conceivable—that is, have a clear insight into the possibility—that he thinks without being extended.²¹ It seems that he would already have to know that he is not a body in order to see that he could think without being extended.²²

We have yet to uncover a Cartesian argument establishing a conclusion at odds with Spinoza's view that thought and extension can and do belong to the same substance. The only starting point to which Descartes is entitled is something general—the conceivability of thought's being instantiated without extension and vice versa—and that, by itself does not lead to the conclusion that thought and extension cannot be instantiated in the same substance. To reach that conclusion, further premises are required. What might they be? And would they have any claim on Spinoza's allegiance?

4. The Principal Attribute Thesis

Partly in recognition of the difficulties Descartes's argument would face if set out as above, Marleen Rozemond (1995) has suggested an alternative interpretation. In her interpretation,

¹⁹ He says as much to Hobbes in the Third Replies at CSM 2, 124.

²⁰ Descartes affirms the principle that when we perceive any attribute, act, or accident, we can infer the existence of a subject to which it belongs in several places, sometimes in contexts suggesting that it is an unstated premise in the *cogito* argument: Principle I.11 (CSM 1, 196); Principle I.52 (CSM 1, 210); Replies to the Second Objections (CSM 2, 114).

²¹ The phrase within dashes indicates the explication of finding something conceivable I advocated in 1983. I contrasted seeing that something is possible (or its seeming possible) with *not* seeing that it is impossible, calling the former “strong conceivability” and the latter “weak conceivability.” Chalmers has given the distinction further elucidation and a better name—“positive conceivability” versus “negative conceivability” (Chalmers 2002).

²² A reconstruction of Descartes similar to mine and a criticism of it similar to mine are given in Shoemaker 1983. Our articles, which appeared in the same year, were independent of each other.

Descartes does not argue from the possibility of mind and body existing apart to there being a “real distinction” between them, but in the opposite direction, from real distinction to separability. How, then, to establish the real distinction? She sees the case for it as resting on a premise that I have so far not mentioned: the Principal Attribute Thesis, articulated as Principle 53 of the *Principles of Philosophy*.

According to this thesis, every substance has exactly one principal attribute, that is, one master property that is essential to it and of which all its other properties are modes or specifications. Here is how Descartes puts it:

To each substance there belongs one principal attribute; in the case of mind, this is thought, and in the case of body it is extension. A substance may indeed be known through any attribute at all; but each substance has one principal property which constitutes its nature and essence, and to which all its other properties are referred. Thus extension in length, breadth and depth constitutes the nature of corporeal substance; and thought constitutes the nature of thinking substance. Everything else which can be attributed to body presupposes extension, and is merely a mode of an extended thing; and similarly, whatever we find in the mind is simply one of the various modes of thinking. For example, shape is unintelligible except in an extended thing; . . . while imagination, sensation and will are intelligible only in a thinking thing. By contrast, it is possible to understand extension without shape or movement, and thought without imagination or sensation, and so on. (CSM 1, 210-211)

One property is a mode or specification of another iff having the first entails having the second, but not conversely, as in the contemporary notion of determinates and determinables.

Alternatively, given the conceivability test for possibility, A is a mode of B iff we can conceive of B without A but not A without B. Being square is thus a mode of being extended, and wondering when the rain will stop is a mode of thinking. A is a principal attribute of S iff (i) A is essential to S, (ii) every property of S other than A is a mode or specification of A in the sense indicated above, and (iii) A is not a property necessarily possessed by everything. I add (iii) so as not to count properties such as *being either extended or unextended* as principal attributes.

It is important to note that Descartes and Spinoza do not use the term ‘mode’ in the same way. For Descartes, modes are *properties*—in particular, accidental properties that are specifications or determinates of more general essential properties. (See Descartes, *Principles* I.56, CSM 1, 210-211, and Rozemond, 7 and 8). For Spinoza, modes are usually *things*—in particular, dependent things like rocks and trees and you and me, who exist “in” the one substance as beings dependent on it. A mode is related to its underlying substance as a wave is related to the ocean; its existence consists in the substance “doing” something or other, like heaving up first here and then there. We are to God as the wave is to the ocean. One might relate the two senses of ‘mode’ as follows: a Spinozistic mode is an individual thing whose existence consists in God’s having some Cartesian mode.²³

How might the Principal Attribute Thesis (henceforth, PAT) be used to prove that thought and extension cannot be combined in one substance? Rozemond gives us all the materials for an argument I rearrange as follows:

A. If thought and extension were combined in the same substance, then, by the PAT, one of the following three possibilities would have to obtain:

1. Thought and extension are *both* the principal attribute, in which case they would have to be the same property; or
2. just *one* of them is the principal attribute, in which case the other would have to be a mode of it; or

²³ Not seeing that Spinoza uses ‘mode’ in a sense different from Descartes’s leads to disastrous interpretations of his philosophy. Pierre Bayle took Spinoza’s claim that everything but God is a mode of God to be the claim that everything but God is a *property* of God. That would make God the only bearer of properties, which led Bayle to accuse Spinoza of holding that Peter’s vile thoughts are really God’s vile thoughts, “the most monstrous hypothesis that can be imagined” (note N in the article “Spinoza” in Bayle [1697] 1965). In fact, a Spinozistic mode can quite well be the bearer of properties not possessed by the underlying substance, just as a wave can move laterally even if the ocean does not do that.

3. *neither* of them is the principal attribute, in which case both of them would have to be modes of some third attribute.

B. All three possibilities can be ruled out.

C. Therefore, thought and extension can never be present in the same substance.

How do we rule out 1-3? One might think that Descartes could do so in each case by using the conceivability test for the mode-attribute relation. We can conceive of a thing's thinking without being extended, so thought is not a mode of extension. We can conceive as well of the converse possibility of a thing's being extended without thinking, so extension is not a mode of thought. Together, these two applications rule out alternative 2. They also rule out alternative 1, since they show that neither thought nor extension entails the other, which implies that they are not the same attribute. Rozemond adds that Descartes simply took it as obvious that thought and extension are not the same thing (20).

But what about alternative 3, that thought and extension are modes of some third attribute?

That fits the characterization I gave in the Introduction of neutral monism, to which we now turn.

5. Neutral Monism

Here is a possible argument against neutral monism: If thought and extension were both modes of some third attribute, then we could not conceive of either of them without conceiving of the third. In that case, we would know what the higher attribute is. But in fact, we do not know this; therefore, there is no higher attribute.

I trust the reader sees what is wrong with that argument: it uses the wrong form of the conceivability test! Instead of saying, "If thought were a mode of higher attribute X, we could not conceive of thought-without-X," which is the correct single-occurrence form, it says, "If thought were a mode of X, we could not conceive of thought without conceiving of X," which is

the incorrect double-occurrence form.²⁴ As we saw above in discussing Arnauld's objection, it is only the first form of the conceivability test that gives valid results. If we restrict ourselves to that form here, there is no guarantee that we will ever discover the third attribute underlying thought and extension. Hence our ignorance of such an attribute would be no proof that none exists.

Rozemond does not discuss neutral monism by name, but she notes another argument Descartes might use against it. She says there is a passage in the Third Replies (which I have been unable to find) in which Descartes says that if we consider modes that presuppose the same attribute together, we will see that they have this attribute in common (21). I agree that there are many positive instances of this—when we consider squareness alongside triangularity, we see that they both presuppose extension—but I am not sure things are always this way. Upon comparing circles in two-dimensional space and spheres in three-dimensional space, we might not be able to discern the genus that also has hyperspheres in four-dimensional space as a species.

Here, then, is a second possible answer to De Vries's question: even if Spinoza accepts two key Cartesian premises—the properly construed starting point (single-occurrence and general) and the PAT—he can avoid Descartes's dualism of substances by being a neutral monist.

Neutral monism has interesting affinities to a view Bennett attributes to Spinoza (141-43). For neutral monism, thought and extension are *modes under a common attribute*. For Bennett's Spinoza, thought and extension are *attributes under which there are common modes* (which Bennett calls "unabstractable differentia"). His idea is that when Spinoza says that a mental

²⁴ It must be admitted that Descartes himself occasionally backslides in the matter of single versus double. In Principle 61 of Part I, he uses the ambiguous "we cannot understand a mode apart from its substance" and proceeds to employ it in the correct single-occurrence way. In *Comments on a Certain Broadsheet*, however, he recapitulates Principle 61 as "we cannot . . . clearly understand a mode unless at the same time we have a conception of the substance of which it is a mode," which explicitly goes over to the double-occurrence pattern (CSM 1, 298).

decision and a brain event are “the same thing” in 2p7s, he does not mean that the events are identical. Rather, he means that what you have to add as a further specification to thought to get a decision is the same feature you must add as a further specification to extension to get the corresponding brain event—a feature ungraspable by us, even if we are familiar with the decision that results when it differentiates thought and the brain event that results when it differentiates extension. Under both neutral monism and Bennett’s view, thought and extension would intersect in a third or “neutral” property, the neutral property being “higher” in the first case and “lower” in the second.

Neutral monism may also be parallel in a formal way to Donagan’s answer to De Vries. Thought and extension may be attributes that “express” the same underlying essence. Whatever “expression” is, there would be a many-one relation between attributes and some more fundamental essence.

6. Why Not Deny the Principal Attribute Thesis?

I strongly suspect that Spinoza would not find neutral monism to his liking. That would certainly be true if the attributes are, as Bennett says, *basic ways of being* or *most general properties*, not reducible to anything else (61). There is still another way, though, for him to accept the Cartesian starting point while rejecting the Cartesian conclusion: he need only deny the PAT.

It is fairly clear that this is what Spinoza does, even though he never says as much in any numbered proposition. A denial of the PAT could take two forms: one could hold (i) that a substance has *more than one* principal attribute, or (ii) that a substance has *no* principal attribute. At first glance, it may appear that Spinoza takes option (i), since he holds that the one substance has infinite attributes. On closer reflection, however, it is clear that none of these can be a

principal attribute as defined by Descartes, for it is not the case that all the other properties of Spinozistic substance are modes of some one among its attributes. It is better to see Spinoza as denying the PAT by taking option (ii).

Is there any reason why Spinoza or anyone else *should* accept the PAT? The views of David Wiggins on the sortal dependence of identity and substancehood afford an interesting potential source of support for the PAT (1967, 1980). Wiggins holds that no substance exists that does not exemplify some sortal property and that no identity ever holds between items that are not “covered” by some sortal. Sortals are the properties one would cite in answering the Aristotelian question “what is it?”; they are essential properties in the classical sense that to cease being F is to cease to be, period; and they are properties that determine the synchronic and diachronic identity conditions for the things that exemplify them.

Wiggins articulates and defends two theses, using ‘f’ and ‘g’ to range over sortals:

(Di) $(x)(t)(x \text{ exists at } t \rightarrow \exists g(gx \text{ at } t))$

and the superficially stronger

(Dii) $(x)\exists g(t)(x \text{ exists at } t \rightarrow gx \text{ at } t).$

The first says every substance at every time it exists exemplifies some sortal or other; the second says that for every substance, there is some sortal it exemplifies throughout its career. Wiggins argues that together with sound metaphysical assumptions, (Di) entails (Dii). My interest here is not with (Di) or (Dii) themselves, but with a premise Sydney Shoemaker employs in giving an improved version of the argument from (Di) to (Dii) (1970).

Here is the Shoemaker premise, labeled by him as (II):

(II) If f and g are co-satisfiable (possessable by the same thing at the same time), they restrict a common sortal.²⁵

The notion of f 's restricting g may be illustrated as follows: 'woman' restricts 'human' and 'human' restricts 'animal'. In these cases, being f is a matter of being g plus something more, as being a woman is a matter of being human and female. In any case, being f entails being g .²⁶

Shoemaker argues for (II) as follows (536-537). Suppose an f and a g coincide for a while—they are in exactly the same place—but at some later time do not coincide. Then the sortals f and g cannot have been co-satisfied at the earlier time—there were two coincident things there, one an f and the other a g , but not one thing instantiating both sortals. Examples: If a person and a body coincide for a while, but then go their separate ways, say because the person comes to inhabit a new body, then there were two things there originally, not one thing that was both a person and a body. If the ship of Theseus and a collection of planks coincide for a while, but not at a later time when the gradually repaired ship is at sea and the reunited collection of planks is in dry dock, then their earlier coincidence was merely that, not identity between a ship and a collection of planks. Generalizing, if an f and a g go their separate ways, then the sortals f and g were never satisfied by the same thing to start with. Contraposing (and letting possibility transmit across entailment), if two sortals are genuinely co-satisfiable, then there is no possibility that their instantiators coincide for a while and then no longer.

To round out the argument, Shoemaker claims that if there is no possibility of divergence as just described, then the sortals f and g share the same criterion of identity. And if they share the same criterion of identity, the best and perhaps the only explanation for this is that they restrict a

²⁵ The other premise in his argument is (I): if something satisfies f and then f' in an adjacent period, f and f' must be restrictions of a common sortal. (I) and (II) together imply that the relation "restricts the same sortal as" is transitive and thus, claims Shoemaker, that (Dii) is true. The last claim, however, is questioned in Tennant 1977.

²⁶ Wiggins also operates with the notion of a *phase* sortal, in which the essence condition is dropped. 'Girl' restricts 'woman', but ceasing to be a girl is not ceasing to be.

common sortal, whose principles for tracing things through time are shared by whatever sortals restrict it. Co-satisfiable sortals must restrict a common sortal.

I leave the evaluation of this argument to the reader or for another time. Here I just comment on its relevance to PAT and to the issue separating Descartes and Spinoza.

How close does (II) bring us to Descartes's PAT? The two theses have clear affinities, but they are not quite the same. For one thing, Descartes's PAT makes a claim of uniqueness that Shoemaker's (II) may not secure.²⁷ Consider the following diagram, in which T, E, and X (f, g, and h) are all sortals possessed by the same substance:

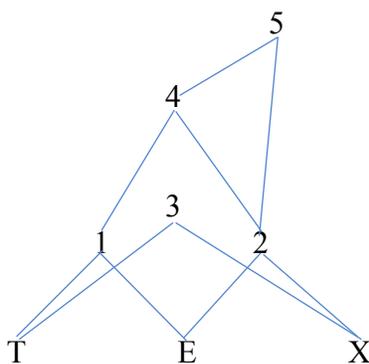


Figure 1

Lines running upward indicate the relation of restriction. According to (II), T and E must restrict the same sortal, which they do, namely 1. Likewise for E and X, which restrict 2, and T and X, which restrict 3. Is there a way to show that 1, 2, and 3 are the same sortal? I don't know, but we can show that T, E, and X do all restrict *some* same sortal, namely 4. 1 and 2 are co-satisfiable, because each is entailed by the satisfied E. By (II) again, 1 and 2 restrict a common sortal, which I label as 4. Now the relation of restricting is transitive, so each of T, E, and X restrict 4, as inspection reveals; so T, E, and X restrict a common sortal, and 4 is the only sortal so far mentioned that the diagram represents them all as restricting. Can we show that there is

²⁷ The PAT also makes a claim of ultimacy—that there are attributes that restrict nothing beyond themselves. Shoemaker's (II) does not secure this part, either.

just one sortal restricted by everything at the bottom? Not that I can see. A Third Man regress threatens to break out at this point. 4 and 2 are co-satisfiable, being entailed by the co-satisfied T and E. Therefore, they both restrict a common sortal, which I label as 5, which is also restricted by everything at the bottom. Though it is possible that 5 is identical with 4 (since the relation of restricting, unlike the Cartesian relation “being a mode of,” is not irreflexive), I do not see how to prove that 5 *must* be identical with 4.

So the Shoemaker considerations do not prove the full-strength PAT with its implication of a unique common sortal. They do not need to do that, however, in order to make trouble for Spinoza: the requirement of *a* common sortal would be trouble enough. Consider Spinoza’s proposition that thought (T) and extension (E) are present in the same substance, along with others of the “infinite attributes”—X, Y, Z, and the other attributes on beyond Zebra. I was already looking ahead to this proposition when I chose labels for the bottom row in Figure 1. If (II) is correct, there would have to be at least one attribute that T and E both restrict.²⁸ What could it be? One possibility is that T restricts E and E restricts itself, which would be a form of materialism. This possibility Descartes rejects, since one can conceive of something’s exemplifying T without exemplifying E, and so does Spinoza if he accepts the Cartesian single-occurrence starting point. Similar remarks apply to the mirror-image alternative of E’s restricting T while T restricts itself. The remaining alternative is that T and E both restrict some third sortal, which is what I have been calling neutral monism. So it looks as though the Wiggins-Shoemaker premise, while not establishing Descartes’s PAT, does constrain Spinoza’s options. If Wiggins and Shoemaker are right, then if Spinoza accepts Descartes’s starting point,

²⁸ This assumes that T and E are both sortals. Perhaps Wiggins would deny this in the case of extension—he says that ‘space-occupier’ is not a sortal (1980, 63). On the other hand, Descartes and arguably also Spinoza hold a space theory of matter, equating extension with materiality; perhaps extension as they construe it is a better candidate for sortal status.

he can have one substance possessing both Cartesian attributes only if he embraces neutral monism.

It is time to summarize. De Vries's challenge to Spinoza is this: how can you admit the Cartesian premise that thought is conceivable apart from extension and then go on to say that thought and extension characterize the same substance, contradicting Descartes's dualism of incompatibles? I have canvassed three possible replies or escapes for Spinoza.

First way out: Spinoza does *not* accept the Cartesian premise—not in the form required for the validity of Descartes's argument for dualism. Making this reply requires making a distinction to which Spinoza seems unfortunately oblivious—the distinction between (i) conceiving of thought without conceiving of extension and (ii) conceiving or finding possible the combination thought-without-extension. Descartes's argument requires (ii). I was unable to determine definitively whether Spinoza would or could deny the conceivability of thought-without-extension.

Second way out: Suppose Spinoza does admit the conceivability of thought-without-extension as Descartes does. He is not yet committed to accepting Descartes's dualism of incompatibles, for Descartes requires further premises. The one proposed by Rozemond is the Principal Attribute Thesis, according to which every substance has just one essential attribute of which all its other properties are specifications. This sounds on its face like a thesis Spinoza would reject—his substance has at least two and perhaps infinitely many attributes, none of which is “principal.” But perhaps there are good arguments for the thesis! I considered an argument (by Wiggins as helped by Shoemaker) that does not establish the full-strength PAT, but which does reduce Spinoza's room to maneuver. It implies that if the same thing is both thinking and extended (and neither of these properties is a specification of the other), there must

be some higher sortal whose in-built identity conditions dictate that the thinking thing is the extended thing. I did not take a stand on the Wiggins-Shoemaker argument,²⁹ but readers who find it persuasive will find Spinoza's denial of it problematic.

Third way out: Suppose Spinoza is under rational pressure to accept both the Cartesian starting point and a weak version of the PAT. He still need not accept Descartes's dualism, for he could accept neutral monism instead: there is a sortal higher than thought and extension, entailed by each of them but identical with neither, of which they are restrictions. Although I do not think Spinoza in fact accepts neutral monism, I cannot see disaster for him if he does. There are precedents or analogs or hints of it in some commentators, including Alexander (to be mentioned again in Part Two), Bennett, and Donagan.

Part Two

7. Identity Puzzles

In this part I turn to puzzles engendered by the identity theses that Spinoza advances as underlying his isomorphism theorem, 2p7. A mode in Spinoza may be either an individual thing, such as a mind or a body, or an event in which an individual thing participates, such as a desiring or a movement. One of Spinoza's identity theses says that each mind is identical with the body correlated with it under the isomorphism. Another says that each mental event is identical with the bodily event correlated with it under the isomorphism. The converses hold as well—each body is a mind and each physical event is a mental event.

The first puzzle is how the identity of minds and bodies is to be reconciled with the distinctness of their constitutive attributes, thought and extension. This puzzle is fairly easy to dispel. The second puzzle is how the identity of mental events and physical events is to be

²⁹ I have not argued the point here, but I am inclined to think myself that there must be some entities whose identity, whether synchronic or diachronic, is primitive and not grounded in any sortal.

reconciled with Spinoza's affirmation that there is causation connecting mental events with mental events and physical events with physical events, but no causation connecting mental events with physical events. If event P1 causes event P2 and $P1 = M1$, must not M1 cause P2? This is one of the problems I advertised in the opening paragraph of this paper, and it is harder to solve.

Let us begin with the easier puzzle: how can my mind and my body be identical if their constitutive attributes of thought and extension are distinct? This is not De Vries's question all over again. De Vries's question was how one and the same *substance* can have two attributes that are conceivable apart from each other and thus distinct. The current question is how *individual modes*, such as my mind and my body, can be identical if their constitutive attributes are distinct. The problem may be developed further as follows, using 'm' for my mind and 'b' for my body:

1. $m = b$
2. Necessarily, m exists iff b exists—from 1 and the necessity of identity.
3. Necessarily, m exists iff God is T*, where T* is some determinate of thought. This follows from how I unpack Spinoza's view that modes exist "in" God; their existence consists in God's being some way under the appropriate attribute.
4. Necessarily, b exists iff God is E*, where E* is some determinate of extension. This holds for the same reason as 3.
5. Necessarily, (God is T* iff God is E*)—from 2, 3, and 4, substituting necessary equivalents within a necessity context.

I am aiming at the conclusion that thought and extension are necessarily coextensive, but we have not quite arrived there. To say that the determinates of thought and extension T* and E* are necessarily such that God instantiates one iff God instantiates the other is not to say that thought and extension themselves are necessarily coextensive throughout the universe.

Nonetheless, it is hard to see why 5 would be true if the attributes were not necessarily coextensive.³⁰ In that case, how can they be distinct?

This question has an answer. If thought and extension do turn out to be necessarily coextensive, Spinoza can still uphold his dualism of attributes at least to the extent of maintaining that thought and extension are *distinct* attributes. Our first puzzle only shows that to make much headway in Spinoza's philosophy, we must accept *hyperintensionality*, a phenomenon that has come to the fore in recent metaphysics (see Nolan 2014), but which is as old as Plato's *Euthyphro*. An extensionalist, like Quine or Davidson, would say that properties and kindred entities are identical when they are actually coextensive. An intensionalist, like Carnap or David Lewis, would say that properties have a tighter condition for identity, being identical when they are *necessarily* coextensive (Lewis 1986, 50-51). A hyperintensionalist says that even necessarily coextensive properties may be distinct, as witnessed by the properties of being three-sided and being three-angled or the properties of being equal to 1 and being equal to the sum of the squares of the sine and cosine of some number. If we bring hyperintensionalism into the interpretation of Spinoza, it is no surprise at all that thought and extension are necessarily coextensive but nonetheless distinct.

A feature of hyperintensionality that goes along with the one just mentioned is that there can be relations borne to one of two necessarily equivalent items that are not borne to the other. A case in point is the now much-investigated relation or operation expressed by 'is grounded in' or

³⁰ Another possible source of the necessary coextensiveness of the attributes is 2d2, a definition of essence: "I say that there pertains to the essence of a thing that which, when posited, the thing is necessarily posited, and by the annulling of which the thing is necessarily annulled." Bennett reads the second clause as saying an essence of x must be possessed by x and the first as saying that an essence of x cannot be possessed by anything other than x (61). Add to this that thought and extension are both essential to God and you can prove that they are necessarily coextensive. Bennett doubts, however, that this is the sense of 'essence' that figures in Spinoza's definition of attributes at 1d4.

‘because’ (Correia and Schnieder 2012). ‘P because Q’ has the following properties: it can hold between P and Q without holding between P and some R necessarily equivalent to Q; it can hold in just one direction between two states of affairs that are necessarily equivalent;³¹ it does not always hold between P and Q when Q entails P; it never holds between a state of affairs and itself. Spinoza’s philosophy absolutely requires such an operation, even if he lacks dedicated vocabulary for expressing finer-than-modal distinctions (unless it is ‘quatenus’, of which more anon).

Let me illustrate. It is a key tenet of Spinoza’s philosophy that God is the only being whose essence involves existence, or who exists by his own nature (1p7, 1p11, and 1p24.). It is also a tenet of his philosophy under most interpretations of it that all the modes exist of necessity, since their existence necessarily flows from the existence of God, which is itself necessary (1p29). If existing of necessity were equated with existing by one’s own nature, the foregoing propositions would be inconsistent. But if existing by one’s own nature meant existing not just necessarily in consequence of one’s nature, but *because* of it (a hyperintensional because), the propositions can be reconciled. Modes are beings that exist just as necessarily as God does, but unlike God, they do not exist because of or in virtue of their own natures.

Here is another example in the same vicinity: though on the necessitarian interpretation of Spinoza’s philosophy, *God exists* and *Peter exists* are both necessary truths (and therefore mutually entailing), there is still room to ask the Euthyphro question whether God exists because Peter exists or Peter exists because God exists. That ‘because’ must be hyperintensional.

For a hyperintensionalist, the puzzle of necessarily coextensive attributes that are nonetheless distinct is no puzzle at all. I must acknowledge, however, that there is an offshoot problem that

³¹ Even though ‘someone carries x’ and ‘x is carried’ are necessary equivalent, Plato still has Socrates ask in the *Euthyphro* whether it is the first because the second or vice versa.

remains and to which I cannot do justice here. The necessary coextensiveness of the attributes is not incompatible with Spinoza's dualism of attributes if such dualism requires only their distinctness. But as I characterized it in the opening background section, his dualism says more. It says there is no logically or conceptually necessary conditional connecting the mental to the physical (as there would be, for example, if behaviorism or phenomenalism were true). We have given reasons above for thinking that Spinoza is committed to there being some kind of metaphysically necessary equivalence between the attributes. Unless this necessity does not involve logical or conceptual necessity, his philosophy is threatened with another inconsistency.

8. The Harder Problem

The second problem raised at the beginning of the previous section may be presented as an apparently valid argument from two premises Spinoza accepts to a conclusion he does not accept:

1. Mental event M1 causes mental event M2.

Spinoza would say there are plenty of cases in which instances of this premise are true; the sight of a dog may cause fright, and so on.³²

2. For some physical event P1, M1 = P1.

This premise is Spinoza's thesis of the identity of corresponding event modes; a visual experience is identical with the cortical event correlated with it.

3. Therefore, P1 causes M2.

This conclusion apparently follows from the premises, but directly contradicts Spinoza's thesis that there are no causal connections in either direction between the mental and the physical.

9. Della Rocca's Solution: Opacity to the Rescue

³² Some commentators suspect that Spinoza does not really believe in mental-to-mental causation. In that case, the problem I am raising can still be raised by starting with a case of physical-to-physical causation and showing that under identity it would be a case of trans-attribute causation.

I now focus on a solution to this problem advanced on Spinoza's behalf by Michael Della Rocca. According to Della Rocca, Spinoza's philosophy is one in which opacity is pervasive, affecting causal contexts in particular and thus making the argument above invalid.

According to Quine, who coined the phrase 'referentially opaque', a context ' \dots is F' is opaque iff terms referring to the same thing may not always be interchanged within the context with preservation of truth value (Quine 1961, 20). For ease of exposition, I will for the time being mix material and formal modes of speech, enabling me to characterize opacity as well in the following way, which is one of Della Rocca's ways (1993, 167-168):³³

If ' \dots is F' is an opaque context, then the argument

Fa

a = b,

Therefore, Fb

is not valid.

We could make the argument formally valid by adding as a premise Leibniz's Law: a = b only if (F) (Fa \leftrightarrow Fb). But then Della Rocca would extend the claim of opacity by saying that Leibniz's Law is not true across the board; it has exceptions whenever ' \dots is F' is an opaque context or, to put it in the material mode, whenever F is a nonextensional property (1993, 173-74).

To all of this, we now add that for Spinoza, ' \dots causes ___' is an opaque context. The argument showing a contradiction between mental-physical event identity and the ban on mental-physical interaction is therefore invalid.

Della Rocca divides the case for the opacity of causal contexts in Spinoza into two phases, first regarding immanent causation and then transitive causation (a.k.a. transeunt causation). (1993, 168-69). Immanent causation is "vertical" causation or the causation by a substance of its

³³ He also has the substitution way at 1993, 171.

own states or modes; transitive causation is “horizontal” causation or the causation of one mode by another.

Della Rocca takes the opacity of immanent causal contexts to be asserted in 2p6:

The modes of any attribute have God for their cause only in so far as he is considered under that attribute [the attribute of which they are modes, (C)], and not in so far as he is considered under any other attribute. (S)

Della Rocca takes this to imply that even though *the extended substance God* causes the existence of a certain planet and is identical with *the thinking substance God*, the thinking substance God does *not* cause the existence of that planet.

Spinoza states the proof of 2p6 as follows:

Each attribute is conceived through itself without any other (by 1p10). So the modes of each attribute involve the concept of their own attribute, but not of another one; and so (by 1a4) they have God for their cause only insofar as he is considered under the attribute of which they are modes and not insofar as he is considered under any other (C).

I do not see how everything that comes after the first occurrence of ‘So’ is supposed to follow. “The modes of each attribute involve the concept of their own attribute.” that much could be taken to follow from the way in which Descartes characterized the mode-attribute relation and the way in which Spinozistic modes are defined in terms of Cartesian modes. “But not of another one:” I do not see how this part follows. Let me use lower-case letters for modes and capitals for attributes. Spinoza has undercut *one* argument for thinking that modes of one attribute involve the concept of another—namely, C(t) requires C(T), C(T) requires C(E), therefore, C(t) requires C(E)—since 1p10 rules out that C(T) requires C(E). But I cannot see how he has positively ruled out that C(t) should require C(E).

Letting the first ‘So’ stand, however, I do grant everything that comes after the second ‘so’. If God considered under E caused t, then by 1a4 (the knowledge of an effect requires knowledge of the cause), C(t) would require C(E), contrary to what we just let stand.

Notice, by the way, that the proof of 2p6 evidently requires the double-occurrence reading of ‘conceived through’ as we reach back to 1p10 and 1a4. That is one more piece of evidence bearing on the question how best to understand that locution in Spinoza.

In the second phase of the argument for opacity, we show that transitive causal contexts are opaque. I agree with Della Rocca’s observation that the basis for proving opacity for immanent causation may be used again to establish opacity for transitive causation: that a mode under a given attribute is caused by another mode only if that other mode is conceived under the same attribute. Suppose otherwise: that t , a mode conceived under T , is caused by e , a mode conceived under E . Then $C(t)$ would require $C(e)$, which requires $C(E)$, so by transitivity $C(t)$ would require $C(E)$, which is contrary to what is asserted in 2p6d.

We now have it that causal contexts are opaque, which is enough to block the argument from event identity to mental-physical interaction. But Della Rocca goes much further in finding opacity in Spinoza. He shows that if causal contexts are opaque, *opacity spreads*, infecting a vast amount of the rest of our vocabulary. In particular, the contexts ‘... is extended’ and ‘... is thinking’ are opaque. We can prove this from two premises. First, for Spinoza, ‘... is extended’ entails ‘... is caused by the extended substance’ (immanent causation, this would be) and ‘... is thinking’ entails ‘... is caused by the thinking substance’. This premise follows from 2p6 along with Spinoza’s belief that necessarily, everything but God is caused by God. Second, the following *Opacity Transmission Principle* is true:

Any context that entails an opaque context is itself opaque.

Della Rocca makes a plausible case for this principle at 1993, 171.³⁴ From these two premises plus the already-argued opacity of immanent causation, it follows that ‘. . . is extended’ and ‘. . . is thinking’ are both opaque, along with any contexts that entail them.

What contexts are *not* opaque for Spinoza? In other words, which contexts are *extensional*? There are some, but by the argument above they must be *neutral*, entailing neither thought nor extension. Examples Della Rocca gives are *happening at such-and-such a time* and *being a complex individual*.

I pause to note that Della Rocca’s stricture on neutral properties agrees with what I said about neutral properties in characterizing neutral monism: they must be entailed by thought and extension, but must not entail either of them in return. I also note that the neutral monist Samuel Alexander claimed that his system is similar to Spinoza’s except that the single substance from whose properties matter and mind both emerge has only two attributes, those of space and time (Thomas 2018). Temporal properties (though not spatial ones, if they entail extension) are among Della Rocca’s neutral properties.

Della Rocca does not appeal to opacity only to deflect an argument against Spinoza’s identity theses; he also invokes it to construct a positive argument for one of the identity theses. It is not an abductive argument like the argument that identity best explains isomorphism, but a deductive argument. It runs as follows:

1. All extensional properties are neutral.
2. Mind and body share all their neutral properties.
3. Therefore, mind and body share all their extensional properties.
4. Things with all the same extensional properties are identical.
5. Therefore, mind and body are identical.

³⁴ His formulation of the principle is longer, but I believe I have accurately compressed it.

I do not say any more about this argument here, but I set it down because I find it audacious and worthy of the reader's consideration.

10. Problems with Opacity

When someone tells me that some philosophical position or some important philosopher's work requires a large dose of opacity, my first reaction is "Say it ain't so, Joe!" If opacity means allowing that there are exceptions to Leibniz's Law, then I find the idea hard to accept and hard even to understand. I now set forth the reasons why I am reluctant to accept opacity in Spinoza and why I think certain presumed exceptions to Leibniz's Law can be resisted.

First, I have always found compelling Quine's stricture that you can quantify into a context only if that context is transparent (1961, 22ff.) For example, if despite the fact that 'Tom' and 'Bob' refer to the same individual, you cannot legitimately infer from 'Mary has a tomcat' to 'Mary has a bobcat', then neither can you quantify into that context and say, 'for some x , Mary has an x cat'.³⁵ Applying this stricture to Spinoza, given the opacity of '... is thinking', we would have to disallow him the inference from 'I am thinking' to 'Something is thinking' (though he could still have the *cogito* inference from 'I am thinking' to 'I exist'). That is a high cost.

Second, by the Opacity Transmission Principle, if '... is thinking' and '... is extended' are opaque contexts, then so are any Cartesian mode contexts that entail them, such as '... is thinking about his mother' and '... was in the bank at noon on Monday'. (I assume in the

³⁵ I got this example years ago from Rolf Eberle; I suspect it is originally of UCLA provenance. Some may object that the example is not an example of opacity on the ground that 'tom' is not even a syntactical part of the 'tomcat' sentence. True, but it still counts as opacity according to Quine's characterization; we have a context surrounding an expression for which we may not substitute and in which the expression does not have referential occurrence. In those ways it is akin to another Quinean example of an opaque context, quotation: compare 'Mary called out "Tom!"; therefore, Mary called out "Bob!"'

example to follow that location is a mode of extension—if it is not, you may change the example to ‘. . . is over six feet tall’.) If such mode contexts are opaque, the following argument is invalid:

The robber was in the bank at noon on Monday.

I am the robber.

Therefore, I was in the bank on Monday.

If that argument is invalid, so is the following contrapositive of it:

The robber was in the bank at noon on Monday.

I was not in the bank at noon on Monday (but in the saloon across town having a drink).

Therefore, I am not the robber.

There goes my alibi! Crucially important reasoning becomes invalid if opacity is rife.³⁶

I would not disallow reasoning in accordance with Leibniz’s Law in the bank robber case. In fact, for reasons I now set forth, I would not disallow reasoning in accord with Leibniz’s Law in *any* case. How can this be, the reader may ask. Are there not undeniable cases of opacity?

Of opacity, yes. But of failures of Leibniz’s Law, no! It is time to stop treating the two as though they were formal and material mode versions of the same thing.

Since the approach I am about to set forth has been espoused by several philosophers once all associated with Wayne State University, it might aptly be called the Wayne Way.³⁷ It received the fullest exposition I know of in the work of Richard Cartwright (1971, 1979).

³⁶ Perhaps there is some way to qualify Spinoza’s thesis to avoid results like this. Perhaps Leibniz Law inferences are valid if the identity statements are intra-attribute identities, allowing them to establish that the body that committed the robbery is not my body (so my body should not be thrown in jail) and invalid only when the identity statements are inter-attribute identities. But that proposal would require a good deal of working out.

³⁷ The philosophers I have in mind are Richard Cartwright (1971, 1979), Alvin Plantinga (1974), and H.-N. Castaneda in talks I have heard him give.

Cartwright distinguishes the Principle of Substitutivity from the Principle of Identity. What he calls Substitutivity may be put in Quine's words as follows:

Given a true statement of identity, one of its two terms may be substituted for the other in any true statement and the result will be true. (Cartwright 1971, 136, quoting Quine 1971, 139).

What he calls Identity he formulates as follows.

If $x = y$, then every property of x is a property of y . (1971, 136)

Identity is often known as Leibniz's Law, and that is what I shall call it in what follows.

Cartwright accuses Leibniz and Quine of conflating the two principles and criticizes those who regard one as simply a material mode version of the other. He argues that counterexamples to Substitutivity are *not* counterexamples to Leibniz's Law. We have a case of this in my example above about Mary and her tomcat. If we use the true identity 'Tom = Bob' to substitute 'bob' for 'tom' in 'Mary owns a tomcat' to obtain 'Mary owns a bobcat', we pass from a truth to a falsehood. Thus the context 'Mary owns a . . . cat' is opaque; opacity in the sense of exceptions to Substitutivity indisputably exists. But it does not follow that any exceptions to Leibniz's Law exist.

Consider the Mary argument:

Mary owns a tomcat.

Tom = Bob.

Therefore, Mary owns a bobcat.

If an inference represented formally as 'Fa, a = b, therefore Fb' is to be sanctioned by Leibniz's Law as Cartwright formulates it, one premise must ascribe a property to a , and the conclusion must ascribe that same property to b . That is not what happens in the Mary argument. 'Mary owns a tomcat' ascribes no property to Tom. In the first place, that premise does not ascribe any property to *Tom*, even if we decapitalize his name; it says nothing about him. In the second

place, there is no such property as *being an item x such that Mary owns an xcat* to be ascribed to Tom or anyone else. So when Leibniz's Law is formulated Cartwright's way in terms of properties, the exception above to Substitutivity leaves it untouched.³⁸

To summarize, the Wayne Way consists of the following points: (i) Leibniz's Law is to be formulated in terms of properties and distinguished from the Principle of Substitutivity; (ii) when Leibniz's Law is thus understood, there are no genuine counterexamples to it; (iii) some apparent counterexamples to it are only counterexamples to Substitutivity, involving opaque contexts that do not express properties.³⁹

Some readers may think that my example of Mary and her tomcat is outré and not really to be compared with the more familiar examples of opacity. But in my opinion, it has the same root failing as those examples—it involves a context that fails to express any property—even if it has this failing for different and more egregious reasons.

Cartwright takes on some of the more familiar examples and shows that they crucially fail to involve property-attributing contexts. Consider the following putative counterexample to Leibniz's Law, based on Frege's puzzle, in which I use Cartwright's numbering:

- (10) It is a truth of astronomy that Hesperus = Phosphorus
- (11) Hesperus = Phosphorus.
- (12) It is *not* a truth of astronomy that Phosphorus = Phosphorus.

³⁸ In 1971, starting at p. 137, Cartwright makes this point with several examples, including Quine's triad 'Giorgione was so called because of his size (true); Giorgione = Barbarelli (true); therefore, Barbarelli was so called because of his size' (allegedly false). There is no counterexample to Leibniz's Law there. The supposed property of being so called because of one's size, allegedly had by Giorgione but not by Barbarelli, is no property at all. The supposed property of being called 'Giorgione' because of one's size *is* a property of Giorgione, but Barbarelli has it, too.

³⁹ In Plantinga, point (i) may be found on p. 15 and in n. 3 to that page, point (ii) on pp. 220 and 238, and point (iii) on pp. 226-27.

That trio constitutes a counterexample to unrestricted Substitutivity, but not to Leibniz's Law.⁴⁰

For it to be a counterexample to Leibniz's Law, there would have to be a property ascribed to Hesperus by (10) that is denied of it by (12). What would that property be? The property of being an object x such that the proposition that x is identical with Phosphorus is a truth of astronomy? But there is no such property, Cartwright argues, since for a given object, there is no such thing as *the* proposition that it is identical with Phosphorus (1979, 203ff.).⁴¹

In a follow-up article some years later (1979, in his 1987), Cartwright airs second thoughts about his 1971 article. Chief among them is that he no longer regards Leibniz's Law as a master principle from which other indiscernibility principles may be derived, but only as one indiscernibility principle among others. But the chief elements of the Wayne Way remain intact. He still maintains that no indiscernibility principle is ever false. Sometimes there are counterexamples to what follows from an indiscernibility principle along with a comprehension principle—for instance, a principle affirming that such-and-such a property exists—but in those cases, it is the comprehension principle that is false

In light of our discussion of opacity and Leibniz's Law, let us now review the inference from mental-physical identity to mental-physical interaction. Here it is again:

1. Mental event M1 causes mental event M2 (e.g., the sight of a dog causes fright).
2. For some physical event P1, $M1 = P1$.
3. Therefore, P1 causes M2.

⁴⁰ Frege's own solution to the puzzle, which reconstrues (10) and (12) as being about the senses of 'Hesperus' and 'Phosphorus' rather than Hesperus and Phosphorus themselves, also admits no exception here to Leibniz's Law.

⁴¹ See Cartwright 1971, 143ff., for a similar treatment of '. . . is necessarily greater than 7'. In brief, that context can be taken *de re*, in which case it expresses a property both of 9 and the number of planets, or *de dicto*, in which case it does not express a property at all.

If we follow the Wayne Way, that inference would be perfectly sound, provided the first premise ascribes a property to M1—the property of causing M2. And why would it not? Why isn't causing fright a property of my seeing the dog and thereby of any neural state identical with the seeing, contrary to Spinoza's ban on interaction? The opacity defense of Spinoza seems to have fizzled.⁴²

But let's give it one more round. There remains a way in which partisans of opacity might block the inference above, which I learned about from Bacon and Russell 2019—a work not necessarily meant to defend opacity, but to articulate its logic. Those who accept Bacon and Russell's logic can accept all the conditions required in the Wayne Way for the validity of the inference above: the unstated version of Leibniz's Law that says every property of anything belongs to anything identical with that thing;⁴³ the contention that there is such a property as causing fright, and that it belongs to M1; finally, the premise that M1 is identical with P1. They can nonetheless deny that the property of causing fright belongs to P1. How so? Is that not a contradictory combination?

Let's put the combination in slightly different words. If a property is possessed by something, it is also possessed by anything identical with that thing—that goes for all properties; causing fright is a property; but it doesn't go for causing fright that if it is possessed by M1, it is also possessed by the identical P1. If that sounds contradictory to you, it is no doubt because you are taking for granted the law of universal instantiation—what goes for all properties goes for the property of causing fright. But in Bacon and Russell's logic, universal instantiation is one of the laws that must be given up. That should come as no surprise if you reflect on two points: (i) as is

⁴² May we save Spinoza from the 1-2-3 inference simply by pointing out that it is sanctioned by Substitutivity, whose falsehood we have conceded? No, for it is also sanctioned by Leibniz's Law, whose falsity we have not conceded. An argument that is an instance of one valid pattern is valid, period.

⁴³ Bacon and Russell formulate this principle in second-order logic as follows: $\forall x\forall y\forall X(x = y \rightarrow Xx \rightarrow Xy)$.

widely agreed, you cannot quantify into opaque contexts—that is, draw inferences in accordance with existential generalization, and (ii) universal instantiation is the dual of existential generalization; they stand or fall together. For Bacon and Russell, they both fall; that is part of the price of believing in opacity.⁴⁴

For those who find that price too high, is there any alternative to saddling Spinoza with opacity? I am now taking opacity to consist not just in the existence of opaque contexts, which I have conceded, but also in the failure of some applications of Leibniz's Law.

11. Hyperintensional Alternatives

In connection with certain passages in Spinoza, I have already identified one alternative to opacity, namely, hyperintensionality. Let us return to 2p6, which Della Rocca puts forth as the chief article of opacity for immanent causation. It is framed thus:

The modes of any attribute have God for their cause only in so far as he is considered under that attribute, and not in so far as he is considered under any other attribute. (2p6)

Immanent causation in Spinoza is generation of modes by God; 'God immanently causes mode 1' is another way of saying 'mode 1 exists *in* God'. A Spinozistic mode exists only if it exists in God, and its existing in God consists in God's being some way—in God's possessing some Cartesian mode or property. I would unpack 'consists in' in terms of the hyperintensional *because*. When a mode exists, it is not merely the case that necessarily, it exists if and only if God is a certain way; it exists *because* God is that way. A finite extended thing exists because God has a certain Cartesian mode under the attribute of extension. The Cartesian mode may according to Spinoza be necessarily coextensive with some Cartesian mode of thought, but it will

⁴⁴ Bacon and Russell go this far with the Wayne Way: they hold that Leibniz's Law in the form ' $\forall x\forall y\forall X(x = y \rightarrow Xx \rightarrow Xy)$ ' is true without exception. But they note that without universal instantiation with respect to both individual and property variables, that law does not imply that if Hesperus = Phosphorus and Asher believes Hesperus is bright, then Asher believes Phosphorus is bright—even if we allow that there is a property expressed by 'Asher believes that ___ is bright'.

not be true that the extended thing exists *because* God has that Cartesian mode of thought. This is one way to understand “the modes of any attribute have God for their cause only in so far as he is considered under that attribute”—and it makes immanent causation a hyperintensional context rather than one violating Leibniz’s Law.

But how would hyperintensionality help us get around the problem of squaring identity with the ban on interaction? I now sketch two possible ways.⁴⁵

The first way may be sketched in the following diagram:

$M1 \rightarrow M2$	$M1 = P1$	$\sim(P1 \rightarrow M2)$
$G_{M1} \Rightarrow G_{M2}$	$N(G_{M1} \text{ iff } G_{P1})$	$\sim(G_{P1} \Rightarrow G_{M2})$

With the arrow standing for causation, the top row displays three facts asserted by Spinoza that apparently conflict with Leibniz’s Law: M1 causes a subsequent mental event M2, M1 is identical with physical state P1, yet P1 does not cause M2. The bottom row displays the three facts that in Spinoza’s philosophy underlie or ground the top facts. The occurrences of modes M1 and M2 each consist (as with any modes) in God’s being a certain way, indicated here by G_{M1} and G_{M2} .⁴⁶ The causal relationship between M1 and M2 is underlain by a corresponding causal relationship between the two mode-constituting facts about God, indicated by the double arrow; I need take no stand here on whether \rightarrow is the same or subtly different from \Rightarrow . The important point for our purposes is that \Rightarrow is a hyperintensional relation, in which case the three

⁴⁵ Thanks to Matt Davidson for getting me to think harder about applying hyperintensionality to the interaction problem.

⁴⁶ Recall that Spinoza uses the term ‘mode’ both for events, like my being in pain, and for things, like me. My pain is a mode of me, and I am a mode of God, so my pain is a mode at second remove. It is plausible and perhaps even demonstrable that the mode-making relation is transitive, so the existence of any event mode like pain consists in some fact about God—not necessarily the fact of his being in pain and not necessarily the same fact that gives rise to me, the subject of the event, or I would be in pain perpetually.

entries in the bottom row are totally consistent with each other and do not conflict with Leibniz's Law. Problem solved.

Or not. What have we actually done? We have shown that the three facts in the top row are grounded in three facts in the bottom row that by themselves are not problematic. But we have not gotten rid of the three top facts, which *do* stand in contradiction with Leibniz's Law. So I am not sure we have made the problem go away. Perhaps we could rest content with this solution if we regarded modes as logical constructions and held, with Arthur Prior, that relations of "identity" we can define for logical constructions are not genuine cases of identity.⁴⁷

The second way makes use of one of Spinoza's favorite words, 'quatenus', which is often translated as 'in so far as it is considered as' and which I render here simply as 'qua'. In this solution, we insist that affirmations of causal relationships between two events or facts are never true *simpliciter*, but only with a 'qua' or two. In a slogan, all causing is causing *qua*.⁴⁸ The first premise in the statement of the interaction problem cannot then be simply 'M1 causes M2', but instead must be

1. M1 qua mental causes M2

or more likely

1'. M1 qua mental causes M2 qua mental

but to spare us some complexity I will stick with 1. The next premise would be the identity

2. M1 = P1

and the conclusion would be

3. P1 qua mental causes M2.

⁴⁷ See Prior 1968, 191-94, where he says the "identity" of virtual classes, unlike the identity of individuals, is only fictitious identity.

⁴⁸ This is a thesis I have seen attributed to Donald Davidson in his 1970, but I do not believe he holds it in the form required here. For Davidson, if *e* causes *f* under description *d*, *e* causes *f*, period.

This time Spinoza could accept 3 as legitimately following from 1 and 2—he commits no breach of Leibniz’s Law—and he could accept 3 as not flouting his interaction ban—he allows that a physical event *qua mental* really can cause a mental event. One thing he could not accept, though, would be

4. M1 qua physical causes M2

even though its flanking terms are mental events. And despite the necessary equivalence advanced in section 7 of thought (mentality) and extension (physicality), he need not infer 4 from 1, provided ‘qua’ is hyperintensional.⁴⁹ *X qua F is H and necessarily, (something is F iff it is G) do not imply X qua G is H.*

What should we make of this solution to the interaction problem? I am not sure, primarily because I am not sure how to understand ‘M1 qua mental causes M2.’ What is its logical form? Is it a three-term relation among two events and a property? Is it a two-term relation whose left flanking term is ‘M1 qua mental’? What manner of relatum would that be?⁵⁰ And however we answer those questions, how exciting or robustly anti-interactionist would be the slogan ‘nothing ever causes anything unless it is a case of X qua mental causing Y qua mental or X qua physical causing Y qua physical’? I must leave the answering of these questions to others or another occasion.

I have identified two solutions to the interaction problem, one involving the hyperintensionality of causation and the other involving the hyperintensionality of ‘qua’. I am not satisfied with either as so far developed, but I believe both should be explored further before we embrace opacity as a way of upholding Spinoza’s ban on interaction.

⁴⁹ Nor need he infer 4’, M1 qua physical causes M2 qua mental, from 1’, M1 qua mental causes M2 qua mental.

⁵⁰ Matt Davidson has suggested to me that ‘qua mental’ converts the relata of the causal claim to events or facts more finely individuated than the originals. But Spinoza himself has no relata more finely individuated than events already specified as mental or physical, which he then identifies.

12. Hyperintensionality and Opacity

At this point some readers may ask, how is hyperintensionality different from or any more acceptable than opacity? Are they not of a piece? Do they both not come under the umbrella of nonextensionality?

They do both come under the umbrella of nonextensionality, but they are not of a piece. One who accepts hyperintensionality is not thereby committed to accepting opacity taken not just as the existence of opaque contexts, but as the failure of Leibniz's Law to deliver its normal results in some applications. In support of this contention, let us examine an argument for the contrary claim—an argument purporting to show that if we use Leibniz's Law, we can show that paradigmatically hyperintensional contexts are in fact intensional.

Let B be the operator 'S believes that' or any other supposedly hyperintensional operator prefixable to formulas, let N be the operator 'it is necessary that', and let corner brackets form names of propositions from formulas, $\langle p \rangle$ being the proposition that p . The argument runs as follows:

1. Assume: $Bp \ \& \ N(p \leftrightarrow q)$. To prove: Bq , which would show that B is intensional.
2. $B\langle p \rangle$ (from Bp and "propositionalization").
3. If $N(p \leftrightarrow q)$, then $\langle p \rangle = \langle q \rangle$. This is shown by the following subsidiary argument:
 - A. If $N(p \rightarrow q)$ then $\langle p \rangle$ entails $\langle q \rangle$. This is analogous to the propositionalization in 2; it takes us from a dyadic operator on formulas to a relation between terms.
 - B. If $N(q \rightarrow p)$, then $\langle q \rangle$ entails $\langle p \rangle$ (similar).
 - C. If $\langle p \rangle$ and $\langle q \rangle$ are mutually entailing, then $\langle p \rangle = \langle q \rangle$. This is what Carnap and Lewis espouse as the identity condition for propositions. (See, e.g., Lewis 1986, 55).
 - D. Therefore, 3 above.
4. $\langle p \rangle = \langle q \rangle$ (from 1 and 3).
5. $B\langle q \rangle$ (from 2, 4, and Leibniz's Law).
6. Bq (from 5 and depropositionalization). Q.E.D.

In this way, ‘S believes that’ may be shown to be intensional by using Leibniz’s Law.

But not so fast—Leibniz’s Law is not the main load-bearing member in that argument. There are two other essential assumptions in it, and I would resist them both. The first is the propositionalization appealed to in lines 2, 3A, and 3B. Like A.N. Prior, I reject an ontology of propositions, and I would parse ‘S believes that / p’ as an operator prefixed to a formula, not a relational expression involving a term on the right (Prior 1971, 16-21).⁵¹ The second assumption is the coarse-grained identity condition for propositions invoked in line 3C. If I *did* accept an ontology of propositions, I would individuate them finely like Chisholm or Frege, not coarsely like Carnap or Lewis. The Carnap-Lewis way is the intensionalist way—and it is hardly surprising that an intensionalist cannot be a hyperintensionalist! But an upholder of Leibniz’s Law who rejects intensionalism is, of course, free to espouse hyperintensionalism. Putting it the other way around, you can be a hyperintensionalist and still believe that Leibniz’s Law has no exceptions.

13. Conclusion

The problem in Part One was how Spinoza can accept Descartes’s contention that thought and extension are conceivable apart from each other without accepting Descartes’s conclusion that they never belong to the same substance. I canvassed three solutions, leaving them all as contenders, though not without obstacles: that Spinoza does not accept the Cartesian premise in the required form; that even if he does, he need not accept the Principal Attribute Thesis that needs to be allied with it; and that even if he accepted the Principal Attribute Thesis, he could let thought and extension belong to the same substance by allowing them to derive from some third attribute, as in neutral monism.

⁵¹ In Prior’s philosophy, we can have nonextensionality and even hyperintensionality without so-called intensional entities; see Prior 1968.

I discussed two problems in Part Two, of which the harder was how Spinoza can identify mental events with physical events, yet deny that there is mental-physical interaction, given that there is mental-to-mental causation and physical-to-physical causation. One neat solution is the opacity solution advocated by Della Rocca, with its supposed implication that Spinoza does not accept Leibniz's Law: M1 can cause M2 and be identical with P1, even though P1 does *not* cause M2. I argued that rejecting that law in the sweeping way Spinoza would require has heavy costs—for instance, that it would undermine the very idea of an alibi. I argued that examples supposedly contradicting Leibniz's Law do no such thing—what they really contradict is the Principle of Substitutivity, from which Leibniz's Law should be distinguished. I argued that key Spinozistic texts that seem to espouse opacity can be read instead as invoking hyperintensionality, that feature of a relation whereby it can take just one of two necessarily equivalent items as relatum. (That feature is certainly less problematic than the opacity feature, whereby a relation can take *a* as a relatum without taking *b*, the very same thing, as a relatum!) I speculated that one or another of two hyperintensional operations, *because* or *qua*, might be of use in solving the interaction problem, but was unable to bring a solution involving either of them into clear and full view. Nonetheless, I continue to hope that a solution other than the opacity solution can be found.⁵²

REFERENCES

⁵² I thank Matthew Davidson and Uriah Kriegel for helpful comments on an earlier draft.

Years of earlier publication are given in parentheses. When so noted below, my page references are sometimes to a reprinted version in a later volume.

Alter, Torin, and Yujin Nagasawa. 2015. *Consciousness in the Physical World: Perspectives on Russellian Monism*. Oxford: Oxford University Press.

Bacon, Andrew, and Jeffrey Sanford Russell. 2019. "The Logic of Opacity." *Philosophy and Phenomenological Research*, 99: 81-114.

Bayle, Pierre. (1697) 1965. *Historical and Critical Dictionary*, translated by Richard Popkin. Indianapolis, Ind.: Bobbs-Merrill.

Bennett, Jonathan. 1984. *A Study of Spinoza's Ethics*. Indianapolis, Ind.: Hackett Publishing Company.

Brandt, Richard, and Jaegwon Kim. 1967. "The Logic of the Identity Theory." *The Journal of Philosophy*, 64: 515-537.

Broad, C.D. 1925. *Mind and Its Place in Nature*. London: Routledge and Kegan Paul

Cartwright, Richard. (1971) 1987. "Identity and Substitutivity." In *Identity and Individuation*, edited by Milton K. Munitz, pp. 119-33. New York: New York University Press. Reprinted in Cartwright's *Philosophical Essays*, pp. 135-47. Cambridge, Mass.: MIT Press, 1987; my page references are to this volume.

Cartwright, Richard. (1979) 1987. "Indiscernibility Principles." *Midwest Studies in Philosophy*, 4: 293-306. Reprinted in Cartwright's *Philosophical Essays*, pp. 201-215. Cambridge, Mass.: MIT Press, 1987; my page references are to this volume.

Chalmers, David. 2002. "Does Conceivability Entail the Possibility?" In *Conceivability and Possibility*, edited by Tamar Szabo Gendler and John Hawthorne, pp. 145-200. Oxford: Oxford University Press.

Chalmers, David. 2003. "Consciousness and Its Place in Nature." In *Blackwell Guide to the Philosophy of Mind*, edited by Stephen P. Stich and Ted A. Warfield, pp. 102-142. Oxford: Blackwell.

Churchland, P.M. 1988. *Matter and Consciousness*, revised edition. Cambridge, Massachusetts: MIT Press.

Davidson, Donald. (1970) 1980. "Mental Events." In Davidson, *Essays on Actions and Events*, pp. 207-25. Oxford: Clarendon, 1980.

Delahunty, R.J. 1985. *Spinoza*. London: Routledge and Kegan Paul.

Della Rocca, Michael. (1993) 1999. "Spinoza's Argument for the Identity Theory." *The Philosophical Review*, 102: 183-213. Reprinted in *The Rationalists: Critical Studies on Descartes, Spinoza, and Leibniz*, edited by Derk Pereboom, pp.165-92. Lanham, Md.: Rowman & Littlefield, 1999.

Della Rocca, Michael. 1996. *Representation and the Mind-Body Problem in Spinoza*. Oxford: Oxford University Press.

Descartes, René. 1911. *Descartes: Philosophical Writings*, 2 vols., translated by Elizabeth S. Haldane and G.R.T. Ross. Cambridge: Cambridge University Press.

Descartes, René. 1984, 1985, 1991. *The Philosophical Writings of Descartes*, 3 vols., translated by John Cottingham, Robert Stoothoff, and Dugald Murdoch. Cambridge: Cambridge University Press. Cited as CSM 1, 2, or 3.

Dicker, Georges. 1993. *Descartes: An Analytical and Historical Introduction*. Oxford: Oxford University Press.

Donagan, Alan. 1973. "Essence and the Distinction of Attributes in Spinoza." In *Spinoza: A Collection of Critical Essays*, edited by Marjorie Grene, pp. 164-181. Garden City, N.Y.: Anchor Books.

Hume, David. (1739) 2000. *A Treatise of Human Nature*. Edited by David Fate Norton and Mary J. Norton. Oxford: Oxford University Press. Cited herein as T 1.3.3.3 for book 1, chapter 3, section 3, paragraph 3.

Kripke, Saul. 1972. *Naming and Necessity*. Cambridge, Mass.: Harvard University Press.

Lewis, David. 1986. *On the Plurality of Worlds*. Oxford: Blackwell.

Nolan, Daniel. 2014. "Hyperintensional Metaphysics." *Philosophical Studies*, 171: 149-160.

Plantinga, Alvin. 1974. *The Nature of Necessity*. Oxford: Clarendon Press.

Prior, A.N. 1971. *The Objects of Thought*. Oxford: Clarendon Press.

Prior, A.N. (1968) 1976. "Intentionality and Intensionality." In *Papers on Logic and Ethics*, edited by P.T. Geach and A.J.P. Kenny, pp. 187-201. Amherst, Mass.: University of Massachusetts Press.

Quine, Willard V. (1961) 1971. "Reference and Modality." In *From a Logical Point of View*, pp. 139-157. New York: Harper and Row. Reprinted in *Reference and Modality*, edited by Leonard Linky, pp. 17-34 (Oxford: Oxford University Press, 1971). My page references are to the Linky volume.

Rorty, Richard. 1965. "Mind-Body Identity, Privacy, and Categories." *The Review of Metaphysics* 19:24-54.

Rozemond, Marleen. (1995) 1999. "Descartes's Case for Dualism." *The Journal of the History of Philosophy*, 33: 29-63. Reprinted in *The Rationalists*, edited by Derk Pereboom, pp. 19-58 (Lanham, Md.: Rowman & Littlefield, 1999) and largely reproduced as chapter 1 of *Descartes's Dualism* (Cambridge, MA: Harvard University Press, 1998). My page references are to Rozemond's book.

Shoemaker, Sydney. 1970. "Wiggins on Identity." *The Philosophical Review*, 79:529-544.

- Shoemaker, Sydney. 1983. "On an Argument for Dualism." In *Knowledge and Mind*, edited by Carl Ginet and Sydney Shoemaker, 233-258. Oxford: Oxford University Press.
- Spinoza, Baruch. (1677) 1982. *The Ethics and Selected Letters*. Translated by Samuel Shirley. Indianapolis, Ind.: Hackett.
- Spinoza, Baruch. 1985. *The Collected Works of Spinoza*, Vol. I. Edited and translated by Edwin Curley. Princeton, N.J.: Princeton University Press. For letters, I cite this source as Curley.
- Tennant, N. 1977. "Continuity and Identity." *Journal of Philosophical Logic*, 6:223-231.
- Thomas, Emily A.E. 2018. "Samuel Alexander." *The Stanford Encyclopedia of Philosophy* (Spring 2018 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2018/entries/alexander/>.
- Van Cleve, James. 1983. "Conceivability and the Cartesian Argument for Dualism." *Pacific Philosophical Quarterly*, 64:35-45.
- Van Cleve, James. N.D. "There Are No Necessary Connections Between Distinct Existences."
- Wiggins, David. 1967. *Identity and Spatio-Temporal Continuity*. Oxford: Blackwell.
- Wiggins, David. 1980. *Sameness and Substance*. Cambridge, Mass.: Harvard University Press.
- Wilson, Margaret. (1980) 1999. "Objects, Ideas, and 'Minds'." In *Ideas and Mechanism: Essays on Early Modern Philosophy*, pp. 126-140. Princeton, N.J.: Princeton University Press, 1999.