

CANADIAN JOURNAL OF PHILOSOPHY  
Supplementary Volume 16

---

## *The Real Distinction Between Mind and Body\**

STEPHEN YABLO  
University of Michigan  
Ann Arbor, MI 48109  
U.S.A.

....it [is] wholly irrational to regard as doubtful matters that are perceived clearly and distinctly by the understanding in its purity, on account of mere prejudices of the senses and hypotheses in which there is an element of the unknown.

Descartes, *Geometrical Exposition of the Meditations*

### **I Substance Dualism**

Substance dualism, once a main preoccupation of Western metaphysics, has fallen strangely out of view; today's mental/physical dualisms are dualisms of fact, property, or event. So if someone claims to find a difference between minds and bodies per se, it is not initially clear what he is maintaining. Maybe this is because one no longer recognizes 'minds' as entities in their own right, or 'substances.' However, *selves* – the things we refer to by use of 'I'

---

\* This paper is dedicated to the memory of George Myro; an early version was read at the George Myro Memorial Conference at UC Berkeley, and a later version at the University of Arizona. Thanks to George Bealer, Jonathan Bennett, Paul Boghossian, Janet Broughton, David Copp, Sally Haslanger, Keith Lehrer, Louis Loeb, Vann McGee, Joe Mendola, Sarah Patterson, Larry Sklar, Barry Stroud, William Taschek, Bruce Thomas, David Velleman, Ken Walton, and two anonymous referees for questions, discussion, and advice. Research supported by the National Endowment for the Humanities.

– are surely substances, and it does little violence to the intention behind mind/body dualism to interpret it as a dualism of bodies and selves. If the substance dualist's meaning remains obscure, that is because it can mean several different things to say that selves are not bodies.

Any substance dualism worthy of the name maintains at least that

(1) I am not identical to my body;

and probably most dualistic arguments are directed at just this conclusion. But philosophers have been slow to appreciate how unimpressive non-identity theses can be. Assuming an unrestricted version of Leibniz's Law (the indiscernibility of identicals), non-identity is established by *any* difference in properties, however slight or insignificant. If, as seems likely, my body will remain when I am dead, then that already shows that my body and I are not the same thing; and even if my body is not going to outlast me, such could have been the case, which again gives a difference entailing non-identity. You may say that this is dualism enough. But bear in mind that analogous considerations show equally that a statue is not identical to the hunk of clay which makes it up; and this is not normally taken as grounds for a dualism of statue and clay. On pain of insignificance, self/body dualism must mean more than just the non-identity of self and body.<sup>1</sup>

What more could be at issue? For all that non-identity tells us, I might still be necessarily *realized in*, or *constituted by*, my body. For this the obvious remedy is to strengthen (1) to

(2) I could have existed without my body.

---

1 Not that this has gone entirely unnoticed. Observing that not only modal but even temporal differences 'establish that a statue is not the hunk of stone, or the conger of molecules, of which it is composed,' Kripke allows that 'mere non-identity...may be a weak conclusion' ('Identity and Necessity,' 101). That is putting it mildly. That people were not identical to their bodies was supposed to be a powerfully antimaterialistic result; but in fact it is compatible with people being as closely bound up with their bodies as statues are with the hunks of matter which compose them!

But even (2) might mean only that I could have been constituted by a different body than actually: which leaves it open that I am necessarily always constituted by some body or other (as the statue is necessarily always constituted by some hunk of matter). Only with

(3) I could have existed in the absence of all bodies (= material objects),

it seems, do we assert a difference between self and body beyond that obtaining already between statue and clay.

Implying as it does that my existence is not essentially owing to the way in which the world's matter organizes itself, (3) approaches on a genuinely challenging form of dualism. Nevertheless the ambitious dualist will want more; for the possibility remains that I *am* in an extended sense essentially embodied, in that my existence depends on there being either bodies or entities *analogous* to bodies (say, ectoplasmic entities of some sort) whose behavior gives rise to my mental life.<sup>2</sup> Functionalists, for example, can allow that I could exist unaccompanied by anything material, as long as there was *something* present with the appropriate causal organization. But it would be a strange sort of dualism which insisted on my aptitude for existing in the absence of physical bodies, only to lose interest when non-physical 'bodies' were proposed in their place.

In the spirit of Descartes, let us speak of my 'thought properties' as all and only those properties which I am directly aware of myself as possessing.<sup>3</sup> To say that I am embodied in the extended sense seems at least to say that there is an entity, my 'body,' which plays host to activities of which I am not directly aware, which activities somehow subserve my state of consciousness. Since these activities are not objects of direct awareness, they ought presumably to

---

2 For the development of this possibility, see Shoemaker, 'Embodiment and Behavior,' 'Immortality and Dualism,' and 'On an Argument for Dualism.'

3 *Thought*. I use this term to include everything that is within us in such a way that we are immediately aware of it' (CSM II, 113; AT VII, 160). More needs to be said about 'immediate awareness' to rule it out that I am directly aware, e.g., of whether my legs are crossed, but this is not a problem I take up here.

be reflected in properties which I possess in excess of my thought properties. So the truth of

(4) I could have existed with my thought properties alone,

should have the consequence that I am capable of existing not only without material things, but in a purely mental condition (i.e., without benefit of anything outside my consciousness). Indeed in a situation in which I possess my thought properties only, it would seem that I exist not just without benefit of anything outside my consciousness, but in the complete absence of any such thing. In recognition of this, we can strengthen (4) to

(5) I could have existed, in isolation, with my thought properties alone,

understood to mean that I could have existed with my thought properties alone and *in the company of no other particulars* (or at least none which are not part of me).

What more could be wanted? Notice that (4) and (5) speak only to how things *could* have been with me, not, or not directly, to how they *are*. In particular, (4) does not rule it out that *as matters stand*, I am constituted by my body, nor even that my body and I are, in the actual circumstances, exactly alike in every ordinary respect. Compatibly with (4) and (5), I might be indistinguishable from my body in point of size, shape, weight, etc., and my body might share all my feelings, thoughts, and desires.

Suppose we call a property *categorical* if its possession by a thing speaks exclusively to what it is like in the actual circumstances, irrespective of how it would, could, must, or might have been (naively, my thought properties are predominantly if not exclusively categorical, and so are most if not all of the traditional primary qualities); and *hypothetical* if it depends on a thing's liability to have been in a certain way different than it is actually (so dispositional, counterfactual, and modal properties, whether mental or physical, are hypothetical).<sup>4</sup> Then the difficulty with (4) and (5) is that they seem

---

<sup>4</sup> For more on the categorical/hypothetical distinction, see Yablo, 'Identity, Essence, and Indiscernibility.'

to express a merely *hypothetical* difference between myself and my body, whereas an ambitious dualism will want to find us *categorically* unlike. Either I do not possess my body's categorical physical properties, like that of taking up space; or my body does not possess my categorical mental properties, like that of experiencing pain; or both.

Beware of taking the point too far; no reasonable dualist believes that I have *no* categorical physical characteristics, or that my body has *no* categorical mental properties. Obviously we do. Even if I do not occupy space myself, I do have the physical property of coexisting, and presumably interacting, with something which does (my body); and my body, though perhaps not itself experiencing pain, coexists, and interacts, with something in which pain authentically resides (myself). Thus the claim must be that my categorical physical properties, and my body's categorical mental properties, are always *extrinsic* (*P* is *intrinsic* to *x* if *x*'s possession of *P* speaks exclusively to what *x* is like in itself, without regard to what may be going on outside of *x*, and *extrinsic* otherwise). From this it is a short step to

(6) All of my intrinsic, categorical, properties are mental rather than physical,

and

(7) All of my body's intrinsic, categorical, properties are physical rather than mental.

Assuming that my intrinsic, categorical mental properties are exactly my *thought* properties, the relation between (4) and (6) is as follows: where (4) postulates a counterfactual condition in which I exist with just my thought properties, (6) says that my *actual* condition is in all *intrinsic, categorical*, respects indiscernible from that counterfactual condition of pure disembodiment.

No doubt the exercise could be taken further. For example, (6) and (7) are somewhat overstated. Even the most extreme dualist will admit that she has (e.g.) her existence, and her duration, intrinsically; and these are not plausibly regarded as mental properties. But this is not something we need to bother about just now (see note 15). Another thing we will be leaving aside is the articulation of still

stronger versions of dualism, for example the necessitations of (6) and/or (7).<sup>5</sup> What I want to ask now is whether dualism in any of these forms, but especially the fourth, fifth, sixth, and seventh, has any chance of being true.

Subject to correction by Descartes scholars, most of us suppose that Descartes maintained dualism in all the versions given. Unfortunately, his principal argument is nowadays seen as bordering on hopeless, and this on the basis of a single apparently decisive objection, roughly to the effect that de re conceivability is a defective guide to de re possibility.

In this paper, I want to pursue two ideas. The first is that Descartes' argument cannot be faulted simply for relying on an inference from de re conceivability to de re possibility; that inference is implicated in too many de re modal claims routinely accepted without qualm or question. So the standard objection needs refinement: even if some de re conceivability intuitions justify de re modal conclusions, others do not, and when the differences are spelled out, Descartes' argument emerges as unpersuasive. The paper's second idea is that, to the contrary, the more the differences are spelled out, the better Descartes' argument looks.

## II Standard Problems with Descartes' Arguments

Descartes believed that he was importantly different from his body, and offered what looks like a variety of arguments for this conclu-

---

5 Obviously I disagree with Bernard Williams when he says that it 'expresses the Real Distinction in its strongest form' to assert the necessitation of (1), i.e., to say that I am necessarily not identical with my body (*Descartes*, 117). Assuming that Leibniz's Law holds necessarily, the same can be said of a statue and the hunk of clay which makes it up; for necessarily the one has different modal properties, e.g., being essentially a statue, than the other. Since Kripke, most metaphysicians treat (non-)identity theses as equivalent to their necessitations; if they are right, then what Williams calls the strongest form of the real distinction is actually the *weakest* (or equivalent to it). Certainly it is far weaker than the claim that necessarily self and body have fundamental categorical differences (this is the necessitation of [6] and/or [7]).

sion. Some of these are less plausible than others. In *The Search After Truth*, there are indications of the much ridiculed 'argument from doubt': I am not a body, 'otherwise if I had doubts about my body, I would also have doubts about myself, and I cannot have doubts about that' (CSM II, 412; AT X, 518). Since I can doubt that my body exists, but not that I do, I am distinct from my body.

Whether Descartes intended precisely this argument or not, it is plainly fallacious, on any readily imaginable interpretation. Perhaps Descartes is reasoning as follows:

*Argument A*

- (1) I can doubt that my body exists, but not that I do. (A)
- (2) Therefore my body and I have different properties. (1)
- (3) Therefore I am not identical with my body. (2)

However, (2) follows from (1) only if 'I can doubt that  $x$  exists' expresses a property of  $x$ ; which, to judge by its admitted referential opacity, it appears not to do.

On the road to Descartes' true argument is a reading which replaces doubt with rational doubt:

*Argument B*

- (1) It is not irrational for me to doubt that my body exists while believing that I do. (A)
- (2) If I was identical to my body, this would be irrational. (A)
- (3) Therefore I am not identical to my body. (1,2)

Again, there is a problem with the second step. Even if my self and body are identical, reason does not constrain me from feeling doubts about my body which I am unwilling to extend to myself, if I am unaware of their identity, and unaware more generally that it is impossible for the one to exist without the other. Before I can draw any conclusions from the rational permissibility of doubting body but not self, I need assurances that my essential properties cannot but make themselves felt in my self-conception. Without these assurances, that I am not irrational in maintaining contrasting attitudes

toward self and body is as likely due to my ignorance of my true nature as to anything else. Yet if the assurances are somehow obtained, then I already *have* my conclusion and the argument is no longer needed. For if I am unaware of being essentially accompanied by my body, then I am not; and so we are distinct.

Even if (as is sometimes alleged) the argument from doubt cannot fairly be attributed to Descartes, his other and more canonical arguments for the mind/body distinction appear to incorporate a similar fallacy. Thus the crucial assumption of the 'Sixth Meditation's dualistic argument is that

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. (CSM II, 54; AT VII, 78)

Since I can understand, or conceive, myself clearly and distinctly apart from my body, I and my body 'are capable of being separated'; hence we are not identical. As an initial guess about what is going on here, consider:<sup>6</sup>

---

<sup>6</sup> In interpreting the quoted passage, I follow the usual practice of disallowing any essential role to God's omnipotence. If we are to take seriously Descartes' doctrine of God's free creation of the eternal truths, God can create anything apart from anything, even  $x$  apart from  $x$ ; and this without regard to what we may or may not find conceivable. Since that doctrine renders irrelevant conceivability considerations which Descartes clearly sees as crucial, and lends itself to the derivation of conclusions much stronger than he would accept, there is no option but to discount it in the present context. Having done so, the divine power to create  $x$  without  $y$  essentially converges on the metaphysical possibility of  $x$  without  $y$ . (Cf. Descartes' remark in the 'Geometrical Exposition of the Meditations' that '...I introduce the power of God as a means to separate mind and body not because any extraordinary power is needed to bring about such a separation but because the preceding arguments have dealt solely with God, and hence there was nothing else I could use to make the separation' [CSM II, 120; AT VII, 170], and in the 'Sixth Replies' that 'to occur "naturally" is nothing other than to occur through the ordinary power of God, which in no way differs from his extraordinary power – the effect on the real world is exactly the same' [CSM II, 293; AT VII, 435].)



*Argument C*

- (1) I can conceive myself as existing without my body. (A)
- (2) If I can conceive  $x$  as existing without  $y$ ,  
 $x$  can exist without  $y$ . (A)
- (3) So it is possible for me to exist without my body. (1,2)
- (4) So I am not identical to my body. (3)

Before asking what might be wrong with this argument, notice an important respect in which it improves on the argument from doubt. All that that argument can hope to establish is that I am not identical to my body. But this goes hardly any distance towards justifying the grand claims of Descartes' dualistic metaphysics: that I am capable of existing without my body, that I am capable of existing without *any* body, that I am unextended, and so on. Although argument (C) terminates in the non-identity thesis, it reaches it by way of the significantly *stronger* thesis that I am capable of existing without my body (and it would not significantly detract from the argument's plausibility if instead of 'my body,' we had written throughout 'any body'). So if it could be made to work, this argument might yield a dualism worth bothering about.

Nevertheless it seems not to work, and for essentially the same reason as before. According to (2), if I can conceive  $x$  as existing without  $y$ , then it *can* exist without  $y$ . But this is plausible only if I can be sure that I am not, in this act of conception, overlooking an essential property of  $x$  which renders its existence without  $y$  problematic or impossible. As Sydney Shoemaker expresses the point, the argument

...involves a confusion of a certain sort of epistemic possibility with metaphysical possibility. In the sense in which it is true that I can conceive myself existing in disembodied form, this comes to the fact that it is compatible with what I know about my essential nature (supposing that I do not know that I am an essentially material being) that I should exist in disembodied form. From this it does not follow that my essential nature is in fact such as to permit me to exist in disembodied form.<sup>7</sup>

---

<sup>7</sup> 'Immortality and Dualism,' 155

Absent prior assurances that his potential for independent existence is not obstructed by unappreciated necessary connections, Descartes is in no position to argue from separability in thought to separability in fact.

Because of difficulties like these, not many philosophers would concede Descartes' claim to have established even so much as his distinctness from his body, much less any *interesting* form of dualism. The problem with Descartes' approach is supposed to be one of principle rather than detail, with the result that most philosophers would now be gravely suspicious of *any* epistemic argument for dualistic conclusions.

### III The Indispensability of Conceivability

Then what kind of argument *is* available to the dualist? Encouraged by recent advances in modal semantics and metaphysics, modern dualists prefer to base their conclusions in *modal* rather than epistemic premises.

No doubt this is an advance of some sort, but it has worrisome aspects. For one, it ignores that the modal premises stand themselves in need of support, which typically they find in conceivability considerations of the sort that Descartes is faulted for having taken seriously. Insofar as they suppress the role of conceivability in modern-day modal arguments, today's dualists let themselves off the hook on which they hoisted Descartes. Second, once the indispensability of conceivability intuitions is allowed, explanations will be required of how it is that *some* such intuitions may be relied on, even if others cannot. Thus grant that the ancients' ability to conceive (say) heat without motion should not have been taken, even by them, to establish that this was possible. Even so, that I can conceive of *myself* existing without the Washington Monument, *does* seem *prima facie* to indicate that the one could have existed without the other (or else how do I know that it could?). Presumably there are some unobvious principles at work here that would explain why the one intuition may be relied on, though the other may not. And so far, nothing rules out that when the operative principles are discovered, Cartesian conceivability intuitions will be vindicated.

As already explained, the usual charge against Descartes' argument from his ability to conceive  $x$  as existing without  $y$ , to the conclusion that  $x$  can exist without  $y$ , is that it seems just to take it for granted that  $x$ 's essential properties do not go beyond those of which Descartes is aware. Objections of this kind were put to Descartes repeatedly, most notably by Caterus in the *First Objections* and by Arnauld in the *Fourth*. Arnauld asks,

How does it follow, from the fact that he is aware of nothing else belonging to his essence, that nothing else does in fact belong to it? (CSM II, 140; AT VII 199)

complaining that

if the major premise of this syllogism [that the conceivability of  $x$  without  $y$  shows the possibility of  $x$  without  $y$ ] is to be true, it must be taken to apply not to any kind of knowledge of a thing, nor even to clear and distinct knowledge; it must apply solely to knowledge which is adequate. (CSM II, 140; AT VII, 200; interpolation mine),

where here 'adequate knowledge' of a thing is knowledge that embraces all the thing's properties (or at least all its essential properties).<sup>8</sup>

Undeniably this looks like an extremely strong objection, maybe even decisive. How wonderful then that Descartes had the chance to hear it and respond. But before looking at what he says, it's important to see that the problem, if there is one, is extremely general. To be consistent, Arnauld should hold that *all* de re conceivability intuitions are suspect, unless the ideas employed are certifiable in advance as adequate, i.e., as embracing all properties, or at least all essential properties, of their objects. What is not often noticed is that if he is right in this, then an enormous part of our de re modal thinking falls under suspicion.

---

<sup>8</sup> Certainly this is how Descartes read Arnauld's use of 'adequate,' and most modern commentators have agreed. However, true Arnauldian adequacy may be a subtler affair than Descartes appreciated (see Bruce Thomas, 'Conceivability and the Real Distinction').

Distinguish two types of *de re* modal claim: *positive* claims, to the effect that something *x* has a property *Q* essentially; and *negative* claims, to the effect that something *y* has a property *R* only inessentially or accidentally. Naturally it is the positive claims which have attracted all the attention (e.g., natural kinds have their deepest explanatory features essentially, artifacts have their original matter essentially, etc.). But it is sometimes just as important if something has a property only accidentally (if, for example, people have their personalities, or their genders, only accidentally); and even where it is not important, it is often true, and often, apparently, *known* to be true. No one would doubt of herself that (e.g.) she could have been born on a different day than actually; and outside of philosophy, no one would question that we know such things. But how do we know them, if not by way of conceiving ourselves without the relevant properties, and finding no difficulty in the conception?

What gives this question its force is the specter of an Arnauldian skeptic who argues from the possible inadequacy of my self-conception, to the conclusion that I am in no position to rule out even such obviously absurd essentialist hypotheses as that I am essentially born on September 30, 1957. If I might, unbeknownst to myself, be essentially accompanied by my body, however clearly I seem to be able to conceive myself without it, why might I not equally be essentially born on that day, however clearly I seem to be able to conceive myself born a day earlier or later? In both cases, the skeptic continues, I have no basis to question the deviant hypotheses unless I have prior assurances that my self-conception embraces all my essential properties. Yet how could I?

In a curious way, this sort of objection reverses a more familiar challenge to *positive de re* modal claims. Suppose I assert that something *x* has some property *Q* essentially, e.g., that this bit of water essentially contains hydrogen. Of course, I might be wrong in supposing that this, or any, water contains hydrogen at all. But now I am interested in the allegation that I might be wrong in another way: I am right that this, like all, water *actually* contains hydrogen, but wrong that it could not have been hydrogen-free. In possible worlds very like this one, it is agreed, it does contain hydrogen; but it is alleged that there may also be worlds in which it contains only oxygen and helium, and yet other worlds in which it contains only helium and aluminum, or helium and aluminum and lead.

Naturally you complain that no grounds have been given for thinking this possible; but then no grounds have been given for thinking it impossible either, and the claim was only that it was possible for all you know. After all, once you have picked *x* out, what essential properties it has is no longer in your hands, but depends entirely on what sorts of counterfactual changes *x* can as a matter of objective modal fact tolerate. How could anything in your way of conceiving *x* rule out that the thing *in itself* is capable of more extreme departures from its actual condition than you had imagined?

Postpone for now the question whether this is a cogent thought, and notice the parallel with Arnauld. Where the present objection is that one cannot rationally exclude that the object of thought has *fewer* essential properties than contemplated, Arnauld contends that one cannot rationally exclude that it has *more* essential properties than contemplated. To answer either objection would be to explain what licenses us in reasoning from premises about what we can conceive of a thing to conclusions about what is possible for it. But let us concentrate on the Arnauldian worry that what I seem able to conceive regarding *x* provides no firm basis for *excluding* properties from *x*'s essence.

Actually, there is a certain irony in Arnauld's position. Leibniz, in his correspondence with Arnauld, alleges that the essence of a thing *x* embraces all of *x*'s properties whatsoever. Since Adam is such that Peter denied Christ some thousands of years after his death, this holds *essentially* of Adam, who would accordingly not have existed had Peter not gone on to be disloyal:

if in the life of some person and even in this entire universe something were to proceed in a different way from what it does, nothing would prevent us saying that it would be another person or another possible universe that God would have chosen. *It would thus truly be another individual....* (LAC, 60, my emphasis)

Unsurprisingly Arnauld objects:

...I find in myself the concept of an individual nature, since I find there the concept of myself. I have only to consult it, therefore, to know what is contained in this individual concept....I can think that I shall or shall not take a particular journey, while remaining very much assured that neither one nor the other will prevent my being myself. So I remain very much assured that neither one nor the other is included in the individual concept of myself.... (LAC, 32-3)

Within limits, it seems obvious, we share Arnauld's assurance. Nobody seriously imagines that it is essential to Arnauld to take, or essential to him not to take, the journey. Still it is hard to see what entitles *him* to the assurance that 'neither one nor the other will prevent me from being myself.' How does Arnauld know that his idea is adequate, i.e., that he is aware of *all* of his essential properties?<sup>9</sup>

Take the Arnauldian skeptic to be the one who questions Descartes' right to reason from separability in conjecture to separability in fact, on the basis that our concepts may for all we know be inadequate; and take the Arnauldian believer to be the one who maintains, against Leibniz, that properly conducted thought experiments can support *de re* inessentialist conclusions. If the skeptic's doubts are allowed to stand, then it is not obvious how the believer can hope to refute Leibniz's suggestion that my essence takes in all my properties whatsoever! Yet surely we side here with the believer. Even without an answer to the skeptic, I think we feel that that he *must* be wrong. Somehow or other, I *must* be in a position to refute the suggestion that I am essentially born on the day of my actual birth, or, even more unbelievably, essentially surrounded by the entire course of actual history.

#### IV The Conceivability Argument

What I want to investigate is whether Descartes had even the beginnings of an answer to the Arnauldian skeptic. For this the natural starting point is Descartes' historical controversy with Arnauld, which centers on the conceivability/possibility principle that

---

<sup>9</sup> To complete the irony, something uncomfortably like this Arnauldian point is put to Arnauld by Leibniz himself: '...although it is easy to judge that the number of feet in the diameter is not contained in the concept of a sphere in general, it is not so easy to judge with certainty...whether the journey which I plan to take is contained in the concept of me, otherwise it would be as easy to be a prophet as to be a geometer...' (LAC, 59). Leibniz thinks that individual concepts frameable by finite minds are rarely adequate, much less certifiably adequate. From this it seems to follow that we must view all our conceivability intuitions with extreme suspicion. Nevertheless, Arnauld confidently asserts that he knows that he might not have taken the journey.

## *The Real Distinction Between Mind and Body*

If I can conceive of  $x$  as lacking some property  $S$ , then it is possible for  $x$  to exist without  $S$ .

For such a principle to be valid, Arnauld thinks, it 'must be taken to apply not to any kind of knowledge of a thing, nor even to clear and distinct knowledge; it must apply solely to knowledge which is adequate' (CSM II, 140; AT VII, 200). In response, Descartes appears willing to grant that the mere conceivability, even the clear and distinct conceivability, of  $x$  as lacking some property  $S$  is not itself convincing evidence of  $S$ 's inessentiality. As Arnauld suggests,  $x$  must be conceived in a suitably comprehensive manner:

a real distinction cannot be inferred from the fact that one thing is conceived apart from another by an abstraction of the intellect when it conceives the thing inadequately. It can be inferred only if we understand one thing apart from another completely, or as a complete thing. (CSM II, 155, AT VII, 220)

(Cf. also CSM II, 86; AT VII, 121.) But in Descartes' view, Arnauld is wrong to think that our conception needs to be certifiable in advance as 'adequate' (CSM II, 155; AT VII, 220). Admittedly, he may have given a contrary impression when he said that a real distinction could not be inferred by 'an abstraction of the intellect when it conceives a thing inadequately'; but he

did not think this would be taken to imply that *adequate* knowledge was required.... All I meant was that we need the sort of knowledge that we have not ourselves made *inadequate* by an abstraction of the intellect. (CSM II, 155-6; AT VII, 221)

To the question, 'what manner of conception is required if we are to be able to rely on the inference from conceivability to possibility?' Descartes therefore answers that we should conceive  $x$  'completely, or as a complete thing'; to which it appears to be a corollary that our conception of  $x$ , even if not adequate in Arnauld's sense, is free at least of that specific type of inadequacy engendered by intellectual abstraction.

In his day as in our own, Descartes' readers have sensed a confusion in his writings between (i) a conception of myself *in which I do not credit myself with corporeal features*, and (ii) a conception of myself *as lacking in corporeal features*. Sometimes it is said that only the

former conception is claimed by, or even available to, Descartes; though it is the latter he needs to argue for the possibility of disembodiment. But Descartes could hardly be clearer that he possesses a self-conception of type (ii); and his repeated insistence on the importance of 'complete conception,' and the avoidance of 'abstraction,' is, as we will see, directed against just the confusion to which he is so often thought to have succumbed.

To conceive something in a complete manner, Descartes explains, he 'must understand the thing well enough to know that my understanding is *complete*'; and his understanding of a thing *x* is called 'complete' if and only if he understands *x* 'to be a complete thing' (CSM II, 156; AT VII, 221). On its face, this could hardly look less enlightening; but let us pursue it. In general, Descartes calls a thing *complete* if and only if it is a *substance*, that is, it is capable of existing on its own (or, since nothing can exist without God's concurrence, capable of existing unaccompanied by anything but God).<sup>10</sup> Intriguingly, though, he here gives a more elaborate explanation, in which epistemological considerations come strikingly to the fore:

... by a "complete thing" I simply mean a substance endowed with the forms or attributes which *enable me to recognize that it is a substance*. (CSM II, 156; AT VII, 221; emphasis mine)

From this it appears that a complete thing is a substance taken together with a set of its properties meeting some further epistemological condition. And the condition is, that those properties should enable him to recognize their bearer as a substance.

Initially, at least, this is extremely puzzling. In Descartes' view, substances are never directly apprehended, but only by way of their properties (CSM II, 124; AT VII, 176); and whenever we apprehend

---

10 Cf., CSM I, 210, AT VIII A, 25, for the problem about God. Even putting that problem to one side, the definition of substance in terms of capability for unaccompanied existence is still misleading, since a substance is always accompanied by its primary attribute and modes thereof. The natural remedy is to define a substance as an entity which can exist without other substances; however, that would be circular. Some have suggested using another Cartesian notion of substance, that of property-bearer, to give a non-circular definition of the first notion (cf. Loeb, *From Descartes to Hume*, 94).



a property, we may infer that there is a substance in which it inheres (CSM I, 210; AT VIIIA, 25). So when Descartes speaks of 'forms or attributes which enable me to recognize that it is a substance,' he cannot, on pain of triviality, mean simply 'forms or attributes which convince me that there is a substance *about*' (all properties do that much). Instead, the properties with which the substance is to be thought of as endowed should *present* to me the substance in a way that allows me not merely to recognize that a substance is there, but also *that it is a substance*. Since to be a substance is to be capable of solitary existence, the obvious thought is that *x* is recognizable *as* a substance, if and only if it is presented by way of properties which reveal to me *how it is that x is capable of existing by itself*. In other words, the properties by which *x* is presented are such that *I find it intelligible that it should exist with those properties alone, in the absence, specifically, of any further properties such as might require the existence of some other substance*. If and only if *x* is thus presented, do I conceive it in a complete manner, or as a complete thing.<sup>11</sup>

Separability in conjecture does not argue for separability in fact if 'one thing is conceived apart from another by an abstraction of the intellect which conceives the thing inadequately... [but] only if we understand one thing apart from another completely, or as a complete thing' (CSM II, 155; AT VII, 220). Thus complete conceptions 'need the sort of knowledge that we have not ourselves made *inadequate* by an abstraction of the intellect' (CSM II, 156; AT VII, 221).<sup>12</sup> Intellectual abstraction is explained in a letter to Gibieuf; it

---

11 Understanding complete things in this way sheds some light on Descartes' otherwise enigmatic remarks about 'incomplete substances,' e.g., 'a hand is an incomplete substance when it is referred to the whole body of which it is a part; but it is a complete substance when it is considered on its own. And in just the same way the mind and the body are incomplete substances when they are referred to a human being which together they make up. But if they are considered on their own, they are complete' (CSM II, 157; AT VII, 222). On the other hand, pursuing the Aristotelian resonances of this and similar passages, one might well arrive at a richer notion of 'complete thing' than that suggested here, e.g., entity with an 'internal principle of activity' (see, for example, *Metaphysics* VII.10 and *De Anima* II.1).

12 In these remarks about abstraction, I am greatly indebted to Bruce Thomas's 'Abstraction and Complete Things.'

...consist[s] in my turning my thought away from one part of the contents of [a] richer idea the better to apply it to another part with greater attention... I can easily recognize this abstraction afterwards when I look to see whether I have derived the idea... from some richer idea within myself, to which it is joined in such a way that although one can think of the one without paying any attention to the other, it is impossible to deny one of the other when one thinks of both together. (K, 123)

Abstraction, then, consists in *prescinding* from some aspect of an idea, such that one cannot *deny* the ignored aspect 'when one thinks of both together.' Thus it is important that Descartes thinks that he can avoid this with the ideas of himself and his body:

If I said simply that the idea which I have of my soul does not represent it to me as being dependent on a body... this would be merely an abstraction, from which I could form only a negative argument, which would be unsound. But I say that this idea represents it to me as a substance which can exist even though everything belonging to body be excluded from it; from which I form a positive argument, and conclude that it can exist without the body. (K, 152)

Evidently Descartes sees the reliability of his modal intuition as hinging on his avoidance of abstraction in favor of exclusion; and, as we know, he attaches a similar significance to his employment of a complete idea of self. Unsurprisingly, then, the completeness of his self-conception as a thinking thing is strongly associated with his ability to exclude his bodily aspects therefrom:

...the idea of a substance with its extension and shape is a complete idea, because I can conceive it alone, and deny of it everything else of which I have an idea. Now it seems to me very clear that the idea which I have of a thinking substance is complete in this sense, and that I have in my mind no other idea which is prior to it and joined to it in such a way that I cannot think of the two together while denying the one of the other; for if there was any such within me, I must necessarily know it. (K, 124)

(Cf. also K, 109.) So when Descartes tells us that in conceiving himself as a thinking thing, his idea of himself is complete, he means (at least) that he is capable not only of *prescinding* from thoughts of body in conceiving of himself, but of conceiving himself as *lacking* in bodily aspects.

Now we should ask, exactly how is this supposed to contribute to the reliability of Descartes' modal intuition? Abstraction is not,

for Descartes, always and everywhere a bad thing. In *Rules for the Direction of the Mind*, he emphasizes the beneficial effects of freeing our conception of a question 'from every superfluous conception' (CSM I, 51ff.; AT X, 430ff.). Nevertheless, abstraction can sometimes lead us astray. Indeed in its most extreme form, where one prescind in thought from *all* the attributes by which a thing is recognized, abstraction is always problematic. Since 'we do not have immediate knowledge of substances,' prescinding in thought from all of a thing's properties leaves us without any proper grasp of what it is that we are thinking about (CSM II, 156; AT VII, 222).

To avoid extreme abstraction, we must conceive our object in terms of some suitable selection of its properties; presumably *which* properties depends on the nature of the investigation. Then what if the investigation is into what is possible for a thing? Given Descartes' rejection of the Arnauldian adequacy requirement, not *all* the thing's properties are needed. But it would seem that we *do* risk a problematic act of abstraction if we prescind in thought from such, or so many, properties that our object cannot be understood as lacking the properties prescinded from (CSM II, 276-7; AT IXA, 216). For this might tempt us into thinking that *x* could exist with no properties other than those included in our conception, when in fact the hypothesis of *x* without those further properties was not fully intelligible. In some such cases, the distinction between *x* and some particular omitted property is merely 'conceptual':

*a conceptual distinction is a distinction between a substance and some attribute of that substance without which the substance is unintelligible ....Such a distinction is recognized by our inability to form a clear and distinct idea of the substance if we exclude from it the attribute in question.... (CSM I, 214; AT VIIIA, 30)*

In others, one assumes, what is 'unintelligible' is not *x* without some particular omitted property (e.g., the wax without extension), but *x* as lacking each of a class of omitted properties (e.g., the wax with no particular shape). Quite generally, though, the complete conceiver must take pains not to exclude from her conception of a thing such, or so many, properties that the thing is 'unintelligible' without them. Drawing on the discussion above, we take this to mean that one avoids problematic abstraction by thinking of *x* in terms

of properties such that the supposition of its existing with them alone is not repugnant to reason.

Avoidance of abstraction, so understood, is necessary, but not quite sufficient, for complete conception. Remember that complete conception requires knowledge of a thing sufficient to let us know that it is *complete*, and a complete thing is described as 'a substance endowed with the forms or attributes which enable me to recognize that it is a substance' (CSM II, 156; AT VII, 221). Thus complete conception additionally requires that the possibility of *x*'s possessing the indicated properties alone reveals it *as* a substance, i.e., as something that can exist on its own. Gathering these threads together, *x* is conceived as a complete thing, if and only if by way of properties **P** such that

*Containment Condition:* *x* is clearly and distinctly conceivable as possessing the properties in **P** to the exclusion of all others.

*Isolation Condition:* For *x* to possess the properties in **P** to the exclusion of all others is for *x* to exist alone (so that its capability to possess the **P** properties exclusively shows that *x* is a substance).<sup>13</sup>

To *be* a complete thing is accordingly to be a substance *x* taken together with properties **P** in terms of which it is completely conceiv-

---

13 Notice how this account preserves the distinction, on which Descartes so much insisted (CSM II, 155-6; AT VII, 220-1), between understanding something adequately, that is, in terms of 'absolutely all the properties which are in the thing,' and understanding it completely. To understand *x*, qua **P**, in a complete manner, is not to know *everything* about it, but only enough so that, at least from the subjective perspective, *x* does not appear to need more than what you know about in order to exist. Thus if adequate ideas embrace all of a thing's properties, then complete ideas need not be adequate. (From the definition of completeness it admittedly follows that, at least from the subjective perspective, no property outside of **P** is essential to *x*, so that **P** will *seem* to give an upper bound on the set of *x*'s essential properties. If to be adequate an idea needs only to include *x*'s essential properties, then a complete idea of *x* will at least appear to the thinker to be adequate. Notice though that the thinker need not yet have any views about which of the properties in **P** are essential to *x* and which accidental.)

able (there is no distinction between being completely conceivable in terms of P, and being complete, qua possessor of P).

Applying this account to the case of interest, to conceive *myself* as a complete thing is to conceive myself in terms of a set P of properties such that I am clearly and distinctly conceivable as possessing P alone, where to exist with P alone is to exist unaccompanied by any other substance.

Does Descartes think that he can conceive himself as a complete thing in this sense? Indications are that he does think that he can do this, by conceiving himself in terms of what I have called his *thought properties*. Indeed, I suggest that he finds, in the fact that he conceives himself, qua possessor of his thought properties, as a complete thing, all that he needs to reach the conclusion that he could have existed, in isolation, with his thought properties alone. Assuming that by 'that of which I am aware,' he means his thought properties, Descartes indicates by his statement that

...it may be that there is much within me of which I am not yet aware... [but] that of which I am aware is sufficient to enable me to subsist with it and it alone.... (CSM II, 155; AT VII, 219; emphasis added)

his satisfaction that his idea of himself as thinking thing meets the containment condition on complete conception.<sup>14</sup> On no further basis than this, he concludes that

---

14 Although there may be a question whether Descartes is fully *consistent* in finding himself clearly and distinctly conceivable as possessing his thought properties exclusively. For his idea of God, together with certain principles revealed by the 'natural light,' proves God's existence as a non-deceiver, which implies in turn the reliability of his senses; whence his experience as of material objects outside himself guarantees their existence. But then how can he clearly and distinctly conceive himself with his actual thought properties, but *without* the properties that he possesses in virtue of his relations to the external material objects which sense reveals, e.g., the property of having a body? Perhaps the answer is that he conceives himself with no other *intrinsic* properties than his actual thought properties (additional *extrinsic* properties are allowed); or that he conceives himself in sole possession of thought properties *other* than those he possesses in actuality. But there is little textual basis for either sug-

I am certain that I could have been created by God without having these other attributes of which I am unaware. (CSM II, 155; AT VII, 219)

In other words, God could have created him with his thought properties alone. Since he finds nothing in his thought properties to suggest the existence of any other substance (CSM I, 213; AT VIIIA, 29), circumstances in which he has them 'without these other attributes of which I am unaware' will be circumstances in which he exists in isolation (this is the isolation condition). Hence he is entitled to conclude that he can exist, in isolation, as a purely thinking thing. And this completes the argument.

*Argument D*

- (1) Qua possessor of my thought properties,  
I am a complete thing. (A)
- (2) I am clearly and distinctly conceivable as possessing  
my thought properties to the exclusion of all  
other properties. (1)
- (3) If  $x$  is clearly and distinctly conceivable as  
possessing exactly the **P** properties, then  $x$   
can exist with exactly the **P** properties. (A)
- (4) I can exist with exactly my thought properties. (2,3)
- (5) For me to exist with exactly my thought  
properties is for me to exist in isolation. (1)
- (6) I can exist, in isolation, with exactly my  
thought properties. (1,3)

Here (1) is the claim of completeness, (2) and (5) are the containment and isolation conditions on complete conception, and (3) is the conceivability/possibility principle by which Descartes hopes to infer his aptitude for solitary mental existence from his thinkability in that condition.

---

gestion, and both sit poorly with the quoted passage (among others). Thus Descartes' dualistic arguments and his antiskeptical arguments appear to be in some tension; and I am forced to ignore the latter in favor of the former.

## V Thinking Things As Complete Things

Evidently argument (D) is formally valid, so its soundness depends on the acceptability of its premises: the claim (1) that I am, qua possessor of my thought properties, a complete thing, and the conceivability/possibility principle (3) which enables me to conclude, on that basis, that I can exist with my thought properties to the exclusion of all others.

To say that I am, qua possessor of my thought properties **T**, a complete thing, is to make two claims: that I am clearly and distinctly conceivable as possessing the properties in **T** to the exclusion of all others; and that to possess the properties in **T** to the exclusion of all others is to exist in isolation. Now the second of these claims is extremely plausible. If I am not isolated, then there is something *y* outside myself, in virtue of my relations to which it seems inevitable that I should possess properties in excess of my thought properties. But the first claim raises, to begin with, an interesting technical difficulty of which Descartes may not have been explicitly aware.

Is it really conceivable that I should possess my thought properties to the exclusion of *all* others?<sup>15</sup> If we understand the word 'property' so that the class of properties is closed under complementation, then nothing *x* can have the properties in a set **P** to the exclusion of all others, unless for *each* property *S*, **P** contains either *S* or its complement not-*S* (*proof*: if it contains neither, then *x* possesses neither, which is absurd). Yet when Descartes claimed he could have the properties of which he was aware but 'without ... these other attributes of which I am unaware,' he certainly did not suppose that for every property *S*, he was aware of himself either as possessing *S*, or as possessing not-*S* (e.g., he didn't think of himself either as extended, or as unextended). For present purposes, then, Descartes would not, or should not, have understood the set of properties as

---

15 Immediately one sees that what Descartes called the 'transcendental' or 'common' attributes (existence, duration, unity, etc.) will have to be allowed as exceptions. For I am not readily conceivable as, e.g., lacking duration. Henceforth, 'property' means non-transcendental property.

closed under complementation. As it happens, he observed a distinction, between *positive* and *negative characteristics*, or *genuine properties* and *mere privations*, which will secure the needed result, if in the definition of a complete thing we read 'property' as signifying genuine properties only.<sup>16</sup>

Not to minimize its difficulties, several things may be said in defense of the revised definition of a complete thing. For one, it relies on a distinction which is, for all its obscurities, important to Descartes, both in his metaphysics (the cosmological proof of God's existence) and in his epistemology (his doctrine of simple natures and materially false ideas). Second, what Descartes is looking for in a complete thing is a substance fitted out with properties sufficient to render it 'intelligible' as a self-standing entity; and intelligibility is aided not by the accumulation of negative characteristics, but of positive. Third, the old definition leads to results which Descartes clearly does not intend. Consider the negative characteristic *U* of being unextended: since *U* is not a member of *T*, to conceive myself as possessing *T* exclusive of all other characteristics is to conceive myself as lacking *U*, and thereby as possessing corporeal properties after all! Fourth, that Descartes never himself contemplates conceivability arguments which trade on negative characteristics such as *U*, suggests that he implicitly understood completeness in terms of positive characteristics. Fifth and lastly, by restricting ourselves to positive characteristics in the definition of a complete thing, we do not limit the definition's generality so much as lessen its redundancy. Let *S* be positive, so that not-*S* is negative; then whatever not-*S* might have accomplished by its presence in *P*, is accomplished anyway by *S*'s (presumed) absence. So much, at any rate, is to the credit of the revised definition. On the minus side, the revised definition inherits all the obscurity of the distinction between positive and negative characteristics. But let us see where it takes us.

---

16 Specifically, if *P* is a set of genuine properties (= positive characteristics), then *x* is a complete thing, qua possessor of *P* iff (a) *x* is clearly and distinctly conceivable as possessing the (genuine) properties in *P* to the exclusion of all other (genuine) properties; and (b) only if it possesses (genuine) properties beyond those in *P* can *x* fail to be alone.



Somewhat tentatively, I propose that to conceive it as possible that  $p$  is to enjoy the appearance that  $p$  is possible, by intellectually envisaging a more or less determinate situation in which  $p$  is understood to obtain.<sup>17</sup> Clarity and distinctness come in as follows: I conceive  $p$ 's possibility *clearly* in proportion as I possess a comprehensive, explicit, and determinate, intellectual vision of what the contemplated situation is like, and how it verifies the condition that  $p$ ; and I conceive it *distinctly* in proportion as whatever is not contemplated as pertaining to the envisaged situation may consistently be understood *not* to pertain (equivalently, nothing which is not contemplated as pertaining is rationally required by factors which are contemplated as pertaining).

Assuming that my conception of a situation in which I exist in a purely mental condition is not *manifestly* incoherent, the role of distinctness is to show that it harbors no *latent* incoherence, i.e., nothing that would generate manifest incoherence if its consequences were followed out; and the role of clarity is to show further that the conception is free of saving unspecificities which, however resolved, would result in incoherence.<sup>18</sup>

---

17 Read this not as an analysis, but only a partial explication, of conceiving; the idea is to give some indication of what my *conceiving* it as possible that  $p$  adds to its merely *seeming to me as if* it was possible the  $p$  (as it might if I was reliably informed that  $p$  was possible). Among the many questions which I leave open are: what is the precise relation between conceiving (it as possible) that  $p$ , and believing that it is possible that  $p$ ? and, is conceiving to be understood as a non-modal attitude which (sometimes) takes *possibly*,  $p$  as its propositional content, or an intrinsically modal attitude which takes  $p$  as propositional content? Without prejudice to this latter question, we use 'conceive that  $p$ ' and 'conceive it as possible that  $p$ ' as synonyms; both indicate an act that is veridical if, and only if, it is possible that  $p$  (analogously, we can agree that the denial that  $p$  is correct iff it is not the case that  $p$ , without settling whether denying that  $p$  is believing that it is not the case that  $p$ ). In this respect, our usage may differ from that of Descartes, who seems willing, at times, to distinguish between conceiving that  $p$ , and conceiving that  $p$  is possible (CSM I, 299; AT VIII B, 351-2).

18 Two remarks. First, on this reading, there is little real prospect of an *absolutely* clear and distinct conception of the possibility that  $p$ , but only of a conception appropriately and sufficiently clear and distinct to allay anxieties about incoherence (notice that Descartes regularly treats clarity and distinctness as mat-

Start with distinctness. Nowadays we are familiar with a range of arguments purporting to show that there *is* a latent and unobvious incoherence in the idea of myself existing with my thought properties alone. Arguments like this are associated with Kant and Wittgenstein, and more recently with Ryle, Strawson, behaviorism, and externalist theories of mental content. Those unaware of, or unconvinced by, the considerations offered may claim to find it conceivable that they should exist with only their mental properties; but if those considerations are finally cogent, then they expose all contrary conceptions as incoherent. Obviously Descartes gave little thought to (e.g.) the Private Language Argument; but the general problem of unobvious entailments and the attendant risk of latent incoherence is one to which he was very much alive. As he observes in several places, '...there are many instances of things which are necessarily conjoined, even though most people count them as contingent, failing to notice the relation between them' (CSM I, 46; AT X, 422). Nevertheless, Descartes is convinced that his conception of himself with only his thought properties *is* relevantly distinct, and so deeply coherent if superficially so.<sup>19</sup> Speaking of his idea of himself as a thinking substance, he claims that he can

conceive it alone, and deny of it everything else of which I have an idea...[I have] no other idea which is prior to it and joined to it in such a way that I cannot think of the two together while denying the one of the other; for if there was any such within me, I must necessarily know it. (K, 124)

---

ters of degree, e.g., at CSM II, 22, 24; AT VII, 33, 35). Second, Descartes' view that there can be clarity without distinctness, but not conversely (CSM I, 208; AT VIIIA, 22), fits naturally with our account; an unclear conception, because it is silent about how certain matters stand, must be indistinct, since it would be incoherent to suppose that they stood in *no* way. Nevertheless, it is convenient to follow Descartes in treating clarity and distinctness as separate requirements.

<sup>19</sup> Admittedly there is a question, already alluded to, how Descartes hopes to reconcile this conviction with his argument for an external world; despite their enormous differences, Descartes, no less than Kant, thinks he sees an unobvious entailment from his subjective condition to external material objects (see note 14).

Of course, this is the very claim that Kant, Wittgenstein, and the others would want to question (could he deny external objects, if he understood their role in internal time-consciousness, or public language, if he appreciated its connection to the normativity of thought?). Since Descartes understands the distinctness claim as central to his argument, the issues they raise are exactly those on which he would, or should, have thought the matter rested. Unless we want to speculate on Descartes' response to the Refutation of Idealism, Private Language Argument, etc., the question cannot be pursued much further here. Suffice it to say that there *is* a question, and that anyone who champions Descartes' reasoning has got to assume that it will ultimately be answered in the negative.<sup>20</sup>

To clearly conceive of a situation in which I enjoy purely mental existence is to have a full, explicit, and determinate conception of what that situation would be like, in particular a conception free of saving unspecificities which however resolved would result in incoherence. At one time, I suppose I found it conceivable that there should be a town whose resident barber shaved all and only the town's non-self-shavers.<sup>21</sup> But this conception escaped inconsistency only by remaining unclear; once the barber's shaving habits were specified, the contradiction became obvious. Is my conception of myself as a purely mental being likewise saved from incoherence only by its inexplicitness?

---

20 Keep in mind that, unless Descartes can be faulted for not anticipating revolutionary developments to come, it was not *unreasonable* for him to claim distinctness for his self-conception as purely thinking thing; and also that there is no consensus, even among contemporary philosophers aware of those developments, that Cartesian solipsism is latently incoherent. In any case, the usual charge against Descartes' argument is *not* that he was wrong, or irresponsible, to claim consistency for his self-conception as purely thinking thing, but that he was wrong to think that such a claim could bear in any convincing way on his aptitude for purely mental existence.

21 Sometimes 'conceivable' is used 'factively,' so that from *p*'s conceivability as possible, its possibility follows. On this usage, I did not conceive it as possible that the town's barber should shave all and only non-self-shavers; I only seemed to do so. As I use 'conceivable,' that *p* is conceivable amounts roughly to its *seeming* to be conceivable in the first sense.

Usually when we are asked to conceive a situation contrary to the actual, we are working to highly partial specifications. Sometimes this leads to trouble, as in the barber case above; but trouble is the exception rather than the rule (which is why nobody complains if my conception of a situation in which Humphrey is President is silent on questions with no apparent bearing on Humphrey's office, e.g., the outcome of the Indian Mutiny). Thus it is all the more striking that when I am asked to conceive myself with exactly my thought properties, this comes very near to providing me with a complete specification of the situation intended; namely, one in which I possess all of the properties which I am in the actual situation directly aware of myself as possessing, and *no more*. Since the properties with which I credit myself in this conception are fixed by my actual state of consciousness, it is not easy to imagine where the problematic indeterminacy could be thought to reside. (Perhaps it goes too far to claim that my conception is fully explicit on every point; certainly, though, it compares extremely well with the competition.)

Tentatively, then, I conclude that I am, qua possessor of my thought properties, a complete thing, and specifically that I can clearly and distinctly conceive myself in a purely mental condition. Postpone for a moment the question whether this is enough to justify me in believing that I could exist in that condition; and ask instead, does it show, at least, that there can be no justification for *doubting* that I could? That depends on a subtle issue of modal epistemology. Descartes thinks that modal opinions are generated by reason; and this faculty he credits with a certain sort of priority relative to the other faculties: its deliverances are correctable only through the further exercise of reason, never by imagination or sense.<sup>22</sup> Thus correction is impossible if the grounds of our opinion are free of 'internal' difficulties, i.e., difficulties in principle disclosable through the exercise of reason. Insofar as clarity and distinctness are the ultimate 'internal' virtues, that my self-conception as purely thinking thing possesses these virtues would seem to show that nothing could justify me in *doubting* that purely mental existence is possible for me.

---

22 See Loeb, 'The Priority of Reason in Descartes.'

For Descartes, 'internal' deficiencies provide the *only* basis on which a modal opinion can be criticized as inaccurate. In recent years, through the work of Saul Kripke, an entirely different basis for criticism has come to light. What Kripke saw, and established beyond reasonable doubt, is that modal opinions *are*, Descartes notwithstanding, correctable through the exercise of sense (e.g., unaided reason finds no difficulty in the conception of a situation involving heat but not motion, but empirical research has turned up facts given which this is seen to be impossible). As a result, purely 'internal' virtues like clarity and distinctness are no longer enough to secure modal intuitions against attack; the most conscientious and clear-headed conceiver can be refuted in a moment by the dullest observer of the passing scene. Obviously this raises new problems, unimagined by Descartes, for the inference from conceivability from possibility, and indeed transforms the issues on which that inference depends in the profoundest way. Nevertheless, the essential lines of his thinking continue to hold up, or so I shall maintain.

## **VI Conceivability and Possibility**

Whether *p*'s lucid conceivability makes it irrational to *doubt* that *p* is possible is one question; whether it rationalizes the *belief* that *p* is possible is another. Why should what I am able to conceive of as occurring be any sort of guide to what can actually occur? Specifically, why should the possibility of my existing in a purely mental condition be thought to follow from my conceivability as existing in that condition? Here there are really two questions, one about why Descartes thought it followed, the other about why we should think so. On the first, I have nothing much to add to what Descartes says himself. In Descartes' view, all of his faculties are the handiwork of an all-powerful, all-knowing, and undeceiving God; and such a God 'surely did not give me the kind of faculty which would ever enable me to go wrong while using it correctly' (CSM II, 37-8, AT VII, 54). Not only his general faculty of judgment, but also its specific application to matters of mind and body, is said to be authorized by the veracity of God. To Gibieuf he writes that

...I do not deny that there can be in the soul or the body many properties of which I have no ideas; I only deny that there are any which are inconsistent with the ideas that I do have... for otherwise God would be a deceiver....(K, 125)

Not that Descartes supposes that divine veracity entirely precludes erroneous judgments about these topics. Through carelessness, inattention, or failure of imagination, unobvious consequences of my self-conception may escape my notice, with the result that I credit as possible a state of affairs which could never arise. But what apparently cannot happen, compatibly with God's veracity, is that the impossibility of this state of affairs should be forever undetectable, i.e., that what I conceive as possible is not possible, though there is no appreciable defect or difficulty in the conception. Mistakes can indeed arise, but when he reflects carefully on the fact 'that God is not a deceiver, and the consequent impossibility of there being any falsity in my opinions which cannot be corrected by some other faculty supplied by God,' Descartes sees that none of these are mistakes which he lacks the means to put right (CSM II, 55-6; AT VII, 80; see also K, 124).

So much for Descartes; why should *we* accept the inference to my possibly existing in a purely mental condition from its conceivability as possible? Strange as it may seem in view of his appeal to God's veracity, Descartes' account contains the seeds of a solution that may find favor even today. Two Cartesian ideas will be important. First, Descartes believes that it is *only* by way of our ideas that we can attain knowledge of what it is possible; so that if these ideas are unreliable, then modal knowledge must remain out of reach. Insofar, then, as we credit ourselves with modal knowledge, there is no alternative but to take our ideas as a guide to the modal facts. Already this is hinted at by the continuation of his remark to Gibieuf, quoted above; he says that soul and body cannot have properties inconsistent with his ideas, or else 'God would be a deceiver, *and we would have no rule to make us certain of the truth*' (K, 125, my emphasis). But the point recurs throughout the letter to Gibieuf, intricately interwoven with the appeal to divine veracity that was featured above:

## *The Real Distinction Between Mind and Body*

[you may object that] ... although I conceive the soul and body as two substances which I can conceive separately, and which I can even deny of each other, I am not certain that they are in reality such as I conceive them to be. Here we have to recall the principle already stated, that we cannot have any knowledge of things except by the ideas we conceive of them; and consequently, that we must not judge of them except in accord with these ideas, and we must even think that whatever conflicts with these ideas is absolutely impossible and involves a contradiction. (K, 124)

In effect, Descartes is saying that we have no other option than to rely on what we find conceivable in drawing conclusions about what can, and what cannot, happen. To be sure, God sees to it that this procedure will not lead us too far wrong. But it is a completely separate point that the vehicle of modal knowledge, if that knowledge can be obtained at all, must be our ideas.<sup>23</sup>

That modal intuition *must* be accounted reliable if we are to credit ourselves with modal knowledge, is a point that retains its plausibility even for those who disagree with Descartes about how that reliability should be accounted for. Unless we are willing to give up our claims to knowledge about something about what could have happened, though it did not, it seems unavoidable that we treat conceivability as a respectable, if not an infallible, guide to possibility. No doubt we are unhappy with Descartes' attempt at a justification for this policy, and hope to find another, but that is a separate ques-

---

23 Cf. K, 123: 'I am certain that I can have no knowledge of what is outside me except by means of the ideas I have within me....But I think *also* that whatever is to be found in these ideas is necessarily also in the things themselves' (emphasis added). Notice too that Descartes considers the argument from his ideas of self and body to be acceptable by *ordinary* standards even *without* the invocation of God's veracity: '...had I not been looking for greater than ordinary certainty, I should have been content to show in the Second Meditation that the mind can be understood as a subsisting thing despite the fact that nothing belonging to the body is attributed to it, and that, conversely, the body can be understood as a subsisting thing despite the fact that nothing belonging to the mind is attributed to it. I should have added nothing more in order to demonstrate that there is a real distinction between the mind and the body, since we commonly judge that the order in which things are mutually related in our perception of them corresponds to the order in which they are related in actual reality' (CSM II, 159; AT VII, 226). (See also CSM II, 272ff.; AT IXA, 207ff.)

tion.<sup>24</sup> The point for now is simply that this *is* our policy; *within limits*, what we are able to conceive as possible, it is our practice to admit as possible. Simple consistency obliges us to consider whether my conception of myself existing with my thought properties alone falls within these limits.

At this point another Cartesian idea becomes important, that we can never reach false conclusions, about modal matters or matters of any other kind, except through the misuse of our faculties. According to the usual story, Descartes claims certain knowledge of this principle on the basis of his certain knowledge of God's veracity. Lacking that recourse, I can't pretend to the same knowledge. Nor do I even believe the principle as stated. What I *do* think is that something *like* a 'no gratuitous error' claim is implicit in our daily practice, in the form of a ban on gratuitous *attributions* of error. Not that doubts must always be backed up by a story about how the thinker has misused her faculties; obviously it is possible to reach a false conclusion through no fault of one's own. But the suspicion that a judgment, modal or otherwise, is erroneous *does* ordinarily need to be grounded in a reason to think that error in *this* case was significantly likely.

Such a claim is of course commonplace as regards perception (the analogy with perception is meant to be suggestive, not probative).

---

24 Traditional conceptualism about modal truth might provide one such justification, but other forms of anti-realism could also serve. Neither is anti-realism forced on us; there are options in the theory of knowledge as well. The account in the text is meant to be neutral between these various possibilities, and indeed to allow that none of them is finally convincing. The problem of justifying reliance on our faculties is quite general, and the potential solutions similar, and similarly unsatisfying, across faculties (e.g., perception, memory, logical and mathematical intuition). Obviously it is not, and could not be, our policy to postpone assent to a faculty's deliverances until its reliability is philosophically assured. (In any case, the complaint against Descartes has always been that his appeal to conceivability involves certain *specific* errors, in light of which the proposed conclusion cannot be drawn *in this case*; it should not be allowed to degenerate into a general modal skepticism according to which we are never justified in relying on conceivability considerations, and so never justified in regarding the non-actual as possible.)



Absent specific and overriding grounds for doubt, perception affords a (defeasible, but that goes without saying) basis for belief. Doubts are of course legitimate if we have independent reason to think that the facts are not as reported, or not of the right kind to be perceived; or that the observer is reckless, or incompetent; or that even competent observers are, on this occasion, liable to go astray. Quite often we can cite some prior error or oversight, which explains the appearances even better than the hypothesis that the facts are as maintained. But what plainly *cannot* be used to justify incredulity is the abstract possibility of error. Obviously this is not meant to constitute any sort of answer to skepticism. The point is only that doubts not backed up in these ordinary ways *are* skeptical doubts; and where skepticism is not at issue, perceptual reports not subject to any but skeptical doubts, are accepted, and I will suppose acceptable, as *prima facie* accurate.

Not to minimize their differences, conception seems analogous to perception in this respect: absent specific grounds for doubt, *p*'s conceivability as possible *prima facie* justifies me in the belief that *p* is possible. Outside of philosophy, this would hardly require argument. Imagine that you claim to be able to conceive of a situation in which you exist, but the Washington Monument does not. Assuming that we ourselves find no difficulty in the conception, are we still in a position seriously to question the possibility of yourself without the Monument? Only, it seems obvious, if we can point to some complicating factor of a kind not yet envisaged (imagine your reaction if we said, 'nevertheless, we wonder whether it is really possible,' though no further complication suggested itself!). Unless we have it in mind to play the skeptic, and dissent from received standards of evidence, to resist now, without grounds for doubt or the prospect of them, would simply be to reveal ourselves as ignorant of what counts as sufficient reason for belief in cases like this.

With these lessons in mind, return to my conception of myself in a purely mental condition. Naturally I wonder whether this conception is veridical, *ie.*, whether it is the conception of a real possibility. Presumably this is because I have heard of cases of falsidical conception, cases where people conceived something as possible which was not in fact possible; and I wonder whether my own case might not be like that. For example, I suppose that the ancients had

no difficulty in conceiving it as possible that Hesperus should have existed without Phosphorus. From this they might erroneously have concluded that the contemplated situation could have obtained (erroneously, because Venus cannot exist without Venus). Maybe I am making an analogous mistake when I conceive myself as a purely thinking thing, and conclude that this is truly possible for me.

But is the analogy a good one? Remember that the ancients found it conceivable that Hesperus should have existed without Phosphorus, only because they falsely believed that Hesperus and Phosphorus were distinct. What is the mistaken belief which accounts for my erroneous intuition, as the ancients' misjudgment belief that Hesperus was not Phosphorus accounts for theirs?

Reflection on the ancients' mistake points toward the following model of modal error. First, I conceive it as possible that  $p$ , although  $p$  is necessarily false. Second, that  $p$  is necessarily false emerges from the truth of some proposition  $q$ . Third, I do not realize this, believing instead either that  $q$  is false, or that it is false that if  $q$ , then  $p$  is impossible; and that is how I am able to conceive, erroneously, of a situation in which  $p$ . Thus:

- (a)  $q$ ;
- (b) if  $q$ , then  $\Box \sim p$ ; and
- (c) my ability to conceive it as possible that  $p$  is explained by my denial of (a), or else by my denial of (b).

( $\Box p$  means: necessarily,  $p$ .) Subject to a qualification to be mentioned presently, every instance of erroneous conception that I am aware of fits this pattern.<sup>25</sup> For example, the ancients could conceive it as possible that Hesperus should exist without Phosphorus (that  $p$ ) only because they denied the truth ( $q$ ) that they were identical; if some contemporary philosophers, aware of this identity, find themselves capable of the same conception, that must be because

---

25 Although such a claim might well be correct, I do *not* claim that all modal error whatsoever fits the model (indeed, I leave it open that there might, in principle, be absolutely undetectable modal errors, to which, a fortiori, the model would not apply); my concern is more with the assertability, than the truth, of 'x is mistaken in conceiving it as possible that  $p$ .'

they deny the conditional truth that if the identity holds, then Hesperus is impossible without Phosphorus (that  $q$  only if  $\Box \sim p$ ). Similarly, Oedipus may suppose that he could have been King even if Jocasta had never lived (that  $p$ ). But that is because he believes that he is not her son (that  $q$ ); and if he persists in his error, that is because he denies, what for argument's sake we assume to be true, that if she is his mother, then he could not have existed unless she had (that  $q$  only if  $\Box \sim p$ ). Examples are easily multiplied, but let us return to the case of interest.

Conceivings are *prima facie* veridical; so I am *prima facie* entitled to think that I am capable of purely mental existence. The question is whether this *prima facie* entitlement can be defeated along the lines just indicated. For my modal intuition is erroneous, if there is a proposition  $q$  such that

- (a)  $q$ ;
- (b) if  $q$  then  $\Box$  (I possess more than my thought properties); and
- (c) my ability to conceive it as possible that I should possess no more than my thought properties is explained by my denial of (a), or of (b).

Certainly it would establish that my modal intuition was erroneous if someone was able to *prove* that it could be explained away in the manner indicated. But so much is not required. To raise legitimate *doubts* about the intuition, it ought to be enough to find a proposition  $q$  for which there is good reason to think that the model *may* apply (for in that case, the intuition is potentially explicable on some other basis than that it is true). Call  $q$  a *defeater* if there is, plausibly, a significant chance that (a), (b), and (c). Then the objector's challenge is to find a proposition which defeats my intuition of the possibility of purely mental existence.

Admittedly, it may be difficult for the objector to present me with a subjectively convincing example of a defeater. For no proposition  $q$  will strike *me* as a defeater unless I can be brought to recognize that I deny something (that  $q$ , or that  $q$  only if  $\Box \sim p$ ) that is not improbably true. And this is not something I am likely to admit.<sup>26</sup>

---

<sup>26</sup> Perhaps this is why the gap between conceivability and possibility can seem so hard to appreciate from the first-person point of view. Intuitively, I can conceive

But this complication need not detain us for long. For I ought to be able to recognize a proposition  $q$ , if there is one, such that it is because I deny that  $q$ , or that if  $q$  then  $\Box \sim p$ , that I am able to conceive it as possible that  $p$ . Having done so, I must admit that if, contrary to what I suppose, it is true that  $q$ , and that  $q$  only if  $\Box \sim p$ , then what I find conceivable is not in fact possible. Whether the objection succeeds must now depend on whether the propositions that  $q$ , and that  $q$  only if  $\Box \sim p$ , possess credibility sufficient to overcome the presumptive reliability of modal intuition.

Certainly there are very many propositions  $q$  such that I deny that  $q$  is a truth which shows me to be incapable of purely mental existence; for example, I deny this of the proposition that I was born on the planet Neptune. Most such denials are irrelevant, since there is no significant chance that that they are in error. But when we turn to propositions  $q$  such that it is *not* wildly improbable that  $q$  is a truth given which purely mental existence is impossible for me, e.g., that I possess more than my thought properties, or that my mental life is grounded in my physical condition, or that I necessarily possess more than my thought properties, or that I am identical to my body, we are met with a certain difficulty. Going into my thought experiment, *I do not deny* that these are truths which rule out the possibility of my purely mental existence; rather, I *come* to these denials as a *result* of the thought experiment. In some cases, the thought experiment leads me to deny  $q$ 's truth, in others its tendency is to show that I am incapable of purely mental existence. But in all cases, the conception precedes, and so cannot be explained by, the denial.

To illustrate, it cannot be said that I am able to conceive myself with my thought properties alone only because I initially deny that I possess physical properties, or that my mental life is grounded in my physical nature; or because I initially deny that if these things are true, then I am incapable of purely mental existence. When I attempt my conception, I *acknowledge* that I possess more than my thought properties, and acknowledge too that my mental life is

---

it, but it isn't really possible' has something in common with 'I believe it, but it isn't really true.' If the assertability of 'x can conceive that  $p$ , but it isn't possible that  $p$ ' is connected, as I am suggesting, with the assertability (for some  $r$ ) of 'x believes that not- $r$ , but  $r$ ,' then the reasons for the analogy become clearer.

grounded in my physical nature. And even if I do not acknowledge that these facts reveal me as essentially unfit for purely mental existence, neither do I deny it; indeed, I attempt the thought experiment in order to discover whether denying it would be unreasonable. Similarly I acknowledge that *if* I am identical to my body, then purely mental existence is impossible for me; and although I do not antecedently acknowledge, neither do I antecedently deny, that I am identical to my body. That is what the thought experiment tells me. So far, then, my conception is not in danger of being explained away.<sup>27</sup>

Someone might object as follows. To erroneously conceive it as possible that  $p$ , why should I have to go so far as to *deny* the proposition  $q$  given which  $p$  is impossible, or to *deny* the proposition that  $p$  is impossible if  $q$  is true? Isn't it enough if I am simply *ignorant*, that  $q$ , or *ignorant* that if  $q$  is true, then  $p$  is impossible? Thus consider a less demanding model of how erroneous conception can arise: there is a proposition  $q$  such that

- (a)  $q$ ;
- (b) if  $q$  then  $\Box \sim p$ ; and
- (c) that I can conceive it as possible that  $p$  is explained by my ignorance that (a), or else by my ignorance that (b).

Perhaps the 'ignorance' model does do a certain justice to cases which the 'denial' model leaves unaccounted for. Imagine, for example, that the medievals, rather than denying that whales were

---

27 Following Kripke, many philosophers believe that (K) for all  $z$ , if  $z$  is the zygote from which I actually derive, then I am necessarily derived from  $z$ . George Bealer observes that if (K) is independently credible, the proposition  $q$  that I derive from  $z$  (my actual zygote) looks like a defeater of my modal intuition; for  $q$  is independently credible, and given the independent credibility of (K), so, apparently, is the conditional proposition that if  $q$ , then  $\Box$  (I possess more than my thought properties). The problem is avoided if by 'I could have existed in a purely mental condition,' I mean only that I could have existed in that condition over some considerable part of my life. Admittedly this response is superficial, if, as may appear, I am now open to a second 'reduplication' argument of the sort typically offered for (K). But that argument, or so I claim, proves difficult to formulate.

mammals, simply had no opinion either way. Mightn't they still have conceived it as possible, erroneously of course, that they should have been something else (say, fish)? If so, then this gives an example of a falsidical conception whose explanation lies not in the fact that  $q$  is denied, but in the fact that it is not believed. Or take the stock example of the conceivability of Goldbach's conjecture, on the assumption that it is, unbeknownst to anyone, false; then it is not because I deny, but because I am ignorant, that some even number is not the sum of two primes, that I can conceive it as possible that the conjecture holds.<sup>28</sup>

Now we have a less demanding, and perhaps (see the last note) a more realistic, model of how modal intuition goes wrong. The objec-

---

28 Not everyone agrees that I can conceive it as possible that Goldbach's conjecture is false. Some will see me as confusing conceivability as metaphysically possible with some sort of epistemic possibility, e.g., it is not known, or not knowable a priori that, not- $p$ ; and others will claim to find a confusion between the conceivability of  $p$ , and its not being inconceivable (van Cleve, 'Conceivability and the Cartesian Argument for Dualism'). To the former, let me say that although 'conceivable' *can* be used to indicate epistemic possibility, what I mean by it is 'conceivable as metaphysically possible.' To the latter, my response is to question the existence of any sharp or principled distinction between its being conceivable, and its not being inconceivable, that  $p$ . Practically all conception is in some degree vulnerable to defeat; as the vulnerability increases, and our consciousness of it grows, we back off the 'conceivability' claim and incline more and more to the 'not inconceivable' formulation. But we do this in response to the gradual intensification of a concern that is never wholly absent, the concern that our intuition is liable to defeat by eventualities which we are not yet in a position to rule out. In the example given, this concern is deeply felt, and that accounts for our admitted hesitation in calling it conceivable that Goldbach's conjecture should be false. But I submit that I feel the same *sort* of hesitation, to a lesser degree, in claiming the conceivability of a situation in which I exist but my car does not (skeptics should consult their TV listings for reruns of the situation comedy 'My Mother the Car'). Having said that, I agree that in the Goldbach example we feel so *much* hesitation that the conceivability claim is at least tendentious. If anything, this strengthens my argument: the 'ignorance' model extends the 'denial' model only in cases where I simply cannot tell whether ( $q$  & (if  $q$  then  $\square \sim p$ )); but in those cases, I am presumably reluctant, anyway, to claim that  $p$  is conceivable. Thus it is mainly in connection with *uneasy* conceivability intuitions that the 'ignorance' model opens up new possibilities for criticism (this is a point I return to).

tor's challenge is to identify a proposition  $q$  for which there is a significant chance that the model applies. Now you may say that nothing could be easier. Let  $q$  be the proposition that I am incapable of purely mental existence; then as long as my intuition is still *sub judice*, there might seem to be a significant chance that (a)  $q$  is true, (b) if  $q$ , then I am incapable of purely mental existence (this is obvious), and (c) my ignorance of (a) explains my ability to conceive myself in a purely mental condition.

Nevertheless, I take it that it gives me *no* real reason not to trust my intuition that I am capable of purely mental existence, to be told that that intuition might be due, in part, to my ignorance of what might, for all I know, be the fact that I am *incapable* of purely mental existence. After all, it could equally be said that I am able to conceive it as possible that I should have had a different birthday, only because of ignorance about the necessity of my actual birthday. In either case, the most that can be claimed is that *if* the alleged defeater is true, and, e.g., it is necessary that I am born on September 30, then if I had not been ignorant of that fact, I would not have found any earlier birthday conceivable. And that is hardly an *objection*; no more than it is an objection to the veridicality of my perceptual impression that there are ducks present, that if I am wrong, and they are decoys, then my ignorance of that fact would figure in the explanation of how I was able to suppose that they were ducks.

Relating this intuitive response to the formal model takes some care; two points need to be distinguished. Even if we allow there is a significant chance that I am incapable of purely mental existence, there seems little chance that my ignorance of this fact could constitute the *explanation* of how I was capable of a contrary conception; the explanation must cite some *other* error or oversight to which my mistaken conception can then be attributed. But that is not the important point; for even if a more informative explanation is constructed, it carries little force if its plausibility depends on the *prior* concession that my conception is not improbably falsidical (this would be like explaining away my perception as of ducks by saying that they were not improbably decoy ducks, decoy ducks being the usual explanation of falsidical duck appearances). If there is any point to saying that the faculty of modal intuition is presumptively reliable, it is that one may not *assume* that a given intuition is untrustworthy, in

making the case that it should not be trusted. Only if there is some basis *independent of the issue under dispute* to suspect that my refusal of some relevant proposition *s* really does put me out of touch with the facts, does the allegation that *s* provide a reason for doubt.<sup>29</sup>

To summarize, the objector's challenge is to identify a proposition *q* for which there are *independent* grounds to suspect that my conceivability as a purely thinking thing is explained by my ignorance of the following fact: that *q* is a truth which shows that this is impossible for me. To see some of the difficulties involved, compare our imaginary medievals' intuition that whales could have been other than mammals, with my own intuition that I am capable of existing in a purely mental condition. Believing (as I will suppose) that whales might after all *turn out* to be mammals, and that if so they are mammals necessarily, these medievals should at least have felt some considerable *uneasiness* about their conception of whales as possibly not mammals. After all, they knew of a hypothesis *q*, amenable to straightforward empirical verification, whose truth would, by their own lights, reveal their conception as not veridical. However I know of no empirical hypothesis *q*, for which it is antecedently at all probable that if *q* is true, then I couldn't have existed

---

29 Someone might object that any consideration with the power to exhibit my unacceptance of the proposition that *q*, or that *q* only if *p* is impossible, as putting me out of touch with the facts, is, eo ipso, *not* 'independent of the issue under dispute' (since that issue is whether or not *p* is possible). But for *s* to be credible independently of the issue whether it is possible that *p* does *not* mean that *s*, if credible, cannot confer credibility on the thought that *p* is impossible; it means that *s*'s credibility is not owing to the *prior* credibility of that thought. Undoubtedly the distinction here alluded to raises fascinating and difficult problems, but its reality seems unmistakable. For example, observation gives me evidence that this swan is black, and this then confers credibility on the thought that not all swans are white. But the fact that 'this swan is black' would not be credible, if 'not all swans are white' were not also credible, has no tendency whatever to show that the former owes its credibility to the latter; and it would be absurd to complain, on the ground that my observation is misleading if all swans are in fact white, that I have failed to supply a reason 'independent of the issue' whether all swans are white, to think that this swan I am now looking at is black. So I see no in principle difficulty about finding reasons independent of the issue whether *p* is possible, for propositions which, if credible, would call *p*'s possibility into question.



as a purely mental being (which is why I do not feel the same sort of mistrust of my modal intuition as I am supposing that the medievals must have felt of theirs). Insofar, indeed, as  $q$  is an empirical hypothesis with some reasonable chance of coming out true, it is antecedently highly unlikely that if  $q$ , then I couldn't have existed as a purely mental being. And something like this holds more generally, I claim, of proposed defeaters  $q$  of my modal intuition: the better the chances are that  $q$  is true, the worse the chances for truth of the conditional proposition that if  $q$ , then purely mental existence is impossible for me.

Let us consider cases. Maybe  $q$  is the proposition that I have physical properties, where these may be either intrinsic or extrinsic. Since there is independent reason to think that I possess at least extrinsic physical properties,  $q$  is independently probable. But I am not aware of *any* independent reason to think that if I possess physical properties, *even if only extrinsic ones*, then I am incapable of purely mental existence. Someone might claim that there is independent reason to suspect that I have *intrinsic* physical properties, specifically, extension; and that there is independent reason to think that if so, then I am extended necessarily, and therefore cannot exist in a purely mental condition. About the second half of this, I am extremely doubtful. Like most people, I regard it as significantly likely that I *am* extended; somehow, though, this does not seem to inhibit me in conceiving myself as a purely thinking thing. But then I need positive argument that this intuition of being possibly-but-not-actually unextended is accountable to some prior error, before I can accept that any independent credibility attaches to the conditional hypothesis stated; otherwise, the objection comes to nothing more than the unsubstantiated allegation that my intuition may be wrong.<sup>30</sup> Of

---

30 Some philosophers may find it tempting to argue as follows: whatever is extended is some sort of body; and whatever is a body is necessarily so, and so necessarily extended. But this reasoning is vitiated by an ambiguity in 'is a body.' If it means 'is of the metaphysical kind <<body>>,' then it is not antecedently plausible that whatever is extended is a body; if it means, 'has the categorical properties of something of that kind, e.g., extension, mass, solidity ...' then it is not antecedently plausible that bodies are necessarily bodies (see section VII).

course, the conditional hypothesis becomes virtually certain if we let  $q$  be the proposition that I am *necessarily* extended. But now it is  $q$  itself which wants independent evidence.

Better, then, to look for a proposition  $q$  which, though not itself modal in character, has modal consequences (specifically, that I am incapable of purely mental existence). Perhaps there are independent grounds to suspect that I am *the same thing* as my body; and that if so, I am incapable of existing with my thought properties alone. (Certainly we seem to have an awful lot in common: shape, size, mass, and so on.) But what is meant by 'same thing'? If it means 'identical,' then the first conjunct needs some reason to believe it. However *categorically* similar my body and I may be, this gives grounds to suspect only that we are *coincident*, not that we are identical.<sup>31</sup> Evidence that we were moreover *identical* would presumably be evidence that my body and self agreed on a wide range of *non-categorical* or *hypothetical* properties, specifically on those for which the agreement is not readily accounted for in terms of our admitted categorical similarity. Counterfactual and dispositional properties are therefore of limited importance, and evidence of identity must to a large extent be evidence of *modal* similarity; yet this can only come from conceivability considerations, which seem in fact to argue the other way! If 'same thing' is understood so as to require sharing of categorical properties only, the problem is merely relocated. Now the apparent categorical similarity of my self and my body *does* give independent grounds for suspecting that we are the 'same thing'; only there is no longer any reason to think that our being so rules out the possibility of my purely mental existence.

So there seems to be at least this much difference between our imaginary medievals' intuition of the possibility that whales are not mammals, and my intuition of the possibility of purely mental existence: unlike the medievals, I am not aware of any independently credible hypothesis whose truth might be supposed, on independent grounds, to have the consequence that my intuition is incor-

---

31 See Yablo, 'Identity, Essence, and Indiscernibility.' Assume for the sake of the objection that there are no temporal differences between my self and my body, e.g., my body doesn't antedate me, nor is it going to outlast me.

rect. Surely it would be absurd and irrational for me to defer, in these conditions, to the abstract possibility that I am in error?

Maybe not. To this point, I have been pretending that the medievals were aware of certain *specific* issues (e.g., are whales mammals?), amenable to independent investigation, whose unfortunate resolution would, by their own lights, have exposed their modal intuition as incorrect. But it may be truer to the normal progress of our dialectic that the conceiver is *not* specifically aware of her conception's vulnerability to its eventual defeater, until the defeater comes along and does its work. Before the discovery of genes, for example, the thought may not have been readily available that scenarios in which animal life was organized along some non-genetic basis risked exposure as not only false, but impossible, by the progress of science. None of this is to deny that the concept of an animal must *somehow* 'prepare the ground' for the eventual recognition that (e.g.) animals necessarily propagate their kind by way of genes. But it is striking how unaware it is nevertheless possible to be of the vulnerability of one's modal intuition to what emerges, in the end, as its defeater. And now the objection comes, can't that be how it is with my intuition of the possibility of purely mental existence?

Ideally lucid conception, were it obtainable, would anticipate, I suppose, every possible scenario for defeat (even before Mendel, ideally lucid conceivers would have realized that such-and-such discoveries would rule it out that animal life could be organized on a non-genetic basis). So understood, ideally lucid conception is not within our powers; but what we are being asked to consider is how very far short of ideal lucidity our conceivings can fall, and how risky it therefore becomes to assume that no defeater would come into view, if it were somehow obtained. Of course this risk cannot be generally prohibitive, or *no* modal intuition would be trustworthy; so the idea must be that there is something in the nature of the thought that I exist in a purely mental condition to encourage the suspicion that in *this* case, if ideal lucidity were achieved, defeat would follow.

What might that something be? Recent work in the theory of content has turned up a variety of cases in which there is a significant gap between grasping a thought content, and appreciating the truth-

conditions it induces.<sup>32</sup> Misidentifying a cunningly groomed shrub as Brendan Sullivan, I entertain the content that *that* individual is not a potted plant; although I consider that I have thought something true in just those worlds in which Sullivan is not a potted plant, I am mistaken: it is the *shrub's* (actual and counterfactual) condition that matters. In this example, indexicality appears to be the culprit. But the same phenomenon can arise with contents that are not on their face indexical: for example, contents involving natural kind concepts; or concepts sensitive to community consensus regarding the use of their standard linguistic expression; or, more generally, concepts whose contribution to truth-conditions is affected by factors potentially unavailable to the thinker.

Since conceivability is a matter of thought content and possibility a matter of truth-conditions, contents for which this gap is especially large (call them 'schematic' contents) seem peculiarly apt to figure in delusive conception. Continuing the example above, I experience no difficulty in conceiving as possible a situation in which *that* individual employs dishonest methods, because I fail to see that that must be a situation in which the *shrub* does this. And now the objector argues that if my conception of myself in a purely mental condition is similarly schematic, then that should provoke concern about its accuracy. For as content grows more schematic, it constrains truth-conditions less and less; and the risk accordingly grows that the truth-conditions present difficulties to which the content offers no clue. Defeat is therefore to be expected, in the form of a proposition *q* spelling out the worldly facts which guide the transition from (benign) content to (malignant) truth-conditions.

However, I will need an argument before I concede that my I-thoughts are dangerously schematic. Remember that conceivability intuitions vary in subjective insecurity, according to how seriously one regards the threat of defeat. Ideally the potential defeaters

---

32 By 'thought content' I mean something peculiarly suited to the classification of the thinker's subjective condition or internal point of view; and by 'truth-conditions' I mean something which determines truth-values over all possible worlds. Depending on context, a given content can induce a variety of truth-conditions; the larger this variety, the larger the gap referred to in the text.

have been identified, and then our confidence depends on the probability we attach to their being truths incompatible with the intuited possibility. But even when the threat is open-ended, subjective insecurity continues to track expectation of defeat, via our sensitivity to schematic elements in the content entertained.<sup>33</sup> Other things being equal, one would expect the perilously schematic character of my I-thoughts to express itself in a pronounced insecurity about my intuition of disembodied existence. Then why do I not feel this insecurity? Various explanations may be possible; but the natural explanation is that my I-thought is not perilously schematic after all.

Even if true, the allegation that my I-thoughts are schematic could hardly be decisive. After all, an I-thought serves as the content of my intuition that I could have existed without Margaret Truman; yet this does not suffice to call the intuition into serious doubt. Thus the abstract possibility of trouble en route from content to truth-conditions, unsupplemented by a plausible scenario about how that possibility might be finding expression in the actual case, seems not to be enough. But then we will need a proposition spelling out how the envisaged complications are supposed to arise (maybe I *am* Margaret Truman); and this proposition would seem to be none other than a defeater, whence a defeater is required in any case. To consider the obvious example, someone might believe that my I-concept picked out the entity, whatever it was, activities in which constituted the ultimate basis of *these thoughts*; and she might attempt to explain my modal intuition away by citing my failure to bear in mind that: I am the entity so described, and the entity so described is my body (but for this, I would see that the possibility of my enjoying purely mental existence is ruled out by my body's inability to do the same). But this is just to offer as defeater the proposition *q* just formulated, which must then be sub-

---

33 Thus I feel far more confident of my intuition that there could have been a planet without mountains, than of my intuition that there could have been (e.g.) a force proportional to the mass of the object it acted on (as gravity was supposed to be).

jected to the same scrutiny as any other proposed defeater. Like them, it is found wanting.<sup>34</sup>

Vague and circumstantial worries about its potential for defeat cannot overcome the *prima facie* credibility of the Cartesian intuition. Pending the discovery of a specific defeater, I propose to acquiesce in the intuition, and to conclude that purely mental existence is possible for me.

## VII Categorical Dualism

Maybe you think that this conclusion is in order, or maybe you think it goes too far; in either case, it is important to remember that the full-blooded Cartesian dualist maintains something even stronger. At the outset I distinguished between the *hypothetical* dualism which asserts the *separability* of selves from bodies, and the *categorical* dualism which claims to find fundamental categorical differences between self and body, such as would imply their *separation in fact* (as statue and clay are *not* separate in fact). In the *Meditations*, at least, Descartes betrays little appreciation of this crucial distinction:

[because] 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 ...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. (CSM II, 54; AT VII, 78)

If by 'distinct,' Descartes means non-identical, then from the premise that *x* and *y* 'can be made to exist in separation,' it does indeed

---

34 As I see it, no independent credibility attaches to *q*'s first conjunct *r*: I am the entity activities in which constitute the ultimate basis for these thoughts. Like many people, I acknowledge that my thoughts are owing to occurrences in my body; yet this does not inhibit me in conceiving myself in a disembodied state. Absent positive argument for *r*, to offer it as independently credible is simply to forget the presumptive reliability of this modal intuition. Perhaps the needed credibility is thought to flow from an (acknowledged?) a priori equivalence between my I-concept and the descriptive condition given. But if this a priori equivalence obtained, then presumably I ought to know it; and not-*r* ought accordingly to strike me as a priori false (which I submit it does not).

follow that they are distinct. But if by 'distinct' he means *categorically* unlike, then he simply does not explain how this is supposed to follow from mere separability. Thus Descartes' argument for hypothetical dualism, even if accepted, is far from establishing the *categorical* dualism which asserts actual separation on the basis of fundamental categorical dissimilarity.

Now Descartes does of course believe that there are important categorical differences between mind and body, in particular that *minds are not extended*, and that *bodies do not think*.<sup>35</sup> To be sure, the situation is somewhat complicated by his contention that there is something – the mind/body union, sometimes called the 'human being' or the 'man' – which is both thinking and extended. But this latter doctrine should not distract us from Descartes' repeated assertions that the components of this union are in categorical respects utterly disparate; the body is extended and unthinking ('I have never seen or perceived that human bodies think; all I have seen is that there are human beings, who possess both thought and a body' [CSM II, 299; AT VII, 444]), and the mind is unextended and thinking ('I deny that true extension as commonly conceived is to be found in God or in angels or in our mind or in any substance which is not a body' [K, 239]). As for the man, he is thinking and extended *only* in the sense that he has disjoint parts of which one is an unextended thinker and the other unthinking and extended:

...the question is whether we perceive that a thinking thing and an extended thing are one and the same by a unity of nature. That is to say, do we find between thought and extension the same kind of affinity or connection that we find between shape and motion, or understanding and volition? Alternatively, when they are said to be "one and the same" is this not rather in respect of unity of composition, in so far as they are found in the same man, just as bones and flesh are found in the same animal? *The latter view is the one I maintain....*(CSM II, 286; AT VII, 424; emphasis added)

---

35 This is not to say that the 'real distinction,' as Descartes conceives it, expresses a categorical dualism; indeed in its canonical statements (e.g., CSM I, 213; At VIII, 28-9) it sounds decidedly hypothetical. What I do think is that, first, Descartes *was* a categorical dualist, second, he was *seriously* unclear about how far categorical dualism outreaches hypothetical, and, third, he had *some tendency* to read his arguments for the real distinction as having established categorical dualism *inter alia*.

Apparently, then, no single thing is both thinking and extended, in the way that triangularity and rectilinear motion *can* jointly inhere in a single thing. Adapting Descartes' terminology slightly, we can say that the mind thinks by *nature*, the man by *composition*, that is, by inheritance from a proper part which thinks by nature; similarly the body is extended by nature, the man by composition. Then Descartes' view is that *the thing which thinks by nature is not extended, and the thing which is by nature extended does not think*. Using 'I' for the thing which thinks by nature, and 'my body' for the thing which is by nature extended, Descartes maintains that I am not extended, nor does my body think.<sup>36</sup>

Given the centrality of these ideas in his thought, it is little short of astonishing that our problem here is not so much to evaluate his reasoning, as to discover what his reasoning could have been. In a work as late as the *Principles of Philosophy* (1644), Descartes still shows a tendency to slide over from separability into separateness:

...even if we suppose that God has joined some corporeal substance to ... a thinking substance so closely that they cannot be more closely conjoined, thus compounding them into a unity, they nonetheless remain really distinct. For no matter how closely God may have united them, the power which he previously had of separating them, or keeping one in being without the other, is something he could not lay aside....(CSM I, 213; AT VIII, 29)

---

36 Three remarks. First, someone might question whether Descartes would assent to 'I am unextended,' on the ground that 'I' refers not to the mind but to the man. Actually, Descartes' usage is unclear on this point, but even if I were the man, it would remain that I was categorically distinct from my body, for I think, and my body does not. In the text, we use 'I' for the thing which thinks by nature; on that usage, Descartes does of course think that he is unextended. Second, Descartes does sometimes allow that mind can be in a very weak sense 'extended,' simply by being in union with body (we might say that mind can be extended 'by union'); however, he makes it very clear that extension by union is not extension in any real or familiar sense (K, 119, 143). Third, when *x* is said to possess an attribute *P* 'by nature,' this does not mean that *P* is a *nature* of *x*, and in particular it does not mean that *P* is a property that *x* cannot exist without, or a basis for its other properties. (For example, it is by nature that the plank is warped, but being warped is not the plank's nature.)



From this one surmises that Descartes takes mind's separability from body to indicate that even in the actual circumstances, soul and body are at best 'closely conjoined.' If 'conjoined' is understood so as to permit overwhelming categorical similarity (i.e. if statue and clay are 'conjoined'), then the conclusion follows, but has no tendency to show that mind is actually unextended, or that body does not think. But if, as seems enormously likelier, 'conjoined' entities are categorically unlike, then it needs an argument to show that my separability from my body entails that we are, as matters stand, at best 'conjoined.'

Most reconstructions of Descartes' reasoning make appeal here to the premise that whatever is embodied is necessarily so.<sup>37</sup> If accidental embodiment is impossible, then from my possible disembodiment, my actual disembodiment evidently follows. Whether Descartes takes the impossibility of accidental embodiment as a premise or not, in the present context its plausibility owes entirely to a confusion between (a) being *a body*, in the sense of belonging to the kind <<body>>, and (b) being *embodied*, in the sense of being categorically (almost) indiscernible from something of that kind. Admittedly bodies are necessarily bodies (and so necessarily embodied); thus if embodiment implies being a body, nothing can be embodied without being necessarily so. But to assume that only bodies can be embodied is simply to beg the question against the categorical monist who alleges that what I am is not a body but an embodied *person*, whose categorical properties are (approximately) those of a certain thinking body, but with modal characteristics all its own.

Nothing remains for Descartes but a last-ditch appeal to the idea that thought *excludes* extension, i.e., that nothing can possess both 'by nature.'<sup>38</sup> Since I undoubtedly think, it would follow that I am

---

37 See van Cleve, 'Conceivability and the Cartesian Argument for Dualism,' Hooker, 'Descartes' Denial of Mind-Body Identity,' and Schiffer, 'Descartes on His Essence.'

38 Notice that this assumption, if Descartes were prepared to make it, would render his subtle conceivability argument entirely superfluous. But then why bother with the conceivability argument at all?

not extended. Some slight evidence that Descartes is attracted to this reasoning comes from his response to a 1647 pamphlet published by his former disciple Regius. Regius remarks that:

...if we are to follow some philosophers, who hold that extension and thought are attributes which are present in certain substances, as in subjects, then since these attributes are not opposites but merely different, there is no reason why the mind should not be a sort of attribute co-existing with extension in the same subject, though the one attribute is not included in the concept of the other....(CSM I, 294-5; AT VIII B, 342-3)

Descartes replies that if we are talking about

...attributes which constitute the natures of things, it cannot be said that those which are different, and such that the concept of the one is not contained in the concept of the other, are present together in one and the same subject; for that would be equivalent to saying that one and the same subject has two different natures – a statement that implies a contradiction, at least when it is a question of a simple subject (as in the present case) rather than a composite one. (CSM I, 298; AT VIII B, 350)

(Cf. also CSM II, 159; AT VII, 227.) Apparently, then, whatever is both thinking and extended *must* be composite:

A composite entity is one which is found to have two or more attributes, each one of which can be distinctly understood apart from the other. For, in virtue of the fact that one of these attributes can be distinctly understood apart from the other, we know that the one is not a mode of the other, but is a thing, or attribute of a thing, which can subsist without the other. A simple entity, on the other hand, is one in which no such attributes are to be found....[Hence] that which we regard as having at the same time both extension and thought is a composite entity, namely a man – an entity consisting of a soul and a body. (CSM I, 299; AT VIII B, 350-1)

Ignore as irrelevant the question why a composite of soul and body should be expected to inherit thought and extension, strictly understood, from its thinking and extended parts (or why, if it did, its unthinking and unextended parts should not equally confer on it thoughtlessness and unextension!). Our problem is much more basic. If by a 'composite' entity, Descartes means a subject of distinctly comprehensible attributes, then that reduces his complaint

against Regius, that whatever has distinctly comprehensible attributes is composite, to the triviality that whatever has distinctly comprehensible attributes, has them. To restore the complaint's substance, 'composite' needs to be returned to its original meaning, namely 'divisible into disjoint parts.' But now the same old worries recur. How does Descartes know that only what is divisible into disjoint parts can possess both thought and extension? What is the argument which rules it out that *some* things, for example, *people*, are thinking and extended *by nature*, that is to say, *otherwise* than by separate inheritance from categorically disparate components?<sup>39</sup>

Obviously it would be disappointing if Descartes had to resort here to a neo-scholastic prejudice according to which every undivided entity *must* be characterized by a single fundamental nature, of which all of its other (non-transcendental) properties are modes. For positive argument, he seems driven back on his apparent conviction that nothing is conceivable as thinking and extended, except by postulating a separation of that thing into purely extended and purely thinking parts; in a word, that nothing is conceivable as thinking and extended *by nature*.<sup>40</sup> Whatever the precise bearing may be

---

39 Admittedly, the clear and distinct comprehensibility of thought without extension does establish that the former is not, in Descartes' sense, a mode of the latter (for modes are not intelligible without their associated attributes [CSM I, 210-1; AT VIIIA, 25]). But on an intuitive level, that we can understand thought without extension shows at most that thought is *not necessarily* a way of being extended, not that it is *necessarily not* a way of being extended. Thus there is room, which the categorical monist may want to take up, for the view that it is in fact by being appropriately extended that one thinks, though thinking can in principle proceed on some other basis, or on no basis at all (thought remains an attribute, since it does not *presuppose* extension). On such a view, we do indeed possess distinctly comprehensible attributes by nature. But it is equally open to the categorical monist to say that we possess thought and extension both by nature, although thinking is *not* a way of being extended, nor conversely.

40 Perhaps Descartes' 'incompatibilist' rears in the 1647 *Notae*, and his 1648 statement to Burman that we possess clear conceptions of mind and body 'as two substances which not only do not entail one another but are actually *incompatible*' (CB [28], emphasis added), reflect a belated recognition of the gap between his premises and his conclusion.

of inconceivability on impossibility (this is something we have not discussed), the problem with this lies elsewhere: it is simply not obvious, if it ever was, that nothing is conceivable as by nature both thinking and extended. In the *Essay Concerning Human Understanding*, Bk. 4, Ch. 3, Part 6, John Locke proposes that it is

not much more remote from our comprehension to conceive that God can, if he pleases, superadd to matter a faculty of thinking, than that he should superadd to it another substance with a faculty of thinking.

From the ensuing controversy, it emerges that Locke was at any rate not simply *wrong* about this, even for his own time.<sup>41</sup> Even if subsequent discussion has done little to relieve the obscurity of bodily thought, it has tended to confirm Locke's judgment that the combination is not strictly inconceivable. But then I am still without a reason to believe that I am not extended, or that my body does not think.

### References and Abbreviations

- C. Adam and P. Tannery, ed., *Oeuvres de Descartes* (Paris: Vrin/C.N.R.S 1964-76) (=AT)
- J. van Cleve, 'Conceivability and the Cartesian Argument for Dualism,' *Pacific Philosophical Quarterly* 64 (1983)
- J. Cottingham, R. Stoothoff, & D. Murdoch, ed., *The Philosophical Writings of Descartes* (I, II) (Cambridge: Cambridge University Press 1985) (=CSM)
- J. Cottingham, ed., *Descartes' Conversation with Burman* (Oxford: Clarendon Press 1976) (=CB)

---

<sup>41</sup> See Yolton, *Thinking Matter*, for a detailed history of the debate Locke provoked by his innocent remark.

*The Real Distinction Between Mind and Body*

- M. Hooker, 'Descartes' Denial of Mind-Body Identity,' in Hooker, ed., *Descartes: Critical and Interpretive Essays* (Baltimore: Johns Hopkins 1978)
- A. Kenny, ed., *Descartes: Philosophical Letters* (Minneapolis: University of Minnesota Press 1981) (=K)
- S. Kripke, 'Identity and Necessity,' in Schwartz, ed., *Naming, Necessity, and Natural Kinds* (Ithaca: Cornell University Press 1977)
- L. Loeb, *From Descartes to Hume: Continental Metaphysics and the Development of Modern Philosophy* (Ithaca: Cornell University Press 1981)
- \_\_\_\_\_, 'The Priority of Reason in Descartes,' *Philosophical Review* **99** (1990)
- H.T. Mason, ed., *The Leibniz-Arnauld Correspondence* (Manchester: Manchester University Press 1967) (=LAC)
- S. Schiffer, 'Descartes On His Essence,' *Philosophical Review* **85** (1976)
- S. Shoemaker, 'Embodiment and Behaviour,' in Shoemaker, *Identity, Cause, and Mind* (New York: Cambridge University Press 1984)
- \_\_\_\_\_, 'Immortality and Dualism,' in Shoemaker, *Identity, Cause, and Mind*
- \_\_\_\_\_, 'On an Argument for Dualism,' in Shoemaker, *Identity, Cause, and Mind*
- B. Thomas, 'Abstraction and Complete Things' (unpublished manuscript University of Michigan)
- \_\_\_\_\_, 'Conceivability and the Real Distinction' (unpublished manuscript University of Michigan)
- B. Williams, *Descartes: The Project of Pure Enquiry* (New Jersey: Humanities Press 1978)
- S. Yablo, 'Identity, Essence, and Indiscernibility,' *Journal of Philosophy* **84** (1987)
- J. Yolton, *Thinking Matter* (Minneapolis: University of Minnesota Press 1983)