CHAPTER TEN

Individuation

In the previous chapter we previewed the topic of this current chapter: How are we to account for the possibility of there being numerically distinct things which are qualitatively identical? If two or more things are qualitatively identical, i.e. share their properties in common, what then accounts for their multiplicity? To be able to pursue such a question has required us to devote virtually all of chapter 9 to an examination of the concept of *sharing properties in common*.

10.1 Physical objects

One of the facts of our experience, perhaps the most familiar, is that the world contains an enormous number of physical objects. It may, as well, contain many other kinds of things, e.g. minds, souls, numbers, universals, and forces. But we will confine our attention in the earlier parts of this chapter principally to physical objects.

Ordinarily we do not much wonder about the commonplace. It is the unusual, the out-of-routine, the unfamiliar, which stands out and demands our attention. One may wonder how it is possible, if indeed it is, that extrasensory perception might occur, or that certain human beings can walk on burning coals, or that the influenza virus of 1918-19 which killed 20 million persons disappeared as abruptly as it had initially appeared. But few persons would be similarly tempted to wonder about something as familiar and commonplace as the existence of material objects. And yet, when one does think about material objects, the more one thinks about them, the more mysterious they seem.

First of all there is the most basic question of all: Why is there something, rather than nothing? Many persons are tempted to give a theological answer: God has created the world. But that answer only raises another: Why did God create the world? There is no satisfactory answer forthcoming from rational theology. Various religions may offer answers; but religious answers are, more often than not, glaringly deficient when judged on the grounds of rational theology and are usually terribly unconvincing to persons who do not share the same religious orientation.

Some modern thinkers, adopting what they have called "the anthropic principle", have turned the question and its answer round about, arguing that unless there were a world, complete with material objects, we could not even ask the question "Why does anything exist?" While the claim is undeniably true, many critics object that it does not provide a satisfactory answer to the question.[1] It is analogous to someone's answering the question "Why does Joan love classical music?" by replying that if Joan's mother had not become pregnant then Joan would not have been born and hence would not have loved classical music. Although Joan's mother becoming pregnant was a necessary condition for Joan's being born and hence a necessary condition for Joan's liking classical music, neither of these events (or states) – Joan's mother becoming pregnant and Joan's liking classical music – critics will insist, explains the other. Similarly, if the world did not exist, none of us could ask why it does; but neither of these events (states) – the world's existing and our asking why it does – critics will insist, explains the other.[2]

Although I have read what a great variety of writers have had to say on this subject, I have never found an answer which has had (for me) a

---

1. The anthropic principle has received considerable attention in the last two decades, both in academic circles (see, e.g., [22] for a sustained defense, and [62] for a criticism – the latter contains a good bibliography) and in the popular media (see, e.g., [74]). But it is not a new principle by any means. Its promoters advance it explicitly as a resurrecting of the teleology of ancient and medieval science (again see [22]). It is a specialized version of the very sort of principle which Bacon, as we have earlier seen (p. 58), was intent to dispel from physics nearly four hundred years ago. As I explained, the arguments for and against adopting such principles cannot be decided by experimental science. These are metaphysical principles which require us to examine the very core of our conception of what a satisfactory model of explanation may be. I, like Bacon, regard such principles, whether in their old-fashioned guise or in their latest raiment, as inappropriate within physics and cosmology. But, clearly, other writers do not share this view.

2. Even more recently (1984), Nicholas Rescher has taken a new departure and has pursued a boldly speculative proposal: that certain very basic physical laws (laws of Nature) – "protolaws" he calls them – necessitate the exist-

significant degree of plausibility. I have become steadily more convinced that the question is ill-conceived. It is as ill-conceived as, for example, one way of posing the question "Where is the universe?" One can, as I suggested earlier (footnote 18, p. 265), answer the latter question by saying (fairly uninformatively): "It is where any of its spatial parts are." But one cannot, of course, rationally demand an answer which specifies where the universe is in relation to something *outside* of the universe. Similarly, one can explain why some specific event occurs in the history of the universe by citing antecedent events (and – depending on one's theories of historical explanation – possibly historical laws, generalizations, or truisms as well). But one can-

---

ence of material objects; and that the truth of these protolaws, themselves, comes about through their maximizing certain 'cosmic values' ([170]). There are two forbidding hurdles lying in the path of this solution.

The first problem lies in the exceptional role assigned to the protolaws. In virtually all contemporary accounts, physical laws state the relationships between existent or possible entities; but they do not 'require' or 'necessitate' the existence of those entities. That is, physical laws are regarded as being logically conditional, not categorical, i.e. as *not* implying the existence of their subjects. For example, the law which states that an isolated planet and a star will revolve around a common focus in elliptical orbits is, in fact, never realized. There are no fully isolated planetary systems anywhere in the universe. Physical laws state (conditionally) "such and such would happen *if* …"; they do not state (categorically) "such and such actually exists and behaves thus …" So we see that Rescher's hypothesized protolaws would be considerably different from 'ordinary' physical laws. To explain the *existence* of material objects, they would have to be an ontologically different kind of thing than familiar physical laws. We would have, as it were, two 'tiers' of physical laws.

The second difficulty arises from Rescher's suggestion that the protolaws maximize certain 'cosmic values' ('cosmic' in that these values have nothing to do with human concerns, i.e. are not ethical, aesthetic, etc.). We are not, however, given an example of any of these cosmic values, and hence are in no position to judge whether they can explain protolaws. But we can anticipate certain severe problems. Insofar as these cosmic values are further characterized as being self-justificatory, it is highly improbable that they could serve as the explanatory basis of protolaws which include, Rescher suggests, quantum mechanics, general relativity, and the like. It is difficult, if not impossible, to imagine self-justifying cosmic values with the kind of specificity needed to explain why quantum mechanics rather than a logically possible, but incompatible, alternative should be true.

not extrapolate indefinitely and think that one can ask meaningfully why the universe as a whole exists. Just as there is nothing outside of the universe against which the spatial position of the entire universe may be gauged, there are no events outside of the universe whose occurrence can be rationally posited as explaining the existence of the universe.

This conclusion may seem overly pessimistic. Many of us hate to think that our understanding must be limited. But the world provides no guarantee of its being fully comprehensible. Given how we standardly explain events and states, it seems to me that we are positively precluded from being able to formulate an intelligible answer to the question "Why is there anything, rather than nothing?" The only possible reply must be: "We cannot answer such a question. It can have no answer."[3]

I take it as a datum that the universe exists and that it has material objects in it. To be sure, some philosophers – in the two thousand years of written philosophy – have challenged even this assumption. I have no particular desire here to recount their arguments. As I have said (p. 240), I do not think that there can be any conclusive refutation of their theories. But I do believe that the contrary posit, that there are material objects, is rationally supportable. In any event, I am sufficiently comfortable with it that I am perfectly happy to accept it as a working hypothesis if for no other reason than to get on to what I do want to examine. (If this latter claim strikes you as evidencing an excessive tentativeness, let me state explicitly that I firmly believe that there *are* material objects. It is just that I am not interested here in making a diversion into a discussion of what, for most persons, requires no justification whatsoever.)

There are material objects – this is the starting point for this chapter. I do not ask *why* there are material objects. But I do want to ask

---

3. Robert Nozick strongly rejects this pessimistic verdict. But he cautions that very odd theories will have to be generated. Characterizing his own theories (1981), he writes: "The question cuts so deep, however, that any approach that stands a chance of yielding an answer will look extremely weird. Someone who proposes a non-strange answer shows he didn't understand this question. Since the question is not to be rejected, though, we must be prepared to accept strangeness or apparent craziness in a theory which answers it" ([146], 116).

how it is *possible* that there are material objects. The two questions are very different.

Physical objects, material things, in spite of their familiarity turn out to be conceptually puzzling. What accounts for the numerical distinctness of material objects? We might discover, in due course, that this question, too, just like the ones I dismissed a moment ago, may be unanswerable; it may be ill-conceived in ways unapparent at the outset. But whether this is so remains to be seen. We must make an attempt to answer it. In that attempt we must be prepared for one possibility: discovering that our initial question was ill-conceived and the enterprise we have embarked upon must eventuate in frustration. But we cannot decide this in advance. 'Nothing ventured, nothing gained' is true; unfortunately, its positive counterpart 'Something ventured, something gained' may be false. Without making the attempt, we will gain nothing. However, in making the attempt, we are not assured of any success. We are not even entitled to the belief that success is possible, still less that it is probable.

10.2 Identity-at-a-time versus identity-through-time

It has become customary to break the question of the numerical identity of physical objects into two stages. We begin (in this chapter) with the question of the numerical identity of objects at some given moment of time. This might be called 'snapshot' identity. For example, we might ask, "Is the book I am now looking at the selfsame book you are now looking at?" This is 'momentary' or 'instantaneous' identity. That is, this is identity which disregards the fact, if it is a fact, that the 'things' being identified may also happen to endure through time. The instantaneous aspect of the identification is perhaps better expressed in a formal statement. The question we will be intent to pursue may be formulated this way:

> Under what conditions may $\boldsymbol{O}_1$-at-T_1 be regarded as being numerically identical with $\boldsymbol{O}_2$-at-T_1?

Identity-at-a-time is often spoken of as 'synchronic identity' and equally as 'individuation'. It is easy to see why the concept of individuation is invoked in this context: to specify the conditions under which $\boldsymbol{O}_1$ and $\boldsymbol{O}_2$ are numerically identical (or distinct) is just to specify the conditions under which $\boldsymbol{O}_1$ is (or is not) to be regarded as being the same *individual* as $\boldsymbol{O}_2$. To specify conditions of numerical identity just is to specify the conditions of individuation.

Having once addressed the question how numerical identity, or individuation, is to be accounted for at some specific moment of time, one can then go on to ask how identity-through-time might be explained. The latter question presupposes an answer to the former. Objects must be individuated – i.e. we must be able to pick out individual objects – before we can hope to be able to trace their evolution through time. Thus we will separate the two problems, tackling individuation (or synchronic identity) in this chapter and identity-through-time (or diachronic identity) in the next.

10.3 Positive and negative theories of individuation

Attempts at solving the problem of individuation divide into two distinct classes: those which argue that what individuates objects is their properties; and those which argue that properties alone can never individuate objects, and hence which find it necessary to posit something 'beyond' or 'behind' an object's properties, a special 'individuator', viz. its 'substance'. Theories which attempt to solve the individuation problem by invoking *nothing more* than a thing's properties are spoken of as 'negative' theories. Negative theories are also sometimes colorfully called "bundle theories" since they argue that there is nothing more to a thing than its being (metaphorically) a 'bundle' of properties. Theories which try to solve the individuation problem by positing that there is *more* (e.g. substance) to a thing than just its properties are said to be 'positive' theories.

Negative theorists will, like Leibniz, take as their point of departure the principle of the identity of indiscernibles (see p. 232 above), arguing that numerical identity can be accounted for strictly in terms of the properties which things instance. In the previous chapter we introduced symbols for numerical identity and for qualitative identity. The principle of the identity of indiscernibles may be stated in a specialized version for identity-at-a-time:

$$(\boldsymbol{O}_1\text{-at-}T_1 \mathrel{Q} \boldsymbol{O}_2\text{-at-}T_1) \rightarrow (\boldsymbol{O}_1\text{-at-}T_1 = \boldsymbol{O}_2\text{-at-}T_1)$$

This principle states that $\boldsymbol{O}_1$ is numerically identical to $\boldsymbol{O}_2$ at some particular moment of time if, at that moment of time, $\boldsymbol{O}_1$ and $\boldsymbol{O}_2$ share all their properties in common. The philosophical task confronting the negative, or bundle, theorists then becomes one of examining under what conditions this principle might turn out to be true.

Questions of numerical identity are interchangeable with questions of numerical difference. For, if we can state under what conditions $\boldsymbol{O}_1$

and $\boldsymbol{O}_2$ are numerically identical, then we have also stated, implicitly, under what conditions $\boldsymbol{O}_1$ and $\boldsymbol{O}_2$ are numerically distinct. Simply: $\boldsymbol{O}_1$ and $\boldsymbol{O}_2$ will be numerically distinct when the conditions for their being numerically identical do not obtain. Thus we may interchange the two questions at will. This explains why we often find theorists, both those promoting a negative theory and those promoting a positive theory, focusing their attention on examples of numerical difference as readily as on numerical identity. From a psychological point of view, it may be somewhat easier to focus on numerical difference than on numerical identity. But both approaches are equally valid.

10.4 The metaphysical and epistemological dimensions of the problem of individuation

There are two quite distinct, although closely allied, problems of individuation.

On the one side, there is the metaphysical problem. What must we assume about physical things *themselves* in order to account for their numerical distinctness? If possible, we want this answer not to invoke any special perceptual or cognitive abilities of conscious creatures since we can readily imagine possible worlds in which there are distinct physical objects and yet which do not contain any conscious creatures. We should no more want to make the numerical distinctness of physical objects depend on some feature or capacity of conscious, perceiving, knowing creatures than we should want, for example, to make the charge on an electron, or the mass of a proton, or the speed of light depend on some fact about *us*.

On the other side, there is the epistemological problem. We human beings *can* and regularly *do* individuate objects, i.e. we are able to distinguish them one from another, and we can even *count* objects (in stipulated regions of space and intervals of time).[4] How are we able to do this? What is there about physical objects and about us which allows us to individuate them? Do we human beings have 'access' to whatever it is which – metaphysically speaking – individuates material objects, or – somehow – do we individuate objects in some other way? But if so, how?

It is a regrettable fact about English that we use the same word

4. Our being able to *count* objects is often taken as equivalent to our being able to *individuate* objects. See, for example, [217].

"individuation" both for the metaphysical basis of numerical difference and for our human ability to discern the multiplicity of things. We could, somewhat artificially, deliberately avoid the verb "individuate" in the latter case and take recourse instead to something like "tell things apart". But the trouble with that maneuver is that it is only a temporary solution. Other writers standardly use "individuation" in both senses, the metaphysical and the epistemological. We will follow suit. But we must be aware that although we use a single word, there are two different concepts at play. What we would like is a solution to both problems of individuation, the metaphysical version and the epistemological version.

10.5 Positive theories: Substratum as individuator

Certain terms in philosophy – "form", "principle", "substance", etc. – have been in vogue for millennia, first in Greek and Latin and, later, in English translation, and have acquired a bewildering number of different meanings. "Substance", for example, occurs frequently in the writings of Aristotle and is used there in no fewer than six different senses (see e.g. [147]). In this chapter and in each of the next two, we will invoke respectively three different concepts of substance. In this chapter, we will examine whether substance need be posited to explain individuation (identity-at-a-time); in the next chapter, whether substance need be posited to explain identity-through-time; and in the last chapter, whether substance need be posited in order to explain personal identity.

Of the many concepts of substance, the principal one for attention in this chapter is that which is sometimes called "the substratum". Although the concept is certainly not original with Locke, it was his version – or at least a version conventionally attributed to him – which came to figure prominently in later discussions and is the one which modern philosophers most often have in mind when they talk of "substratum". Locke offered this explanation as to why he thought it necessary to posit substance:

> … when we talk or think of any particular sort of corporeal [material] substances[5] … though the *idea* we have … of them

---

5. Locke, in this first instance, is not using "substance" in his own technical sense, but in the colloquial or ordinary sense, in which one might talk, for

> be but the complication or collection of those several simple *ideas* of sensible qualities, which we … find united in the thing …; yet, because we cannot conceive how they should subsist alone, nor one in another, we suppose them existing in and supported by some common subject; *which support we denote by the name substance*, though it be certain we have no clear or distinct *idea* of that *thing* we suppose a support. ([124], book II, chap. XXIII, §4)
>
> … not imagining how these simple *ideas* can subsist by themselves, we accustom ourselves to suppose some *substratum* wherein they do subsist, and from which they do result, which therefore we call *substance*. (§1)
>
> … substance is supposed always *something* besides the extension, figure, solidity, motion … or other observable *ideas*, though we know not what it is. (§3)

This argument for the introduction of substance is a so-called argument to the best solution: it argues for the existence of something on the basis that that something must be posited as the best (if not the only possible) solution of some puzzle. Substance, it is clear, because it is itself propertyless, could not possibly be known either by sense or by scientific experiment. Its existence is established, not by empirical means, but – it is alleged – by rational means.

As he originally introduces, and justifies, the concept of *substance*, Locke does so for a particular purpose. The problem he is addressing in book II, chap. XXIII is that of trying to explain how the several properties of an individual thing occur at one place and do not seem separable from one another. The properties of gold, "yellowness, great weight, ductility, fusibility, and solubility in *aqua regia*,[6] &c., [are] all united together in an unknown *substratum*" ([124], §37). This first role for the concept of *substance* has, in more modern times, come to be referred to as 'substance-as-ontological-glue'. Substance, in this

---

example, of gold as being a kind of substance or of water as being a kind of substance. His own examples are of horse and of stone. Roughly, this first, ordinary, sense might be thought something akin to a natural kind of thing.

6. A mixture of nitric and hydrochloric acids

first sense, is regarded as that 'stuff' which 'binds together' a thing's properties and gives it its characteristic unity.

Four chapters later (XXVII), however, Locke addresses a quite different set of problems, viz. those of individuation (although only very briefly) and identity-through-time (at rather greater length). Again he invokes the concept of *substance* in trying to solve the problems he is examining. But Locke seems – like so many other philosophers who have taken recourse to the concept of substance – to overlook the fact that he is assigning a role to substance additional to its earlier one. For there is no logical connection whatever between the concepts of *substance-as-ontological-glue* and of *substance-as-individuator*. That a thing's substance 'binds together its properties' does not imply that that substance also accounts for the thing's numerical distinction from other things. Indeed, even if one were disposed to be sympathetic to the notion of substance, one could still ask: "Is the ontological glue which holds together a thing's properties the same substance or distinct from that substance which accounts for that thing's numerical uniqueness?" To have solved the former problem is not only not to have solved the latter problem, it is not even to have addressed it.

We must, then, distinguish between substance-as-ontological-glue and substance-as-individuator. It is only the latter which is of present concern.

Subsequent philosophers, most especially George Berkeley, found Locke's concept of *substance* unacceptable. Contemporary philosophers (e.g. J.L. Mackie [128], chap. 3) are somewhat dubious about Berkeley's objections, suggesting that perhaps Berkeley and many other philosophers have misinterpreted Locke, that Locke was merely reporting commonly held views and indeed may have been skeptical about their cogency himself. These are questions of historical scholarship best left for another sort of book. Here we need merely ask whether a concept of *substratum* along the lines historically attributed to Locke will solve the problem of individuation. In a trivial sense it will. For so long as substance is described as being something "I know not what", then one can simply assign to it by fiat whatever is needed to solve the individuation problem. Substance, on this account, becomes the individuator by stipulative* definition. Two qualitatively identical things will be numerically distinct, on this account, because their respective substances differ. It is of the essential nature of each substance to be numerically distinct from every other substance.

Whether or not they are tilting against a notion of substance which was uniquely Locke's, or whether they are objecting to a notion which is merely representative of a certain type of approach to solving the problem of individuation, a great many philosophers will eschew the concept of substance-as-individuator and will try to solve the individuation problem without resorting to such a notion. The objections to substance-as-individuator are threefold.

First is the fact that substance seems to be nothing more than an invented notion imbued with just those features needed to solve the problem. From a purely explanatory point of view, the methodology underlying the introduction of the concept of *substance* is ineffective. It is as if one were to try to explain, for example, why opposite poles of magnets attract one another by saying that each pole has within it 'an affinity for its opposite'. A new description has been given of the phenomenon, but nothing more by way of explanation.

The second objection stems from the fact that inasmuch as substance-as-individuator is essentially propertyless, i.e. its assigned role is to lie 'behind' or 'beneath' the properties of a thing, then it is essentially undetectable by sense or by scientific instrument.[7] It is as unempirical a concept as one can possibly define.

And third, and perhaps most significantly, substance-as-individuator solves the metaphysical version of the individuation problem but not the epistemological version. We can and do individuate physical objects. And just as surely we do not perceive their substances. We must, then, be able to individuate physical objects *on some other basis*. If we can explain how we are able to do this, then perhaps we will have no need to posit substance-as-individuator. Whatever it is which allows *us* to individuate things perhaps might serve as well to account for the metaphysical numerical difference of things.

10.6 Negative theories: Qualities and relations as individuator

Leibniz's attempted solution was a classic negative theory: he tried to solve the problem of individuation solely by recourse to the properties of things. His solution, recall (p. 233), was that every thing whatsoever has a set of properties unique to it, i.e. there are, as a matter of

---

7. The model of substance as lying 'beneath' the qualities (the monadic properties) of a thing may have some minimal virtue in evoking certain suggestive images, but it is hard to understand how substance might be supposed to ground the relations in which particulars stand.

fact, no qualitatively identical things in the universe. Although it is logically possible that there should be two or more qualitatively identical things, Leibniz argued that their existence would be incompatible with God's perfection. God sees to it that there are no qualitatively identical things, that any two things, however much apparently alike, do in fact differ in some detail or other, i.e. that *each* thing consists of a *unique* bundle of properties.

Apart from its unacceptable appeal to highly dubious theological principles to solve a metaphysical problem, there are two other flaws in Leibniz's solution. The first is that his solution, even if it were to be accepted as the solution to the metaphysical version of the individuation problem, would not solve the epistemological version. For the undeniable fact is that we are often able to individuate items *at a glance* without taking cognizance of any particular differences in their qualities, let alone – as Leibniz would have it – differences, in some instances, in their microscopic features. The second flaw is that Leibniz conceived of 'qualitative identity' and 'qualitative difference' in terms of monadic *qualities* alone. In dismissing the reality of dyadic (two-place), triadic (three-place), and higher-place relations (see section 9.2.6), Leibniz cut himself off from a possible solution which has attracted a great many contemporary negative theorists, viz. that what individuates physical things is not their qualities, but their relations.

But before we turn to examine relations as individuator (sections 10.6.2 and 10.6.3), we must examine some present-day attempts to argue, once again, for monadic qualities as individuator.

10.6.1 Relational properties; haecceity

Various recent philosophers have – like Leibniz – tried to argue that the principle of the identity of indiscernibles is true. But they do so nowadays on the basis that the principle is logically necessary, eschewing any appeal whatever to theological principles about God's perfection, etc.

Thomas Foster, for example, argues that two or more objects *must* differ in some property or other ([73]). His arguments depend on his claim that so-called relational properties (see above p. 258), e.g. *is-older-than-the-Eiffel-Tower* or *is-to-the-left-of-**a***, are bona fide monadic qualities applying to single individuals and are not merely disguised relations – e.g. *is-older-than* or *is-to-the-left-of* – holding between ordered pairs*[8] of individuals. (Remember, the terminology is a bit confusing. *Relational properties* are not relations: they are monadic qualities.) His argument may be summarized this way:

> Consider two numerically distinct objects, ***a*** and ***b***, that are separate from one another. Then ***a*** has the relational property (monadic quality) of *being-separate-from-**b***. But if the principle of the identity of indiscernibles were not true, i.e. if it were possible for two distinct things to have *all* properties in common, then ***b***, too, would have to have the property of *being-separate-from-**b***, i.e. of being separate from itself. But this latter condition is impossible; nothing can be separate from itself. Thus the principle of the identity of indiscernibles is necessarily true and any two numerically distinct things *must* differ in their (monadic) properties. That is, qualitative identity between numerically distinct things is a logical impossibility.

There are several objections to Foster's argument.[9] Chief among them is his treating *n*-place relations as monadic qualities rather than as relations.

Unlike Foster, some other philosophers have argued that it is logically possible for two or more numerically distinct things to share all their properties in common. Max Black, for example, has described a world consisting solely of two numerically distinct, but qualitatively identical, iron spheres ([31]). According to Black, such a world is logically possible; according to Foster, such a world is not.

"It all depends on what is meant by *sharing all properties in common*", we might say. If we allow that *is-separate-from-**b*** is a genuine quality (monadic property), then, of course, the two iron spheres must differ in their properties, since sphere ***a*** will have this property, and sphere ***b*** will not have it, indeed it is impossible that ***b*** should have this property. On the other hand, if we insist that *is-separate-from* is a *relation* holding between two things, then – clearly – ***a*** stands in this relation to ***b***; and ***b***, in its turn, stands in this very same relation to ***a***. In an account, then, which resists treating *n*-place relations as relational properties (i.e. as monadic qualities), the two iron spheres – just as

8. For definition of "ordered pair" in Glossary, see under "set".

9. Foster, himself, reviews some which I do not discuss. See [73].

Black hypothesizes – can share all qualities in common and can also share all *n*-place relations in common.

The debate turns, then, on whether or not to allow so-called relational properties (e.g. *is-separate-from-**b***) to be regarded as qualities (monadic properties). The major thrust of philosophical writing over the last century has been explicitly away from such a notion. Since Russell's initial forays into this area ([176.1]) in 1900, the theory has been that such relational properties as *is-separate-from-**b*** or *is-west-of-Toronto* should be regarded not as qualities instanced in single particulars (e.g. ***a*** and Edmonton, respectively), but as disguised proxies for two-place relations instanced in ordered pairs of particulars, e.g. in the ordered pairs ⟨***a***, ***b***⟩ and ⟨Edmonton, Toronto⟩. (We have earlier reviewed [section 9.2.6] the very considerable difficulties, insuperable perhaps, in trying to 'reduce' relations to monadic qualities.)

But even if there were not the pressures of logical considerations against Foster's relational properties, there would remain the objection that such 'peculiar' hybrid properties do not effectively solve the epistemological version of the individuation problem. If two numerically distinct physical objects, $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$, differ only in instancing, respectively, such properties as *being-separate-from-*$\boldsymbol{a}_2$ and *being-separate-from-*$\boldsymbol{a}_1$, it is very hard to see how we might take cognizance of such properties in discriminating between those objects. Suppose you were to take these objects, place them in an opaque box, shake them about in that box, and dump them out onto a tabletop. Clearly I would be able to see that there were two objects. But which is $\boldsymbol{a}_1$ and which is $\boldsymbol{a}_2$? Even though I would be able to see there are two objects, I could have no way of knowing whether the one on the left had the property of *being-separate-from-*$\boldsymbol{a}_1$ or had the property of *being-separate-from-*$\boldsymbol{a}_2$. *If* the item on the left is $\boldsymbol{a}_1$, then it has the property of *being-separate-from-*$\boldsymbol{a}_2$; and *if* the item on the left is $\boldsymbol{a}_2$, then it has the property of *being-separate-from-*$\boldsymbol{a}_1$. But which of these relational properties an item has will depend on its being $\boldsymbol{a}_1$ or $\boldsymbol{a}_2$; i.e. its relational properties depend on, and do not determine, its numerical identity.[10]

---

10. This little thought experiment – where the objects might exchange places – must not be confused with the problem of *re*-identifying $\boldsymbol{a}_1$ at a later time, T_2, with itself at an earlier time, T_1. (This latter problem will be dealt with in chapter 11.) The point here is to challenge the theory that we discriminate between items, e.g. $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$, by recognizing that $\boldsymbol{a}_1$ has the property of

Another approach, along lines similar to Foster's, is to restrict the class of relational properties to one, very special, class of individuation-conferring relational properties. In this latter account, individuation is secured for each individual, $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, …, $\boldsymbol{a}_n$, by that thing's instancing a monadic quality utterly unique to it alone, viz. the property *being-identical-to-*$\boldsymbol{a}_1$, *being-identical-to-*$\boldsymbol{a}_2$, …, or *being-identical-to-*$\boldsymbol{a}_n$. Such properties are said to be *haecceitist*[11] *properties* or *individual essences* (see e.g. [2], [126], and [122]).

All the objections already leveled against relational properties apply to haecceitist properties inasmuch as haecceitist properties are relational ones. But there is an additional objection to haecceitist properties (or individual essences): they seem to defy the very concept of what a *property* is supposed to be. If we examine our working inventory of 'ordinary' properties and relations – *redness*, *triangularity*, *being to the left of*, etc. – we find that each and every one of these is *general* in the sense that there is a large (potentially infinite) class of things which are eligible candidates for having that property or relation.[12] But haecceitist properties are strikingly different. There is one and only one thing in the entire universe which could possibly instance the 'property' of *being-identical-to-*$\boldsymbol{a}_1$, namely $\boldsymbol{a}_1$ itself. Similarly, there is one and only one thing in the entire universe which could possibly instance the 'property' of *being-identical-to-*$\boldsymbol{a}_2$, namely $\boldsymbol{a}_2$ itself; etc. Haecceitist properties seem to lack the very feature – generality – which one might well believe is essential to the concept of *property* itself.

Moreover, proposing haecceitist properties as a solution to the individuation problem seems to present a circularity. We want to know

---

being-separate-from-$\boldsymbol{a}_2$ and that $\boldsymbol{a}_2$ has the property of *being-separate-from-*$\boldsymbol{a}_1$. I contend that it is impossible to determine which item has which property prior to having discriminated between them. Put another way, I am arguing that to see that $\boldsymbol{a}_1$ in fact has the property *being-separate-from-*$\boldsymbol{a}_2$, one must already have seen that there are two objects. Relational properties do not account for numerical difference; it is, in fact, the other way around.

11. "Haecceity" (pronounced *hex′-ee-i-ty*) is the translation of the Latin "haecceitas" (literally "thisness"), a term coined by the medieval philosopher John Duns Scotus (c. 1266-1308).

12. This is true even of ordinal properties. Any number of mountains *could have been* the twelfth-highest mountain, although at most one actually is. (We will examine ordinal properties in the next subsection, 10.6.2.)

what it is about $\boldsymbol{a}_1$ which accounts for its numerical difference from $\boldsymbol{a}_2$. In the haecceitist account we are offered the ‘property’ *being-identical-to-*$\boldsymbol{a}_1$. But what kind of property is this? On the face of it, it looks like an invented ‘property’ having just the requisite features needed for the individuation of $\boldsymbol{a}_1$.

But even if one were to allow that haecceitist properties provided a solution to the metaphysical version of the problem of individuation, it is unclear that they solve the epistemological problem. Can one literally see such properties? If I am able to tell by looking at two highly similar physical objects (e.g. newly minted pennies pressed from the same dies) that they are numerically distinct, do I do this by recognizing that the first (on the left) has the property of *being-identical-to-*$\boldsymbol{a}_1$ and the second (on the right) has the property of *being-identical-to-*$\boldsymbol{a}_2$? Again, as in the case of ‘ordinary’ relational properties, suppose someone switches $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ without my seeing her do so. I can still see that there are two objects although I might well now believe (mistakenly) that the one on the left has the property of *being-identical-to-*$\boldsymbol{a}_1$. Even though I now have got their haecceitist properties wrong, I have still managed to individuate the objects. How could I have done this, if haecceitist properties account for individuation?

If there are haecceitist properties, then there must be as many of them as there are individuable things in the world. Moreover, to be able to individuate things never before encountered (e.g. the individual flakes of corn in my cereal bowl), I must be able to respond to the stimulus of their haecceitist properties, without of course ever having encountered such properties before in my life. I find such a theory implausible. From a psychological point of view, that is, of being able to learn one’s way about in this world on the basis of past experience, one must assume an ability to generalize from acquaintance with a variety of properties: redness, triangularity, etc. But haecceitist properties (remember these are monadic qualities, not *n*-place relations) are never instanced in more than one thing. How could one ever, then, prepare oneself to respond correctly to a haecceitist property not yet encountered? How could I possibly have an ability *now* to individuate flakes of corn tomorrow, if individuation comes down to recognizing the haecceitist property (individual essence) of each individual flake? I cannot see how this would be possible.

My objections to haecceitist properties are not conclusive. To me, proposing that $\boldsymbol{a}_1$’s having the property of *being-identical-to-*$\boldsymbol{a}_1$ is what accounts for the numerical difference of $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ strikes me as not particularly attractive as a solution. It ‘feels’ too much like simply having invented a variant description of the very thing to be explained and then offering that description as explanation. But can we do any better? We can find out only by trying.

10.6.2 Ordinal properties

The theory that ordinal properties individuate has the considerable attraction that it invokes no esoteric or ‘peculiar’ properties like relational or haecceitist ones. Ordinal properties are gratifyingly ordinary and familiar. Examples include ‘is the tallest man’ and ‘is the fourth largest freshwater lake’ (Quinton’s examples, [165], 15) and ‘is the first dog to be born at sea’ (Strawson’s example, [200], 26). Ordinal properties assign their bearers to a unique position in an ordering (first, fourth, etc.). Although ordinal properties – like all non-haecceitist properties – are general in that there is a potential infinity of things which are candidates for instancing the property, they differ from other properties in that no more than one thing in actual fact does instance the property. For example, vast numbers of things are red, and a still larger number are candidates for being red. But of these many things – the red things and the possibly red things – no more than one can be *the reddest thing*; no more than one can be *the fifth reddest thing*; etc. In this latter regard, ordinal properties would seem to offer especial promise of solving the problem of individuation: perhaps each and every thing instances a unique ordinal property, i.e. has an ordinal property which is proprietary to it and it alone.

Ordinal properties are not simple qualities; they are, in effect, ‘collapsed’ relations. When they are ‘spelled out’ using the resources of modern logic, they can be seen to invoke relations such as *is taller than* or *is born prior to*. For example, the sentence “Rob Roy was the first dog to be born at sea” can be reconstructed this way: “Rob Roy was a dog; Rob Roy was born at sea; all dogs, other than Rob Roy, that were born at sea were born *later than* Rob Roy.” Similar paraphrases, or reconstructions, in terms of relations can readily be given for “second”, “third”, etc.[13]

---

13. Although ordinal properties are ‘collapsed’ relations, they are not relational properties. The difference is that relational properties, such as *is west of Toronto*, refer to specific individuals (e.g. Toronto). Ordinal properties do not. In the reconstruction just given of “Rob Roy was the first dog born at sea” there is no reference to any individual other than Rob Roy.

As we pass from “first” to “second”, and from “second” to “third”, etc., As promising as ordinal properties may initially appear as offering a solution to the problem of individuation, it turns out that, in the end, they do not. There are two problems.

The first problem is already familiar. Just as with the theory that numerical difference is to be accounted for by difference in qualities (monadic properties), the theory that numerical difference is to be accounted for by difference in ordinal properties suffers at the hands of the requirement that it solve the epistemological problem of individuation. For we regularly individuate things without, for the most part, being in the slightest aware of their ordinal properties. You can glance at two pencils, fresh out of the box, lying on your desk and tell that there are two of them without knowing, or indeed having any way of finding out, which is the older, the slightly longer, the heavier, etc.

The second problem, too, recalls a difficulty we have already seen. Except for the possible exception of haecceitist properties, there is no convincing logical principle requiring that any two things must differ in one or more qualities. Similarly, there is no logical principle requiring that any two things must differ in one or more ordinal properties. Of course no two things can be 'the tallest man' and no two things can be 'the first dog born at sea'. But there is nothing to prevent there being two or more men, of identical height, being taller than all other men, or there being two or more dogs all being born simultaneously at sea and prior to the birth of all other dogs born at sea. In short, there is no logical guarantee that numerically distinct objects *must* differ, or any factual guarantee that numerically distinct objects *do* differ, in at least one ordinal property, from one another.[14] Even if it is not in fact

---

the complexity of the reconstruction grows exponentially. The reconstruction of "Riff Raff was the second dog born at sea" is: "There was some *x* such that *x* was a dog and *x* was born at sea and *x* is other than Riff Raff; Riff Raff was a dog, and Riff Raff was born at sea, and Riff Raff was born *later than x*; and for any *y*, if *y* is a dog and *y* is born at sea, and if *y* is other than *x* and *y* is other than Riff Raff, then *y* is born *later than* Riff Raff."

14. Quinton thinks that there is another problem as well, viz. "we can … only ascribe ordinal properties to things in theory if they are finite in number" ([165], 16). This is a mistake. Consider an infinite class of objects whose lengths are as follows: 1/2, 2/3, 3/4, 4/5, 5/6, etc. Each item in this class has a unique ordinal position, 1st, 2nd, 3rd, etc. While it is true that there is nothing which is 'the longest member in the class' (i.e. is the last item of the ordering), it is still nonetheless true that every one of the infinite number of items in the class has a unique ordinal property. (To be sure, there are some infinite classes which are 'open at both ends', e.g. {… 1/5, ¼, 1/3, 1/2, 2/3,

actual, we can at the very least imagine a possible world (again, for example, Max Black's possible world of two qualitatively identical iron spheres [p. 286]) where every object has exactly the same qualities as every other and no object has an ordinal property unique to it alone. Yet, by hypothesis, in such a world there are two or more objects. Their numerical difference is yet to be explained.[15]

10.6.3 Spatial and temporal relations

Although the theory that physical things are individuated by ordinal properties cannot be sustained, our examination of that theory has helped to reveal just what it is that we seek: some specifiable and recognizable property, or bundle of properties, which each thing instances and which – by its very nature – cannot be instanced by more than one thing. *Is* there any specifiable bundle of properties which has this feature?

Many negative theorists argue that a physical thing's position in space and time precisely has this sought-for feature, and, thus, position in space and time is what finally individuates physical things. It is important to understand that it is *conjoint* position in space and in time which is alleged to be the individuator. Position in space alone is insufficient; so too is position in time. For the property of being at

---

3/4, 4/5 …} which may be ordered by the relation "is larger than", but whose members cannot be assigned ordinal properties. We often describe such classes just in terms of their lacking certain *ordinal* characteristics, saying of them that "they lack a *first* member and lack a *last* member". Thus, it is not the infinitude per se of a class which precludes its members instancing ordinal properties, but its being open-ended at both ends.) This correction of Quinton's error is irrelevant to the two objections made in the text above.

15. There is a third problem as well, perhaps in the end even more serious than the two just mentioned. The essential occurrence of the relation *other than* in every reduction of an ordinal property to 'standard' (non-numerical) relations bears comment. (See the example, p. 290, of the reconstruction of "Rob Roy was the first dog to be born at sea".) What does "other than" mean? The most natural way to interpret this relation is that it is equivalent to "not identical to" or "is numerically distinct from". In that case, then even to invoke an ordinal property is to have presupposed a solution to the individuation problem. That is, ordinality, it would appear, is logically dependent on individuation; we will not be able in a noncircular manner to explicate numerical difference in terms of ordinality.

some specific place, e.g. being at P_1, is a property which any number, indeed a potentially infinite number, of physical things might instance (by being at that place at different times). Over the last ten years or so, there will have been hundreds of different books, one after the other, on my desk (i.e. at P_1). The property of *being at* P_1 is one which many, many physical things have (at one time or another) instanced. Similarly, the property of *being (or existing) at* T_1 is one which many things all instance.[16] At this very moment (i.e. at T_1), vast numbers of physical objects – some close at hand, others at appreciable distances – instance the property of *being at* T_1.

But combine the two properties – *being at a place* and *being at a time* – and the number of bearers collapses. Pick any specific place at any specific time: there is at that place either no things or exactly one; there cannot be two or more things there. Quinton waxes metaphorical on this point, soaring to heights of inspired prose:

> There is no limit to the number of things which can be present at a particular map-reference [position in space], provided they occur there at different times. Equally there is no limit to the number of things that can be in existence at a particular moment of time, provided that they are to be found at different places. But this boundless promiscuity of positions in space and time considered separately is replaced by the most rigorous propriety when they are conjoined. A complete, that is to say spatial *and* temporal, position is either monogamous or virginal, ontologically [metaphysically] speaking. ([165], 17)

Positional properties – i.e. *being at some specific place at some specific time* – have one marked advantage over ordinal properties for purposes of individuation. While there is no guarantee that any given physical object instances any ordinal property whatever (recall the example in the preceding section of the two equally-old dogs born at sea), every physical object, it would seem, *does* occur at a unique position in space and time. While most conjoint positions in space and

16. Special relativity theory tells us, contrary to our naive intuitions about physics, that the relation *is simultaneous with* must be treated with extreme care. Two events at different places may turn out to be simultaneous with one another viewed from one vantage point (what physicists call a “reference frame”) and not simultaneous when viewed from another.

time are empty (there is – colloquially speaking – more empty space in the universe than occupied space), it is still nonetheless true that, 'viewed the other way round', each and every physical object in the universe occupies a position in space and time unique to it alone.

Position in space and time would seem to solve not only the metaphysical version of the individuation problem but the epistemological version as well. *We* individuate objects by attending to their positions in space and time. There are two black pens currently on my desk. For all intents and purposes they share all their monadic qualities in common. If you were to swap them, i.e. exchange their positions, without telling me or showing me, I would not detect that you had done so. And yet I have no difficulty at all in seeing that there are two of them. How *do* I tell them apart? How do I count them? Not by detecting some subtle differences in their qualities. My sight reveals nothing whatever different in their qualities. Nor do I discern their numerical distinctness by taking cognizance of their 'substances'. In this latter regard, I would not know even how to undertake such a task. Rather, I detect their numerical difference at a glance, simply by seeing that they are, at one particular time, each in a place different from that of the other.

Thus, it would appear that we have finally solved both the metaphysical and the epistemological versions of the problem of the individuation of physical objects. Position in space and time is what individuates. There apparently is no need to postulate a mysterious 'substratum' as individuator.

But philosophers are seldom inclined to 'leave well enough alone'. In philosophy, just as in science, there is no natural stopping point. Having given one explanation or theory, there is then always the inclination and desire to delve deeper yet, to try to understand what might account for the truth of the latest theory. In biology, for example, Gregor Mendel (1822-84) – knowing nothing of the 'mechanisms' of heredity – was able to advance a theory which correctly predicted the statistical transference of features from parent to offspring. But what lay behind these statistical laws? It was to be many years before chromosomes were to be discovered in the nuclei of cells, and further decades still before DNA was to be discovered in the chromosomes.

At one level of analysis, the puzzle of individuation is solved: it is position in space and time which individuates. But this seeming solution to the problem of individuation is not quite the end of the matter. For upon probing, this latest answer, in its turn, is found to harbor some further puzzles themselves in need of examination. Just why is it – we might be inclined to wonder – that position in space and time is, as Quinton rakishly puts it, virginal or monogamous? Why is it that no two physical things are in the same place at the same time?

And there is a second problem as well. If, as was argued in chapter 8, we want to posit only a relative space and not an absolute space, then the position of any physical object is not a monadic quality, but a two- or many-place relation. But if position in space (and equally position in time) is a relation, it would seem to be a relation between *numerically distinct* objects. If so, that would, in turn, suggest that to individuate any one object we would antecedently have had to individuate another. The account of individuation in terms of spatial and temporal position would appear to be in imminent danger of becoming circular or of presupposing an infinite regress.

We turn to these latter two problems in the next two subsections.

10.6.4 Impenetrability

The principle that no two physical things can be at the same place at the same time is often referred to as 'the principle of the impenetrability of material objects'. This principle merits careful examination.

There are two perfectly straightforward ways in which two numerically distinct things *can be* in the same place at the same time.

First, any physical object is at any one time at all the places where any of its parts are. My bedside radio contains, as one of its several parts, a loudspeaker. That loudspeaker is located at some specific place, P_1. In response to the question "What is located at P_1?" it is perfectly proper to reply in either of two ways: (1) "There is a loudspeaker there"; and (2) "There is a radio there". If we designate a *small* place, small in the sense that it is contained within a part of a larger thing, then there will be two things at one place: the part of the thing, and the thing of which it is a part. Remember (section 8.8), a thing does not have to have all of its spatial parts present at a place for it to be at that place. The Mississippi River exists at Natchez although other spatial parts are hundreds of kilometers distant.

Second, if we choose a place which is *large* enough, then it can easily contain two or more objects. My office at the university is a place. It contains not just two, but several objects: a typewriter, two filing cabinets, a desk, a telephone, etc. In this latter example, several objects all occupy the same place at the same time.

Clearly, in their attempt to solve the problem of individuation, when negative theorists invoke the principle of the impenetrability of material objects, they mean to exclude such cases. It is, however, a tricky business to state the intuitive principle rigorously so that it does not immediately fall victim to counterexamples. It is easy to see why. As a first attempt, one might try to state the principle more precisely by formulating it this way: "A place which contains the part of some object $\boldsymbol{O}_1$ cannot simultaneously contain a part of some other object $\boldsymbol{O}_2$." But this first try is immediately refuted by the fact that you can place your telephone in the drawer of your desk. There would then be in one place two different things: your desk and your telephone. This latest difficulty arises from the fact that many material objects contain vacuities within their boundaries (e.g. an empty drawer in your desk), and it is perfectly feasible to place other objects within these vacuities.

Persons who have promoted the principle doubtless have had in mind a certain limited number of examples. These examples likely include such commonplace facts as these: Pots do not sink into stove-tops; the metal surface of the burner 'excludes' the pot. One's fingers do not penetrate the keys of a typewriter; although they may depress those keys, human fingers do not enter into the 'internal space', i.e. the interior, of the keytops. One might drive a nail into a wooden beam. But the steel of the nail does not merge into, i.e. form a composite with, the wood fibers; instead the nail 'pushes aside' the wood of the beam and replaces the contents of that region of space, which had been of wood, with steel.

The problem with trying to capture these latter examples in a defensible principle is that not all material things behave in this way. Stovetops and typewriter keys may, but galaxies do not. Although a galaxy may contain thousands of millions of stars, the distances between individual stars are so enormous that it is possible for one galaxy to pass through another and for both to emerge from the collision, not unscathed, but at least identifiable as galaxies ([192], 345-7). But are galaxies material objects? They certainly satisfy most, if not all, of the conditions we pre-analytically ascribe to material objects: their parts are material; they have a certain physical 'cohesiveness' in that they move about in space preserving their general size and shape; they are held together by physical forces; etc. And yet galaxies *are* able to pass through one another.

Someone might protest: "the difference is that galaxies are mostly empty space; bona fide (i.e. 'real') physical objects are far more dense (compact); there are no great 'open spaces', as it were, in stovetops, our fingers, or steel nails." But the trouble with this line of rebuttal is that it is false. The surprising truth is that, viewed at the atomic level, our flesh, our stovetops, our steel nails *are* mostly empty space. Compared to the pressures prevailing in the interior of white dwarfs (collapsed stars), those affecting physical objects in or near the vicinity of the Earth's surface are relatively slight. The material of white dwarfs is the same basic sort of stuff of which stovetops and fingers are composed, viz. electrons, protons, neutrons, etc. But because of the difference in pressure, the density of the interiors of white dwarfs[17] is a millionfold that of water. And thus, while the average density of the physical objects which are common in our own local environment is, in fact, vastly greater than the average density of a galaxy, it remains very much less than the average density of certain stellar interiors. In short, the fact that stovetops 'exclude' the pots placed upon them is a more-or-less 'local' feature of some familiar, ordinary, room-temperature objects. It is not a feature which holds of all material objects, e.g. it does not hold of galaxies.

Over the years, writers have expressed a great many differing attitudes about the logical (i.e. modal) and the epistemological status of 'the' principle of impenetrability. Newton (1687) had thought that we learn the principle by experience: "That all bodies are impenetrable, we gather not from reason, but from sensation. The bodies which we handle we find impenetrable, and thence conclude impenetrability to be a universal property of all bodies whatsoever" ([144], vol. II, 399). Locke (1690) concurred, and expanded Newton's claim, arguing that not only is the idea that all objects are impenetrable learned from experience but, further, the very concept of *impenetrability* itself 'arises' from experience.

> The idea of *solidity* we receive by our touch; and it arises from the resistance which we find in body to the entrance of any other body into the place it possesses, till it has left it. There is no *idea* which we receive more constantly from sensation than *solidity*. Whether we move or rest, in what posture soever we are, we always feel something under us that supports us[18] and hinders our further sinking downwards; and the bodies which

---

17. 10^6 g/cm^3 ([192], 126)

18. This claim – that we always feel something under us that supports us – is overstated. Most of the time we are not consciously aware of the support of

> we daily handle make us perceive that, whilst they remain between them, they do, by an insurmountable force, hinder the approach of the parts of our hands that press them. … If anyone think it better to call it [i.e. solidity] *impenetrability*, he has my consent. … And though our senses take no notice of it, but in masses of matter, of a bulk sufficient to cause a sensation in us: yet the mind, having once got this *idea* from such grosser sensible bodies, traces it further and considers it, as well as figure, in the minutest particle of matter that can exist, and finds it inseparably inherent in body, wherever or however modified. ([124], book II, chap. IV, §1)

These passages both argue that the principle of impenetrability is learned by experience (i.e. a posteriori) and strongly suggest, but do not state explicitly, that their authors regard the principle as a scientific truth and not as a necessary (or logical) truth.

Certain contemporary philosophers take a quite different point of view. Quinton, for one, takes the principle to be a necessary truth. (Here he is using the expression "metaphysical truth" much as I use the expression "necessary truth"*.)

> If we are confronted by two distinct things between which we can find no strictly qualitative difference of length or weight or colour we can always distinguish them by reference to their respective positions. What proves this is the familiar but highly important metaphysical truth that no two things can be at the same place at the same time. … Individuals [i.e. material objects] are, to use an old word, *impenetrable*, which does not mean that they are never soft or porous. ([165], 17)

(In his reference to "soft or porous", Quinton is likely being motivated by a similar sort of distinction discussed by Descartes ([55], vol. II, 225-6).)

Quinton offers no argument for his claim that the principle of the impenetrability of material objects is true. He is content to declare it a

---

the floor on which we are standing or of the chair on which we are seated. Locke has here confused what we are *capable* of feeling, whenever we like, with what we *do* feel at any time.

'familiar' and 'highly important' necessary truth and to proceed from there.[19] But just as sure as Quinton is of the necessary truth of the principle, other philosophers have been just as sure of its contingency. Friedrich Waismann (1897-1959) offers the following possible-worlds tale:

> Suppose there were two chairs, *A* and *B*, with exactly the same characteristics. The fact that they are in principle distinguishable depends on the property of impenetrability. Suppose now that we lived in a world in which experiences of the following sort were everyday occurrences. When the two chairs are put so that they touch, and pressed together, they gradually merge into one chair; then turn back into two [qualitatively] identical chairs. ([209], 201)

For Waismann, the principle of impenetrability is no necessary truth. He has no difficulty whatever in imagining a possible world in which that principle is false. In his hands the principle is nothing more than a physical law of this particular world.

But can Waismann's possible-worlds tale really be sustained? It is, we note, exceedingly brief. If we were to try to fill in details, what might we find? In particular, although Waismann talks, ostensibly, about 'chairs', can we really regard 'chairs' which merge into one another as 'real' physical objects, or is Waismann playing fast and loose with the concept of *physical object*? Can things behave like Waismann's chairs, i.e. act in violation of the principle of impenetrability, and still be regarded as physical objects?

I think that Waismann was exactly right. But I also think that his possible-worlds tale needs some further elaboration to make it a plausible counterexample to the alleged logical necessity of the principle at stake.

The legions of philosophers who have supposed that impenetrabil-

---

19. He does, however, slightly qualify his claim, allowing for one exception, viz. the first of those we discussed a moment ago: "The only apparent exception to this rule is not an exception in principle. A whole is at every place and time that its parts are. Wholes and parts share positions. But they share only some of their positions. If A has B as a part they are not indistinguishable in position. For although all of B's positions are also A's, not all of A's positions are B's as well" ([165], 17).

ity does, in some way, figure in our concept of what it is for something to be a material object were on the right track. Where many of them were mistaken, however, was in supposing that *every* material object must *at all times* be impenetrable. In contrast, I would argue that a 'thing' can earn the status of material objecthood by instancing the property of impenetrability from time to time and from circumstance to circumstance.

> Imagine a possible world where select objects are found to be able to merge into one another. Two chairs which share all their qualities (monadic relations) in common, for example, set upon a 'collision' course are found not to rebound from one another, but to merge into one place and subsequently to separate. The spectacle admittedly would be highly surprising to a native of this universe. Any of us, on the first occasion of seeing such an event, would probably be inclined to reckon it a hologram or a conjuring trick. But as we become more familiar with the physics of that world, we recognize that this is no hologram, no conjuring trick. These chairs which are able to merge into one another are genuine physical objects. Even though they are *inter*penetrable with one another, they are not interpenetrable with other things. They are tangible, publicly observable, and as permanent as any other fixtures in the environment, e.g. tables, trees, and lampposts. We can, if we like, sit on these chairs, move them about, paint them, weigh them, burn them, etc. In short, they satisfy a great many other properties typical of material objects. The only caution is that we would have to keep these chairs from touching one another, for when they do, they temporally 'collapse into one another'.

We must not generalize too extensively on this latter possible-worlds tale and posit a world in which every material object was interpenetrable with every other, for then we would lose entirely our grasp of what it is for something to be a material object. A chair which was interpenetrable with everything else could not be sat upon, or painted, or cut into pieces, etc. It would not be a *physical* object at all. While we can allow interpenetrability for some physical objects under some circumstances, we cannot allow it for all under any circumstances.

In our being able to tell this latter possible-worlds tale, the principle of impenetrability is revealed to be no necessary truth at all, but only a contingency, i.e. it is a proposition which is true in some possible worlds, but not in all. This does not mean that the principle cannot be invoked in an attempt to solve the problem of individuation. What it does mean, however, is that the *justification* for using that principle must rest in experience, not in a priori reason. If typical physical objects in this world are individuable because they are impenetrable to one another, then that is a physical law of this particular world, it is no necessary truth.

What is emerging in our researches is the realization that our ability to individuate physical things, indeed numerical difference itself in physical things, is attributable to certain contingent facts about this particular world. We have found a solution to the individuation problem, but it is by no means one suited for any possible world whatever. Had this world been much different from the way it is, our ordinary means of individuating, indeed our very concept of individuation, would have to be different.

Just for a moment, let us explore how our techniques of individuating physical objects would have to differ in yet another world, where objects are more penetrable than they are in this world. We will pick up, and continue, our last possible-worlds tale.

> Suppose we are able to construct a machine which turns out objects which look like single chairs. But we discover that these are actually Waismannesque chairs: whenever these (seemingly single) chairs are left in bright hot sunlight for an hour, they suddenly separate into two qualitatively identical chairs (i.e. into two chairs sharing all qualities in common). The original chair undergoes something like biological mitosis (cell division) except that the resulting products each have the same mass as the original. (In the world being described, mass is an intensive property [see section 9.2.4], not an extensive property.) The original chair, we are happy to report, cannot separate when someone is sitting on it because that person's body would shield the chair from the sunlight. Chairs, as they come off the assembly line, are called "coupletons"; after they undergo splitting, they are called "singletons". Singletons cannot be further 'split'.
>
> In merely *looking* at a chair, no one is able to tell whether it is a coupleton or a singleton. To ascertain whether it is a coupleton or a singleton, one must either know that it is fresh off the assembly line or subject it to test, by placing it in bright hot sunlight, to observe what becomes of it.

> In such a world, the problem of individuation cannot be treated more-or-less independently, as it is in *this* world, from the problem of identity-through-time. How many chairs are in a room at any one time depends not only on how many seemingly distinct chairs there are but essentially on how many of those apparently single chairs are in fact coupletons and how many are in fact singletons. But to answer that question depends on determining facts about the chairs' history or future. In that world, no philosopher has ever suggested that the problem of individuation could be handled separately, indeed antecedently to, the problem of identity-through-time. Where *we* see two problems, they see only one.[20]

In searching for a solution to the problem of individuation, we discover, once again, how our concepts are tailored to the contingencies of the world in which we find ourselves. There is no guarantee that the physical objects of a world, more particularly the ready-at-hand most familiar objects of our experience, would exhibit the degree of impenetrability that we find in this world. We can describe other possible worlds in which objects are more penetrable, and still other worlds in which objects are less (e.g. in which galaxies are torn apart when they collide). We simply find ourselves ensconced in a world where familiar physical objects are so constituted that the principle of impenetrability can be used fairly successfully in our creating a workable concept of individuation. But the principle is no necessary truth, and it probably cannot be stated very precisely. It is a *vague* principle, adopted and adapted for our everyday needs.

---

20. One can, in fact, describe a series of possible worlds where solutions to the problem of individuation become progressively more remote from the solution in this world. We can, for example, describe possible worlds in which some objects undergo mitosis but without there being anything in their history or among their manifest properties (see section 9.2.2) to account for which will and which will not. In this latter world, whether a chair is a coupleton or a singleton is a matter solely of a disposition. But some chairs will have been burned without their ever having been exposed to sunlight. Given the occurrence of such circumstances, where some physical objects are never subjected to test, individuation could be explicated only counterfactually, i.e. it will have to contain a component of this sort: "*x* is a coupleton if it were to be subjected to test and were to undergo mitosis."

10.6.5 Does individuation presuppose absolute space?

Far more troubling than the vagueness and the contingency of the principle of the impenetrability of material objects is the precise role played by spatial and temporal relations in the negative theorist's solution of the problem of individuation.

If one adopts a theory of relational space and relational time, then a physical object's having a determinate spatial and temporal position involves its standing in spatial and temporal relations to *other* things in the universe. We locate physical objects in space by specifying their whereabouts *in spatial relation to* other things. My scissors are on the desk blotter which, in turn, is on my desk. The city of Burnaby is immediately east of Vancouver. And we locate physical objects in time by specifying their whereabouts *in temporal relation to* other things. The library at Simon Fraser University was built before the Administration Building, and it, in turn, was built before the engineering laboratories. Even when we set up coordinate systems, by which we might say that something is located at 116°32′W and 42°N, or that something exists on 18 June 1823, we are tacitly depending on those coordinate systems having been fixed by reference to specific places (Greenwich and the two poles) and to the conventionally assigned date of the birth of Jesus.

But if fixing the position of some physical object in space and time involves a reference to *other* physical things, then the negative theorist's proposed solution to the problem of individuation would seem to involve either an infinite regress or circularity. To be able to individuate one thing, it would seem that we would have had to have individuated another.

One might think that this current difficulty constitutes good reason for positing an absolute space and an absolute time. But, on reflection, we can see that an absolute space and an absolute time will not solve the problem either. The trouble with an absolute space and an absolute time is that they are each *amorphous*, i.e. each point of space is indistinguishable *to us* from any other point of space, each point of time is indistinguishable *to us* from any other point of time. Thus, even if one were to adopt a theory of absolute space and time, one would still have to take recourse to relations to determine an object's position. Newton, who did posit an absolute space and time, saw this difficulty clearly: "… because the parts of space cannot be seen, or distinguished from one another by our senses, therefore in their stead we use sensible measures of them. For from the positions and distances of things from any body considered as immovable [i.e. taken as the fixed point of our coordinate system], we define all places; and then with respect to such places, we estimate [judge] all motions, considering bodies as transferred from some of those places into others. And so, instead of absolute places and motions, we use relative ones …" ([144], vol. I, 8). No theory of absolute space and time can solve the epistemological version of the individuation problem, but a theory of relative space and time seems to solve the epistemological version of the individuation problem at the unacceptable cost of an infinite regress or of circularity. Is there any way out of this bind?

Quinton thinks there is. He argues that, for each observer, there is *one* position which is, in a sense, epistemologically 'privileged' or 'primary', viz. that person's 'here-and-now': "The position where I am at the present moment is, then, the absolute point of origin of all my positional characterisations of things. It is the one position I do not have to pick out by its relation to something else and by their relation to which in the end everything else is individuated" ([165], 20). Even if you were to wake out of a deep coma, not knowing what year it was, or where on the face of the Earth you were, you would presumably still be able to individuate local objects, e.g. the furniture of your room, by reference to your own 'there-and-then'.

And yet, I think that in two different ways Quinton's solution, invoking each individual observer's own 'here-and-now', is unsatisfactory.

What Quinton has done is to underscore the fact that of all the possible points which one might choose as one's starting point, or so-called fixed point, on which to construct a phenomenological (perceptual) coordinate system, there is one point which is not arbitrary but which virtually forces itself upon us, viz. the point from which we 'view the world'. For visual space, this 'point' is our eyes. We 'look out on' the world from sockets on the front of our heads. But while this is true, it – in its own way – raises problems as serious as those it was meant to solve.

Quinton has 'reduced' the general epistemological version of the problem of individuation to that of individuating other things with respect to one's own body. But this is no ultimate solution. It is tantamount to my saying, "I can solve the individuation problem if you will allow me to begin with individuating some one thing, and then I can individuate others with respect to it." The question remains: "How are you to get the process started?" It does not matter whether the fixed point is one's own body, the Tower of Pisa, or lines of longitude and latitude. The problem remains basically the same: how can any of this be done without an infinite regress or circularity?

At some point in each of our own personal histories, we had to individuate our own body. None of us is born into this world knowing that he/she has a unique body. The connection between our bodies and our perceptions and kinesthetic sensations is something we had to *learn*, by trial and error, during the first few years of our lives. We had to *learn* the difference between consciousness and externality; we had to *learn* where our bodies 'leave off' and other things begin. Having done that, we were then in a position to rely on the correlation between certain kinds of sensory data and the external objects which cause that data as our guide in individuating, at a glance, external physical things. But that we are *now*, as adults, able to do this must not conceal from us the fact that we had to spend a considerable time sorting this all out as youngsters. What we do now habitually, we once did only by trial and error. We may now, as adults, use our bodies, and their unique positions in space and time, as the locus of our individual perceptual coordinate schemes, but this is not innate, it is learned. And what still needs explaining is what we must suppose true of the world so as to warrant our creating such a conceptual scheme.

You perhaps may detect that I am pushing the epistemological version of the individuation problem back toward the metaphysical version. That we are *able* to discriminate physical objects needs explaining. But it needs explaining, I would suggest, in a way which depends not only on facts about us, but equally on facts about physical objects which account for their being individuable *logically prior* to our discriminating between them. And this brings me to the major criticism I have of Quinton's solution of the epistemological version of the problem of individuation: it abandons the metaphysical problem.

No ultimately satisfactory solution of the problem of individuation can make individuation depend solely on abilities of human beings and overlook whatever it is about physical objects themselves which accounts for their numerical difference. If *we*, human beings, are able to discriminate among physical objects, they must have been *discriminable* before we exercise our abilities. We want an account of individuation which explains the numerical distinctness of physical objects not only in worlds in which there are conscious perceivers but in worlds devoid of consciousness altogether. Numerical distinctness of physical objects is not just a feature of worlds in which there are conscious creatures, but of lifeless worlds as well.

In short, Quinton's account of how *we* individuate objects leaves unexplained what it is about objects themselves which permits us to discriminate among them.

The dilemma is this. The metaphysical version of the individuation problem would seem to require positing an absolute space, a space 'in' which each and every physical thing would have a unique position distinct from that of every other thing and independent of the knowledge or perceptual abilities of any conscious creature. But an absolute space cannot solve the epistemological version of the individuation problem. We human beings cannot discern any difference between the points of an absolute space. We would seem, then, to have to take recourse to a relative space. But a relative space just *is* the set of spatial relations instanced by numerically distinct objects. That is, a relative space is logically dependent on our discriminating physical objects: the logical and psychological order is that objects are primary, and that space is derivative, i.e. space is 'constructed out of' the relations obtaining among physical objects which are antecedently individuated.

There are only two ways out of this dilemma. The first is the way adopted by Newton. He embraced *both* theories of space (see above, pp. 303-4). He posited an absolute space in which God, but none of us, would be able to know the absolute position of each thing. And he posited a relative space which we human beings use to individuate and to gauge the motion of physical objects.

Such a 'Newtonian' solution is to be resisted. In positing an absolute space, we invite the inevitable incoherence which we explored at some length in Buber's befuddlement. Having struggled, both in physics and in philosophy, for more than three hundred years to shed the incoherent theory of an absolute space, we should hardly want to reintroduce that theory in a desperate attempt to solve the metaphysical version of the individuation problem. And in positing a theory of relational space to solve the epistemological problem, we precipitate a vitiating circularity. In short, were we to adopt both the absolute and the relative theories of space, the one as a supposed solution to the metaphysical problem and the other as a supposed solution to the epistemological problem, we would create a composite theory incorporating the worst features of *both* of the originals. Moreover, the resulting theory would have more than just internal defects. Any such blended solution offends aesthetic sensibilities. It is untidy and profligate. It violates the methodological principle[21] not to multiply entities beyond need. It would grate against our desire for comprehensiveness in our theories. It would be unacceptably piecemeal.

I think there is another way, one which solves both the metaphysical and the epistemological versions of the individuation problem and which does not require the positing of an absolute space or saddle us with the circularity of a relative space.

10.6.6 The solution: A more radical negative theory

The problem of the individuation of physical objects will forever remain insolvable so long as negative theorists persist in trying to explain numerical distinctness in terms of space and of properties and relations instanced at positions in space. This is the post-Leibnizian picture with which negative theorists have been working for generations. But it is a mistake. To proceed along these lines is to have failed to appreciate the profound difference between a theory of relative space and a theory of absolute space. The typical negative approach to trying to solve the problem of individuation has been, in effect, an unwitting illicit amalgam of the two incompatible theories. These attempts have uncritically borrowed from the theory of absolute space the notion that the points of space, or perhaps I should say the places within a space, were distinct from one another independent of the objects 'in' that space. But this hybrid theory cannot be made to work. If one is going to adopt a theory of relative space one must be prepared for a more wholesale conceptual reorientation.

In a theory of relative space, space is itself, both in a logical and in a psychological sense, 'constructed out of' the spatial relations among numerically distinct physical objects. Objects are not 'in' space – that is the old, absolute theory. In the theory of relative space, what exists are physical objects at varying distances from one another and moving about with respect to one another; space is then 'constructed' out of these objects. What is logically primitive, and the conceptual foundation upon which the rest of the theory is constructed, is the existence of numerically distinct physical objects. It is little wonder, then, that it becomes impossible to explain the numerical distinctness of physical objects in terms of space: the conceptual order is being inverted. One can never explicate the primitive concepts of a theory in terms of that

21. Known as 'Ockham's razor'

theory's derived concepts: the explication becomes circular. And this is precisely the situation we find in many theorists' attempts to solve the problem of individuation. Through a historical accident, our having earlier posited a theory of absolute space and having not completely thrown off its conceptual shackles, many theorists persist in approaching certain problems as if they were still working with the earlier theory. They try to explain individuation by invoking properties of space, failing to recognize that a theory of relative space *presupposes* the very existence of numerically distinct physical objects. An incisive discarding of the remnants of the theory of absolute space is what is needed.

We must give up trying to explicate numerical distinctness in terms of space. It is not space which is our 'starting point' and the numerical distinctness of physical objects which needs to be explicated. It is the concept of *physical object* which lies 'at the bottom', as it were, and it is space (relative space, that is) which is 'constructed' – both metaphysically and epistemologically – on that foundation. In a world where there are no physical objects, there is no physical space. To fail to understand that proposition, and thus to fail to understand that the concept of *physical object* is more primitive than that of *space*, is to turn the conceptual order on its head and to saddle oneself with necessarily insolvable problems.[22]

Edwin Allaire, thirty-five years ago, argued similarly that no appeal to spatial relations could provide an ultimately satisfactory solution to the problem of individuation: "Relations – I'll stick with spatial ones – *presuppose* numerical difference; they do not account for it" ([7], 254). But Allaire was not content to make the numerical difference of physical objects an unanalyzable concept, as I have done. Instead, he posited what he and some others (e.g. Gustav Bergmann) called "bare particulars" to account for the numerical distinction: "Bare particulars are … the entities *in* things accounting for the numerical difference *of* things" ([7], 253). While bare particulars may, at first, seem like nothing other than Locke's substratum given a new name, Allaire explicitly denied that they were, and argued that, unlike the problematic substratum, bare particulars are supposed to be accessible to expe-

22. Although it is notoriously difficult to design experiments to reveal clearly the conceptual framework children devise, there does seem to be anecdotal evidence that children form the concept of *physical object* much before, indeed perhaps years earlier than, an abstract concept of *physical space*.

rience. In a paper written two years earlier ([6]), Allaire had argued that we are 'acquainted' with bare particulars, meaning by "acquainted" what Russell had meant when he introduced a technical definition for that term in 1912: "We shall say that we have *acquaintance* with anything of which we are directly aware, without the intermediary of any process of inference or any knowledge of truths" ([178], 46).

It is clear Allaire and I agree that we are directly acquainted with the numerical difference of physical objects. We human beings can literally *see* that there are numerically distinct objects.[23] But I decline to take Allaire's last step: positing 'bare particulars' to account for that numerical difference. I do not believe that there is any 'entity' (Allaire's own term) under or behind or accompanying or 'in' the physical object itself which accounts for that physical object's numerical difference from other physical objects. Still less do I believe that we are acquainted with bare particulars or that they are 'presented' to us like the properties of things, while not, of course, being properties themselves (see [7], 256). Such a theory of perception is empirically untestable and philosophically suspect. A bare particular, like a substratum, is nothing but an *invented* entity, posited by fiat as a solution to the problem. But just as in the case of substance, the posit amounts to idle wheel-spinning. Allaire was correct in arguing that spatial relations presuppose numerical difference of objects. But I resist, as creating a mere illusion of further analyzing the concept of numerical difference, his positing of bare particulars. At some point, analysis must end. I suggest that Allaire has gone one step, an unnecessary step, beyond that stopping point.

How, in the end, do we solve the problem of individuation? We solve the problem by recognizing that the concept of *physical object* is a primitive notion in our conceptual scheme and that the numerical distinctness of physical objects cannot be explicated in terms of anything (e.g. relative space) which is derivative in that scheme. This solution, which argues that the original question was in some sense improper, is bound to dissatisfy some readers. (Recall our discussion of the 'Barber paradox', p. 152 above.) On the face of it, the question

---

23. I am not making the stronger claim, however, that we literally see *how many* physical objects are within our field of view. There may be ten objects in my field of view, and I may be aware of all of them without being aware that there are ten.

"What accounts for the numerical distinctness of physical objects?" looks as if it ought to have a straightforward answer. Grammatically it is like the question "What accounts for the green color of grass?" But, upon examination of the problem, it can be seen that the concept of *physical object* plays a very special role in our conceptual scheme, one so basic that it does not allow for further analysis, at least not in terms of the categories *space* or *position in space* which require, themselves, to be explicated in terms of it.

In short, what it all comes down to is a radical, and I think necessary, conceptual about-face. It is not *space* which is the primitive concept in terms of which *physical object* is to be explicated; it is the concept of *physical object* which is primitive and in terms of which *space* is to be explicated. In the fifteenth edition (1952) of *Relativity*, Einstein said this in his preface, previewing a newly added appendix: "… space-time is not necessarily something to which one can ascribe a separate existence, independently of the actual objects of physical reality. Physical objects are not *in space*, but these objects are *spatially extended*" ([64], vi). And in that appendix itself, he writes: "It appears to me, therefore, that the formation of the concept of material object must precede our concepts of time and space" ([64], appendix v, 141). What is here true for physics is true as well for metaphysics. And once we have understood the order of priority, there is no special problem of individuation. At the very least, it is not a problem 'over and above' that of explicating *physical object* itself. We do not have a concept of *physical object* of which it is then proper to ask what accounts for the numerical distinctness of physical objects. The concept of *numerical distinctness* lies at the bedrock of our conceptual scheme; there is nothing more basic than it in terms of which it might be explicated. Numerical difference is not to be accounted for in terms of a 'presented' entity (a bare particular), of a 'know not what' (a substratum), or of any specially favored properties (e.g. spatial relations). To have the concept of *physical object* is already to have the concept of *numerical difference*. And the latter concept is not further analyzable.

In finding that a particular concept is not further analyzable, we have learned something of enormous importance about our conceptual scheme. In learning that the problem of individuation cannot be solved by appeal to substance, to bare particulars, to qualities, or to relations, we have learned that we had mistakenly got the order of logical precedence wrong in our thinking about these notions. We had imagined physical objects strewn about in space and their numerical difference accounted for in terms of their different positions in that space. Our studies reveal, however, that such a view is ultimately untenable. Our starting point must be that of numerically distinct physical objects lying at varying distances from one another and moving about relative to one another. From such a view we can construct a workable concept of space. But such a view is grounded upon the primitiveness of the notion of physical object and numerical distinctness. We cannot – on pain of circularity – try to explain those notions by invoking properties of space.

We may think that we have abandoned the theory of absolute space, have fully switched over to the theory of relative space, and have seen clearly the implications of the latter theory. But in cold fact, many of us have not fully done so. Some philosophers believe that the problem of individuation can be solved by invoking spatial relations, but in so doing they overlook the fact that spatial relations *presuppose*, and do not ground, the numerical difference of physical objects. Simply put, there are no places logically or epistemically independent of, or prior to, the existence of numerically distinct physical objects. The order of dependence is from physical things to places; not the other way around.

10.7 Nonphysical objects[24]

Identity is a two-place (dyadic) relation. In that regard, it is like *is taller than* and *is envious of*. But unlike all other two-place relations, identity is unique in that it can hold only between one thing and itself; it cannot – like *is the same height as* or *is the same age as* – ever hold between two numerically distinct things. One might think, then, that it would be particularly easy to establish in practice when the relation of identity holds. But it turns out that on occasion it is exceedingly difficult to do so. This is so because we often have several logically independent ways of individuating one and the same thing. We may individuate some 'thing' in one way and may individuate some 'thing' in another way, and we might not quite be sure whether in doing so we

---

24. I have explored the issues of this section previously in "Can the Theory of the Contingent Identity between Sensation-States and Brain-States Be Made Empirical?" in *Canadian Journal of Philosophy* 3, no. 3 (Mar. 1974), 405-17. I thank the editor of that journal for permission to reprint selected passages below. Most of this section is, however, newly written for this book.

have individuated two numerically distinct things or have individuated one thing twice.

For example, I sometimes talk about 'my red car' and I sometimes talk about 'my Oldsmobile'. Each of these descriptions individuates[25] exactly one thing in the universe: I have only one red car, and I have only one Oldsmobile. But if you were to overhear me using both expressions you might have no way of knowing whether my red car is (one and the same with) the Oldsmobile, or whether I have (at least) two cars, one red but not an Oldsmobile, and the other an Oldsmobile but not red. If, however, you were to have access to my cars and were to examine my red one, you would quickly discover that it is an Oldsmobile; or, if you were to examine my Oldsmobile, you would quickly discover that it is red. You would be able, as an empirical matter of relative ease, to establish the identity of my red car with my Oldsmobile.

With an example such as this in mind, one might get the mistaken idea that identity can always be readily established by direct observation. We need only, one might mistakenly believe, simply examine the referent of the one individuating description and examine the referent of the other individuating description and the identity (or not, as the case may be) of the 'things' referred to would be altogether obvious. But the example of my red car and of my Oldsmobile is not, in the end, a fair representation of the general methodology involved in identifying *A* with *B*. For there are large classes of 'things' whose identity poses methodological and metaphysical puzzles which are not solvable by simple observation or direct empirical examination. To identify one of these things with 'another' we will have to grapple with metaphysical issues which transcend mere observation.

In the case of the synchronic identity of physical objects, the focus of the problem is usually on trying to create a viable account of their numerical difference in the face of qualitative identity. But in the case of nonphysical objects, the focus is usually on the other side of the same coin, on the problem of trying to account for numerical identity in the absence of perceived qualitative identity.

We have already seen one example of this latter problem when I discussed, in chapter 8 (pp. 181-3), the case of identifying the causes of our visual sensations with those of our tactile sensations. Each of

25. Or, more exactly, is a uniquely referring expression which may be used to individuate

these kinds of sensations, the visual and the tactile, exists in a sensory 'space', i.e. comprises its own self-contained modality* (as it is sometimes described), and the problem becomes one of trying to 'identify across sensory modalities'. As I argued, there is no necessity that the data of various sensory modalities correlate highly with one another; it is a brute contingency of this world that they do correlate as highly as they do. Capitalizing on this correlation, as children we learned to synthesize a unitary sensory space out of these distinct kinds of data, so that now, as adults, we hardly if ever give a moment's thought to the profound differences between visual and tactile space. Apart from shape,[26] the features of things which we discern visually – e.g. their hues, their distances, their transparency or opacity – do not overlap with the features we discern tactually – e.g. their texture, their degree of pliability, their viscosity, their weight. And yet, in spite of the nearly complete difference in the two sets of data furnished by our eyes and by our hands, we posit the *identity* of the sources of the one with the sources of the other, saying that the thing seen is one and the same with, i.e. is numerically identical to, the thing felt. This identification is not a matter of observation (like identifying my red car with my Oldsmobile), but involves making a posit which carries us well beyond anything 'given' in perception or demonstrable by empirical means. The identity of the visual with the tactile is a posit which helps 'to make sense' of the sensory data, but is not itself anything perceived, nor, for that matter, is it anything perceivable.

If we fail to recognize that identification is possible, even when the descriptions of the 'things' to be identified are profoundly unalike, we may fall into the error some philosophers have made in arguing that it is impossible *in principle* scientifically to identify mental states with brain states.

For several generations, a certain element of popular metaphysics was in advance of that of professional philosophers. For the past hundred years at least, many persons who would not regard themselves as philosophers had adopted a particular view about the nature of human

26. Although both sight and touch seem to furnish the 'same' data pertaining to shape, we know that identifying visual shape with tactile shape is something that we had to learn to do; that ability is no more innate or logically mandated than is our ability to identify colors and temperatures (in those few cases where they do correlate, e.g. where an object is heated to incandescence). (See p. 181, esp. footnote 20.)

consciousness which was at distinct odds with what many skilled philosophers were then arguing. This popular view held that consciousness was, in some sense, to be identified with the (higher) workings of a central nervous system. Many philosophers, however, actively resisted this popular theory, not because they were reactionary by nature, but because they thought that the theory was logically flawed. They pointed out that mental states and brain states ostensibly have *different properties*. For example, brain activity consumes a certain amount of electrical power and gives it off as heat, and is – like any other physical activity – subject to the physical laws of thermodynamics. But mental states do not seem to have these features at all. Then, too, a mental state, such as your seeing a red apple or hearing a loud noise, does not seem to be paralleled by a corresponding feature in your brain. There are presumably no apple-shaped red patches in your brain, nor, presumably, are there any loud noises there either. Invoking the principle of the indiscernibility of identicals,[27] many philosophers were then prone to argue that it was not just false, but – more strongly – that it was *impossible*, that mental states could be brain states. But the entire debate was reversed at the end of the 1950s with the virtually simultaneous publication by three different philosophers – U.T. Place ([153]) in 1956, H. Feigl ([70][28]) in 1958, and J.J.C. Smart ([193]) in 1959 – of what has come to be known as the Identity theory of mental states and brain states. These philosophers argued that the standard, historically long-lived objections to the identifying of mental states with brain states were not valid.

Their rejection of the standard philosophical arguments against identifying mental states with brain states turned on two flaws they claimed to find in those arguments. The first was that even if a description of a mental state did not, in any obvious sense, demand a reference to properties appropriate to physical objects, that in and of

27. The principle of the indiscernibility of identicals states, recall, that whatever are numerically identical are qualitatively identical. A logically equivalent statement of this same principle reads: "Whatever are not qualitatively identical are not numerically identical." It is the latter version which is being invoked in the present context. If we wanted a name for this alternative formulation, we could call it the principle of "the nonidentity of discernibles".

28. The reprint (1967) of Feigl's original monograph contains a supplementary bibliography. Between the original and the supplement, the compilation comprises 565 items published on the mind-body problem.

itself did not mean that mental states were not states of physical things: the description may be incomplete, and incompleteness is no indicator of incompatibility. That we do not typically describe mental states using terminology drawn from neurophysiology, physics, and chemistry does not prove that those states cannot be so described. The second flaw these philosophers claimed to find in the standard argument dealt with the objection that when a person sees, for example, a red apple, there is no apple-shaped red patch in his brain. The way around this supposed difficulty was to counterargue that it was not the 'contents' or the 'objects' of consciousness which were to be found in the central nervous system, but the *state* or *activity* of consciousness. What was being identified was not a red patch in one's visual field with a publicly observable red patch in one's brain – by all accounts the latter patch does not exist – but it was the *activity* (or state) of seeing red which was being identified with some *activity* (or state) in the central nervous system.

In the ensuing years, the Identity theory has gone through a variety of refinements as philosophers have struggled to improve it and to make it more specific.[29] The precise details are best left to other books. Here my sole concern is with certain arguments which allege that the Identity theory – in any version whatsoever – can never aspire to the status of being a scientific theory.

Brain states are not physical objects, but are *states* of physical objects. Mental states, too, are not physical objects, but are states of something: of physical objects, if the Identity theory is true, or of something else (of minds or of persons), if the Identity theory is not true.

Is it possible to identify one property, P′, with a property P″? Is it possible to identify one state, S′, with a state S″? Two sorts of objections have been leveled against the proposal that the Identity theory might be testable by laboratory experiment. One of these is a general

29. The original proponents of the Identity theory claimed that the identity was *contingent*. Some subsequent theorists, also sympathetic to the Identity theory, have argued that if the relation is identity, then it cannot be contingent, but must instead be *necessary*. (See, for example, Kripke, [116].) This issue need not be pursued here. In this section, I am not concerned with the modal status of the Identity theory (i.e. whether it is contingent or necessarily true), but with the metaphysical presuppositions we bring to bear in determining the theory's epistemological status.

objection pertaining to the very possibility of establishing empirically the identity of any properties whatever. The other has to do with some alleged special peculiarities of the identification of mental states with brain states.

On the first score, some philosophers have argued that property- (and state-)identification is possible only when there is an equivalence in meaning between the expression used to refer to the one property (or state) and the expression used to refer to the other. For example, one might argue that the properties *azure* and *cerulean* are identical because the term "azure" means the same as the term "cerulean". But since "*a* is having a throbbing pain in his hand" and "*a*'s brain is in state *x*" do not *mean* the same thing (however "*x*" may be interpreted), it may be argued that the experience of having a pain can never be identified with any brain state whatever.

But against this theory is the hard evidence provided, within the history of science, of actual examples where property-identification has been successfully carried through. Scientists have, on occasion, identified one property, P′, with 'another', P″, without there being any meaning-equivalence between the terms used to refer to each. Scientists have, for example, identified electric current with the flow of electrons in a conductor; diabetes with elevated glucose levels in the bloodstream; radioactivity with the emission of subatomic particles; and visible light with electromagnetic radiation of wavelength 4000 to 7000 angstroms. In each case, at the time when the identification was made, there was no meaning-equivalence between the terms used. "Electric current" did not originally mean "flow of electrons". "Diabetes" did not originally mean "elevated glucose level". Indeed, diabetes was recognized as a distinct illness long before modern chemistry distinguished glucose from other sugars. It was an *empirical* discovery that diabetes is elevated glucose levels; it was no linguistic or semantic discovery.

We can see, then, that it is sometimes possible to identify properties with one another, even if the expressions we use to refer to them do not mean the same thing (or, more specifically, even though the expressions we use to refer to these properties are logically independent of one another). This latter is an important result, for without it, it would be impossible from the outset to subject the Identity theory to empirical test.

But having found that, on occasion, it is possible to identify some properties with one another even though the expressions referring to those properties are not equivalent, certainly does not by itself guarantee that we will be able to contrive an empirical test of the hypothesis that mental states are brain states. Indeed, at least one philosopher, Peter Herbst, argues that it is impossible in principle, because of the peculiar nature of mental states and of brain states themselves, ever to demonstrate empirically their identity with one another.

> Let us then investigate a proposition that there is a particular mental entity which is … identical with a particular brain state. In order to be able to test it, we must know which mental entity is supposed to be identical with what brain state. Therefore we need at least two clear and independent identifying references to serve as the basis of our proposition of identity. They must each be sufficient to individuate an entity, or else we cannot say what is identical with what, and they must be independent of each other, or else the identity proposition expressed in terms of them becomes tautologous. ([93], 57-8)

But having pointed out what he takes to be a logical requirement for putting the theory to empirical test, viz. having independent ways of individuating brain states and mental states, Herbst proceeds to express the gravest pessimism about our ever being able to carry out the task in practice.

> … it will not do to individuate experiences of having-a-sensation by their alleged neurophysiological properties. For exactly the same properties would also have to individuate the brain state, and therefore the two identifying references would fail of logical independence. Thus, for purposes of testing the empirical identity thesis [Identity theory], the ascription of neurophysiological properties to experiences is not only question-begging but useless.
>
> By what shall we individuate them? It is no use trying to individuate them by their spatio-temporal position alone, because, for one thing, we are not in a position to assign spatial positions to them unless the identity thesis is true, and it cannot be shown to be true unless sensation-experiences can be individuated.
>
> … individuation by neurophysiological properties is question-beggingly useless; individuation by spatial position likewise, and individuation by temporal position not useless but insufficient. ([93], 58-9)

The model Herbst is working with lies just beneath the surface. He is assuming that, in order to be able to demonstrate in a scientific experiment the identity of *A* with *B*, we must have some unique description of *A*, i.e. a way of individuating (singling out) *A* and we must have some *other* way, a logically independent way, of picking out *B*. For, as he points out, if the two descriptions are not logically independent of one another, if, that is, either one logically entails the other, then the identification of *A* with *B* is not a matter of scientific experiment at all but is simply a matter of logic (in Herbst's terminology, the identity proposition is 'tautological', i.e. true as a matter of the meanings of the terms involved). Applying these general requirements specifically to the Identity theory would imply that for that theory to be testable by empirical means we would have to have a way of individuating mental states and a way of individuating brain states which do not – as a matter of logic – entail that the one is the selfsame thing as the other. Then, having picked out, or isolated, the 'two' states, we must be able to demonstrate that the states (or properties) so individuated are really one and the same state (or property).

But he argues that we will, in principle, be unable to individuate mental states and brain states in a manner suitable for experimentally identifying the one with the other. This is so, he argues because (1) if we were to individuate mental states by their physical properties, i.e. by the neurophysiological features we use to individuate brain states, then we will have prejudged the very theory we are trying to prove; similarly, (2) if we were to individuate brain states by mentalistic properties (e.g. by their being painful, or clever, or incoherent), then, again, we will have prejudged the very theory we are trying to prove. But what about individuating brain states and mental states each by their spatiotemporal properties? Might we not be able to identify one with the other if we can show that they are both in the same place at the same time? Herbst argues (3) that we cannot assign precise spatial and temporal positions to mental states in advance of having accepted the Identity theory. Suppose you are now (having the experience of) recalling what you ate for breakfast today. Where, precisely, is this current memory experience? It is insufficient to offer a vague answer of the sort "somewhere in my head", since the identification at which we are ultimately aiming is with some very specific state located in particular nerves and lobes of the brain. In general, if we are depending on spatial positions as the basis for making an identification, it is unsound to identify *A* which is located 'in the vicinity of P_1' with *B* which is located 'exactly at P_1'. Herbst thus argues that an answer to the question about the location of experiences can be given with the requisite degree of specificity only if one has *already* accepted the Identity theory. But if one accepts the Identity theory, then one's testing procedure has become viciously circular, presupposing as a premise the very thing to be established as a conclusion.

We seem, then, to be faced with what looks to be an intractable dilemma: unless one assumes that the Identity theory is true, one cannot precisely assign spatiotemporal positions to experiences; but unless one can assign spatiotemporal positions to experiences, one cannot *test* the Identity theory. If this argument is accepted, it would appear, then, that the theory that there is an identity between mental states and brain states is in principle untestable.

Is there any way out of Herbst's stark dilemma? I think there is, but it requires that we back up and reject the naive methodology which Herbst has presupposed. His idea of the methodology involved in property-identification is the analog of the example we described of identifying one physical object with 'another': recall the case of your identifying my red car with my Oldsmobile. He seems to believe that to identify one property with another we must isolate instances of each, and then in examining them, we must be able to 'discover' their identity (in much the same way, for example, that we might 'discover' that two objects have the same length). But when we turn once again to the pages of the history of science, looking now to learn how scientists actually go about making property-identifications, we discover that their methodology is nothing remotely like what Herbst has envisaged.

In the later half of the nineteenth century, James Clerk Maxwell and Ludwig Boltzmann, in their celebrated dynamical (kinetic) theory of gases, were able to identify the temperature of a gas with the [total translational] kinetic energy of the molecules of that gas.[30] The term "temperature" certainly did not *mean* (at that time at least) "total kinetic energy". The identification plainly was not a matter of discerning any meaning-equivalence between the *terms* used. But neither did the identification take place in the manner presupposed by Herbst's model. These brilliant theoreticians did not proceed by first experimentally individuating the temperature of a gas and experimentally individuating the [translational] kinetic energy of its molecules, and

30. More exactly, with the mass of the gas multiplied by the root-mean-square speed of its constituent molecules

then discovering at a second stage – by some sort of unexplained observation – that these two properties of the gas instance the relation of *identity*. The identification of the temperature of a gas with the kinetic energy of its molecules came about, rather, on the basis of a certain highly controversial *theory* about the nature of gases, and on a great number of assumptions about the behavior of gas molecules, e.g. about their relative sizes, their numbers, their interactions, and the nature of their activity on striking the walls of a container. The subsequent empirical confirmation of their bold hypothesis of identity rested upon the predictions their theory made of observable macroscopic phenomena such as rates of diffusion, measures of specific heats, and viscosity. The empirical confirmation of the identity relation has never – not even now, more than one hundred years later – rested on an independent measure of the kinetic energy of the molecules. And yet the identification is taken to be so well established as to be no longer a matter of debate.

There are, as well, many other counterexamples to Herbst's flawed methodology.

One of the great mysteries in classical Newtonian physics was the unexplained proportionality between so-called inertial mass and gravitational mass. Physical objects attract one another with a force proportional to the product of their 'masses'. Physical objects also are accelerated by forces in proportion to their 'masses'. These are two – seemingly – quite distinct properties, and physicists distinguished *two* concepts of mass: a physical object's *gravitational mass*, and a physical object's *inertial mass*. Newton, himself, noted that these two 'masses' are apparently proportional to one another, but offered no explanation for it (see [144], book I, defs. I-III; book III, prop. VI). In the nineteenth century, the Hungarian physicist Lóránd von Eötvös[31] constructed an apparatus to measure how closely gravitational mass correlated to inertial mass. His torsion balance (c. 1890) was accurate to one part in 100 million (i.e. 10^{-8}). Within the limits possible with his device, he found no discrepancy whatever between the two measurements ([68]). Later measurements (1971) extended the accuracy to one part in a million million (i.e. 10^{-12}) ([195], 534). But clearly no such measurements, however refined, are capable of showing anything

31. Pronounced *ût′-vûsh*

more than a correlation. It is impossible for measurement, even if perfect, to demonstrate the identity of the properties being measured.

Since the late-seventeenth century, experimental physicists had been able to individuate the inertial mass and the gravitational mass of objects in logically independent ways and with increasing precision. In so doing they had satisfied the first and, on the face of it, the only problematic part of Herbst's methodology. Were Herbst's methodology sound, it would seem, then, that physicists ought to have been able, in a relatively effortless manner, to go on to determine whether they were observing one property or two. But having individuated inertial mass and gravitational mass, from as early as 1687 through to the early-twentieth century, physicists were up against a brick wall. There was no observation possible, or any direct test conceivable, which could answer for them whether they were observing one property of physical matter or two distinct, but highly correlated, properties. The question defied answer by any appeal to direct observation. The eventual identification of inertial mass with gravitational mass, when it was finally made by Einstein in 1916, did not occur as a result of his making finer measurements or by observing some telltale feature overlooked by other physicists.

Einstein's posit that inertial mass was not just highly correlated with gravitational mass, but was, in fact, one and the same property, proceeded by his identifying the so-called gravitational field as being itself an inertial field. (See [63] and [64], esp. chaps XIX and XX.) *Fields* are not observable entities. They are what are often called 'theoretical' or 'hypothetical' entities. This is not to say that they are unreal or fictitious, although some philosophers and physicists have been wont to so regard them. To say that an entity is 'theoretical' is to say that it cannot be observed 'directly'; that its existence is posited and confirmed by its explanatory role in a scientific theory. Gravitational fields are posited to explain the mutual attraction of physical bodies (just as electric fields are posited to explain the attraction and repulsion of charged bodies) (see [64], 144-8). Inertial fields are posited to explain the acceleration of bodies subjected to forces. By identifying gravitational fields with inertial fields, it followed as an immediate consequence that the gravitational mass of a body would be the selfsame as its inertial mass.

The identification of gravitational mass with inertial mass did not, then, come about in the sort of naive manner imagined by Herbst. One did not individuate inertial mass, individuate gravitational mass, and then 'experimentally discover' or 'directly observe' that the relation of identity held between the two properties. The identification came about because at a higher theoretical level, far removed from the observational base, Einstein posited the identity of gravitational fields with inertial fields.

Why does making such a posit count as doing empirical science? Because *predictions* which can be derived from the theory in which such a posit occurs are testable. In identifying gravitational fields with inertial fields, Einstein's theory – unlike Newton's – entailed that the axis of the orbit of the planet Mercury would gradually rotate (precess) ([63], 163-4), that the light emitted from a massive star would experience a so-called red-shift, and that light rays would bend in a gravitational field. These dramatic, and unexpected, predictions fit observed astronomical data better than did Newton's theory. And thus Einstein's theory won confirmation at the expense of Newton's theory, and Einstein's posit of the identity of gravitational fields with inertial fields was taken to be *indirectly* confirmed.

The empirical route to the confirmation of identity – of temperature with total kinetic energy, of gravitational mass with inertial mass – is no simple or direct matter. It is certainly not a matter to be settled by 'direct observation'. Identity is often established, not by direct observation, but indirectly, through layer upon layer of theory and of assumption.

Such historical examples shed light on the possibility of confirming the Identity theory because they tell us that if an identification of mental states with brain states ever should be made empirically, it will not come about by individuating instances of states of the one sort, by individuating instances of the other, and then by discovering that there is a relation of identity between them. The identification, if it ever is made, will be far more roundabout and far more protracted. And it will involve not only observation and experiment, but also vast amounts of theory.

The identification of mental states and brain states is, at present, a working hypothesis. It provides the motivation for a far-reaching research program, but its fine details are decades or more away from being stated. It will take generations to spell out precisely which brain states are supposed to be identical with which mental states. There will never be any particular laboratory findings, there will never be a crucial experiment, to which future historians will be able to point and say, "That experiment finally tipped the scales and showed once and for all that mental states are brains states". The transition, if it occurs, can only come about through a gradual and steady accumulation of vast storehouses of data along with theorizing of ever greater sophistication and refinement.

Sophisticated scientific theories seldom, if ever, appear suddenly, fully articulated. The Identity theory of mental states with brain states, if it is to prove successful, will doubtless follow the historical course we have earlier seen of the kinetic theory of heat: a succession of scientists will adopt it as a working hypothesis, and these generations of researchers will steadily improve the theory, filling in details, expanding its compass, piece by piece, over a course of time. In the case of heat theory, the transition to the kinetic theory took two hundred years.

Much of the future impetus for the Identity theory will come as much from research in Artificial Intelligence as it has, historically, derived from neurophysiology and experimental psychology. As engineers and theoreticians working together create electronic (and perhaps chemical and atomic) devices to emulate the cognitive processes of human beings, there will be a steadily increasing incentive to regard our own actual mental states as being nothing other than brain states. Present-day computers are not conscious; they do not think. But are computers capable *in principle* of consciousness? Of course it all depends on what one means by ‘a computer’ and what one means by ‘consciousness’. If by ‘a computer’ one means an electronic device operating in a linear manner (one step after another), i.e. as a so-called Von Neumann device ([4], 32-3), then there seems to be accumulating evidence, both experimental and theoretical, that such a computer will never be able to mimic the conscious and cognitive processes of an adult human being. But if we mean by ‘a computer’ nothing more than ‘a manufactured device containing no organic materials’, then it is very much an open question whether such a device could replicate the mental processes of human beings. To date, there do not seem to be any compelling reasons to believe that it is impossible in principle. We are beginning to be able to build computers which, in a rudimentary way, imitate some of the cognitive processes of human beings: in pattern recognition; in sensorimotor skills; in ability to play chess; in language translation; etc.

If, over the next several decades (centuries perhaps), computers can be built which imitate still better the abilities of human beings – e.g. our ability to understand a spoken language, our ability to communicate, our ability to reason, our ability to recognize features of our environment, our ability to learn, our ability to generalize – there will be less and less reason to refrain from attributing consciousness to them.[32] And once we have begun to attribute consciousness to our computers, we will very likely find that we regard our own mental processes as being nothing other than the processes of our own built-in computers, i.e. of our brains and their associated peripheral nervous systems. The Identity theory will have come to be accepted with the kind of natural inevitability at present enjoyed by the theory that the temperature of a gas is the [total translational] kinetic energy of that gas and by the theory that inertial mass is gravitational mass.

There is a loose end in this scenario, however. To isolate it, we will back up one last time, virtually to the conceptual roots, at it were, of the debate about the Identity theory. "Why", it is profitable to ask, "in spite of the fact that so many philosophers and scientists remain undecided about the truth of the Identity theory, have so many others uncritically accepted it?" Why do so many nonphilosophers and non-scientists even now believe the Identity theory to be true? What evidence promotes that belief?

The evidence cited is well-known. Many persons believe that the Identity theory is true because they are familiar with laboratory findings that certain electrical stimulations evoke vivid memories; that other electrical stimulations cause or assuage pain; that certain chemicals introduced in the bloodstream will cause vivid hallucinations, euphoria, panic, sleep, etc. In short, most of us already know that a great deal of what goes on in our consciousness, perhaps all of it, is profoundly intimately related to what is happening in our brains. And

---

32. See Turing's "Computing Machinery and Intelligence" ([206]) and Scriven's "The Compleat Robot: A Prolegomena to Androidology" ([187]). Turing's suggestion that we should want to attribute consciousness to a machine which could 'imitate' a human being in a question-and-answer game (the 'imitation game') has been subjected to severe criticism, probably the most well-known being that by John Searle in his "Minds, Brains, and Programs" ([188]). Searle's critique, in its turn, provoked a firestorm of counterargument (see, e.g., Hofstadter's reply, along with a reprint of Searle's original paper, in [99], 373-82). To a certain degree, Searle's criticisms are becoming moot, insofar as they were directed principally against a particular model – the so-called computational model – of Artificial Intelligence instantiated in a Von Neumann machine. But both of these targets are gradually giving way to newer techniques and bolder architectures which one may, reasonably, regard as approaching more and more the structure and functioning of the human brain.

thus many persons have assumed that our mental states just simply *are* brain states.

But of course the latter conclusion is not strictly warranted by the empirical data which are cited in its support. Descartes, too, believed that most if not all of what goes on in our consciousness is intimately related to what goes on in our brains, but he, unlike many persons today, did not believe that mental states were brain states. Indeed, he believed that it was impossible that they could be. He was, as we have said (pp. 93ff.), a dualist, believing that there was causal interaction between the mental states and brain states, but that they could not be identified one with another.

The very possibility of maintaining dualism, as did Descartes, in the face of an exceptionless correlation between mental states and brain states tells us that something more must be added to the account of the methodology of identity. Given that one can always, in principle, be a dualist about any alleged identification – one could, for example, be a dualist about temperature and kinetic energy, about inertial and gravitational mass, etc. – we must explain why identity is sometimes the preferable hypothesis. In short, the hypothesis of identity always competes against the hypothesis of 'mere correlation'. Any two quantitative properties (or states) which are identified must be correlated. Why should we ever want to pass beyond positing mere correlation to posit a stronger relationship, viz. identity? What empirical evidence could we ever have for warranting a hypothesis of identity over and above one of mere correlation?

Suppose someone were to be a dualist, not with regard to mental states and brain states, but with regard to inertial mass and gravitational mass, arguing that all that has ever been experimentally demonstrated (e.g. in Eötvös's experiment and its successors) is a remarkable correlation between the two properties and that there has never been, nor could there be, an experimental demonstration that showed anything more, i.e. that it is impossible to demonstrate that there is an actual identity between the two. How could such a challenge be met?

It is at this point that two powerful metaphysical principles involved in the identification of properties with one another must come into play. In the end, the choice of identity over mere correlation is made in part on the basis of the desire for ontological or metaphysical economy. We want not to multiply entities beyond necessity. Mere correlation posits two ontologically distinct entities; identity posits but one, and is thus preferable. But the latter choice is also mandated by our desire that our theories be explanatorily powerful. In positing 'mere correlation', an intractable problem remains: "Why are the two states correlated? What is the nature of the connection between the two?" Descartes, himself, was crucially aware of this latter difficulty and struggled to offer a cogent answer, but neither he, nor any other dualist, ever could satisfactorily fill that gaping hole in the theory. In a theory which posits identity, in contrast, there is nothing further to be explained. If mental states *are* brain states, then there is nothing to explain in their being correlated. Correlation follows immediately, as a logical matter, from identification.

By transcending the empirical data of correlation, in particular by positing an identity, we satisfy at one stroke two intensely powerful metaphysical desiderata: we effect an economy in our ontology and we avoid the need for further explanation carrying, as it might, the requirement of positing still further kinds of entities and hidden interactions. In short, when it comes down to positing an identity or a 'mere correlation', if there is not good reason to desist from making the identification, identity – not 'mere correlation' – is the preferred hypothesis. The naive model, which portrays correlation as being demonstrable and identity as being a relation which takes 'something more', over and above 'mere correlation', to warrant its being posited, has turned the metaphysical requirements upside down. Identity is the preferred hypothesis; 'mere correlation' a decided second-best, to be invoked only when there are grounds to believe that identity does not obtain.

A scientific version of the Identity theory of brain states and mental states has not yet been well enough confirmed to warrant its acceptance to the degree, for example, that we accept certain identifications made in kinetic theory and in general relativity. The Identity theory is at present a remarkably fruitful research program. But it stands to a fully articulated theory much as Bacon's theory, "Heat is Motion", stood to the theory of Maxwell and Boltzmann (two centuries later) which explained specifically temperature, pressure, viscosity, entropy, free energy, etc., in terms of the physical properties of the microstructure of a gas. (Incidentally, Maxwell's and Boltzmann's theory was hardly the last word. Kinetic theory has continued to evolve, through the Dirac and Einstein-Bose repairs, and will continue to evolve for the foreseeable future.)

The Identity theory is in principle empirically testable. Testability is, however, a matter of degree. The Identity theory, which is today only minimally testable, may well grow steadily more testable as it becomes, slowly, over many decades, better articulated. Testability never resides in a single, one-shot, laboratory experiment. It is a much more global affair, encompassing observation, creative imagination, theory construction, constant revision, and a metaphysical model of what constitutes ontological economy and explanatory power. The testing of the Identity theory will involve hundreds of researchers painstakingly assembling countless tens of thousands of pieces of data into a comprehensive whole. It will no more be tested by a single experiment, or a few, than was Newtonian mechanics or kinetic theory. And like these latter theories, it will have a penumbra of metaphysical assumptions, assumptions which are as vital to the theory as any of the most directly testable predictions of that theory.