

## CHAPTER 23

- Reading 23.1 **Posner, M.I. (1993).** *Seeing the mind. Science, 262: 673–674.* (Extract p673)
- Reading 23.2 **McDowell, J. (1985).** *Functionalism and Anomalous Monism. In Actions and Events* (ed. E. Lepore and B. McLaughlin). Oxford: Blackwell pp. 387–398. Reprinted in McDowell, J. (1998) *Mind, Value, and Reality*. Cambridge, MA.: Harvard University Press pp. 325–340 (Extracts pp. 328–331, 336–7)
- Reading 23.3 **Davidson, D. (1980).** *Mental events. In Essays on Actions and Events*. Oxford: Oxford University Press, pp. 207–227. Reprinted in D. Rosenthal (ed.) (1991) *The Nature of Mind*. Oxford: Oxford University Press, pp. 247–256 (Extracts pp. 247–8, 248–9, 250).
- Reading 23.4 **Kripke, S. (1980).** *Naming and Necessity*. Oxford: Blackwell. Extract reprinted in D. Rosenthal (ed.) (1991). *The Nature of Mind*. Oxford: Oxford University Press, pp. 247–256 (Extract pp. 242–3)

**Reading 23.1****EXERCISE 1**

From: Posner, M.I. (1993). Seeing the mind. *Science*, 262: 673–674. (Extract p673)

The microscope and telescope opened vast domains of unexpected scientific discovery. Now that new imaging methods can visualize the brain systems used for normal and pathological thought, a similar opportunity may be available for human cognition. Some of the data generated with these imaging techniques fits with current ideas, but much of the new information will require new theories. Here, I review three areas of cognitive psychology in which our ideas are changing as a consequence of these new results: the localization of mental operations, the identification of separate brain control systems (attentional networks), and the convergence of sensory input and mental imagery in the same brain areas. New theories based on these findings are likely to change our interpretation of the meaning of what has already been seen and to suggest new experiments.

The major new anatomical tools all image aspects of blood vessel function that reflect nearby neuronal activity. The two most successful so far are positron emission tomography (PET) and magnetic resonance imaging (MRI). Although the fundamental links between blood flow and neuronal activity remain obscure, it is clear that the measurement of cerebral blood flow can tell us where neurons are more or less active in comparison to a control condition. Of course, an increase in neuronal activity may signal either inhibitory or excitatory synaptic events, so

these measures must also be related to the efficiency of task performance. In addition, blood flow lags behind neural activity in time, requiring the use of other methods with time resolution in the millisecond range if changes in anatomy are to be traced dynamically during mental tasks. Methods based on noninvasive electrical and magnetic measurements will be important in studying these dynamic changes when they can be related to the emerging functional anatomy from PET and MRI studies.

It is a popularly held belief in psychology that the cognitive functions of the brain are widely distributed among different brain areas. Even though the organization of the nervous system suggests that sensory and motor functions are localized to specific brain regions, the failure of phrenology and difficulties in locating memory traces for higher functions have led to a strong reaction against the notion of localization of cognitive processes. Nevertheless, imaging studies reveal a startling degree of region-specific activity. The PET studies show clearly that such visual functions as processing color, motion, or even the visual form of words occur in particular prestriate areas. This localization extends beyond sensory systems. When thought is analyzed in terms of component mental operations, a beautiful localization emerges. In word-reading studies, words activate specific posterior visual areas that are not affected by consonant strings, and specific frontal and temporo-parietal areas are active when subjects are required to indicate the use of a noun (for example, hammer-pound) or its classification into a category (hammer-tool). The brain activation accompanying this form of semantic processing is illustrated in the figure, in which the subjects were asked to determine the use of visually presented nouns.

## Reading 23.2

### EXERCISE 2

From: McDowell, J. (1985). *Functionalism and Anomalous Monism*. In *Actions and Events* (ed. E. Lepore and B. McLaughlin). Oxford: Blackwell, pp. 387–398. Reprinted in McDowell, J. (1998) *Mind, Value, and Reality*. Cambridge, MA.: Harvard University Press pp. 325–340 (Extracts pp. 328–331, 336–7)

### Extract 1

3. I remarked that any particular instantiation of deductive rationality will embody a more or less imperfect grasp of what, in general, follows from what. This variable gap between actual and ideal may make it seem that any constitutive force that can be attributed to this particularization of the concept of rationality could not extend to the structure of deductive reason itself, as I have said that Davidson's argument requires. The idea would be that constitutive force could be ascribed only to some minimally necessary structure, exemplified in the actual psychological economy of anything that could be recognized as a rational mind: exactly the sort of thing that Loar's "L-constraints" aspire to capture. Some such idea must underlie Loar's misinterpretation of what Davidson claims. I believe it betrays a prejudice about the character of the understanding we can achieve by employing the conceptual apparatus that is governed by the constitutive force of rationality.

Davidson's claim is, in effect, that if someone offered to reflect the patterns required by rationality in a structure described in non-intentional terms, then, in view of the fact that the constitutive concept functions as an ideal or norm, he would be committing a kind of "naturalistic fallacy". (The label is suggestive, but unfortunate in implying that there is something non-naturalistic about propositional attitudes conceived as irreducible to anything non-intentional; I shall return to this.) The prejudice I have in mind would preclude giving this thought its proper significance, by inducing a refusal to recognize that it is something with the status of an ideal that is being credited with a constitutive role in governing our thinking about propositional attitudes. To recognize the ideal status of the constitutive concept is to appreciate that the concepts of the propositional attitudes have their proper home in explanations of a special sort: explanations in which things are made intelligible by being revealed to be, or to approximate to being, as they rationally ought to be. This is to be contrasted with a style of explanation in which one makes things intelligible by representing their coming into being as a particular instance of how things generally tend to happen. (In the usual way of formulating the philosophical issue we are concerned with, "the physical" need do no more than point to the subject-matter of those sciences that aim at explanations of the second sort.)<sup>1</sup>

<sup>1</sup> This is why I have put "physical" and its cognates in quotation marks. The issue is a live one quite independently of whether "physics" is reducible to, or even interestingly dependent on, physics.

Loar's weakening of the concept to which, in purported agreement with Davidson, he attributes the constitutive force that shapes our understanding of propositional attitudes, from Davidson's ideal to his own highest common factor of the actual, reflects a determination to assimilate all explanation to the second of these two sorts.

This has a damaging effect, which I can illustrate without lifting the restriction to deductive rationality. With the restriction in force, what is in question is a mode of understanding in which one finds a belief intelligible on the basis of its following deductively (or being intelligibly but falsely thought to follow deductively) from other beliefs that one knows the believer holds. Attaining this kind of understanding requires bringing to bear the notion of deductive consequence, and it must be that notion itself, not some thinned-down surrogate; if we allow ourselves the idea that the relevant explanations work by locating explanandum and explanans within a structure, it must be the ideal structure of deductive reason, not the less demanding sort of structure that could be determined by something on the lines of Loar's "L-constraints". Now Loar envisages a theory that would flesh out the admittedly thin structure that his "*a priori* rationality constraints" would impose by adding further functional relations; these further relations, to be established by theoretical inquiry rather than excogitated *a priori*, are envisaged as belonging to "the part of the theory outside common sense", and hence as not being required to "correspond to cognitive, intentional, or conceptual relations" (p. 79). So there is no requirement that anything in the theory outside the "L-constraints" should even aim to mirror the structures of deductive inferences, let alone amount to expressing the concept of deductive consequence. The result is that outside the sphere of beliefs, or absences of beliefs, that are related as the "L-constraints" expressly stipulate, a theory such as Loar envisages would not even aspire to deliver a kind of understanding of beliefs, or absences of beliefs, that depends on the thought that other beliefs are deductively cogent reasons for them. And even within that sphere, the undemandingness of the "L-constraints"—their innocence of any ambition to capture the ideal structure of deductive reason—means that any understanding that such a theory could offer of beliefs, or absences of beliefs, on the basis of beliefs that are as a matter of fact deductively related to them would not be the kind of understanding I have described. Such a theory would not have the general normative notion of deductive consequence at its disposal; so its explanations could not exploit that notion, but could draw at most on the idea of certain transitions, and refrainings from transitions, that minds are as a matter of fact prone to. By Davidsonian lights, even that formulation is unwarranted; since the idea of rationality is not credited with its constitutive role, there is nothing to ensure that it is minds that the theory is about.

In connection with his "*a priori* rationality constraints", Loar writes at one point of the way in which a theory that embodies them says that beliefs "ought rationally to be related" (p. 221). That "ought" may seem to make room for the distinctive kind of

## 4 CHAPTER 23 READING 23.2

understanding that I am claiming Loar's position cannot countenance, but the appearance is misleading. It is helpful to consider the question *why* beliefs ought rationally to be related as the "L-constraints" stipulate. In real life we need have no difficulty in answering this question. But a theory of the sort Loar envisages would contain no materials for addressing it. Such a theory would owe no allegiance to "common sense" apart from the "*a priori* rationality constraints" themselves; and the explanations it yielded would purport to be self-sufficient—there is no room for the suggestion that their explanatory power might be enhanced if we explained the demandingness of the "L-constraints" in terms of the general notion of deductive cogency. Made within such a theory, the claim that beliefs ought rationally to be related in accordance with the "L-constraints" would not reflect the conformity of beliefs so related, in particular, to a categorical norm, intelligibly operative in other cases also; obedience to the "L-constraints" would have, rather, the status of a hypothetical imperative—something without which a system of states could not be recognized as characterizing a rational mind, and so as a system of beliefs, at all. Now it is no doubt correct to attribute that status to Loar's "L-constraints". But if we grant an explanatory role to an ideal that transcends them, we can explain why they have that status in terms of the thought that violations would lie outside the boundaries of what is intelligible—a terrain of whose topography we have a pre-theoretical ("common-sense") grasp that outruns anything captured by Loar's "*a priori* rationality constraints". In Loar's picture, by contrast, "common sense" is conceived as doing no more than presenting us with the "rationality constraints", all else being the province of a theory-construction that, having embraced them, lies under no further obligation to respect thoughts expressible in intentional (content-involving) terms. So Loar's "ought rationally" does not reflect an acknowledgement of the distinctive kind of understanding I have described; this shows up in its being, from the standpoint of a theory of the sort Loar envisages, a brute fact—a sheer inexplicable datum of "common sense"—that the "*a priori* rationality constraints" mark limits of intelligibility.

When we come to consider applications for the concept of rationality outside the sphere of deductive consequence, it seems no less plain—though proof is no longer in question—that it would be a fantasy to suppose that the full normative force of the concept, in its extra-logical applications, could be captured in a structure specifiable from outside intentional content. Not a fantasy that Loar indulges in: my point is that his response to

Davidson reflects, rather, an inability to see that it is the full normative force of the concept to which Davidson attributes a constitutive status. (pp. 328–331)

### Extract 2

It is often thought that if essential subjectivity poses any threat to an objectivistic conception of the mental, it is a threat restricted to phenomenal or qualitative mental states or events: states or events about which there is an answer to the question what it is distinctively like to be in them or undergo them.<sup>2</sup> But the Davidsonian irreducibility of propositional attitudes—which are not states whose essence resides in their qualitative character—traces back to subjectivity too: not only in that the mode of understanding they subserve is a matter of comprehending the specific content of a particular subject's outlook on the world, but also—a thought involving a special status for a less individually conceived point of view—in that we cannot find any use for a distinction between what makes sense and what could come to make sense to us, if necessary as a result of our learning from those whom we thereby come to find intelligible.<sup>3</sup> Loar's position cannot incorporate either of these considerations, in both cases for reasons we have already touched on: it misses the first because it holds that the explanatory capacity of a propositional attitude is in general independent of its bearing on the world, and it misses the second because, in so far as it deals at all with a distinction between what makes sense and what does not, it is committed to treating the limits of intelligibility as a brute and presumably objective datum.

These considerations bring out a connection between the ideal-involving kind of explanation and the irreducible subjectivity of propositional attitudes. Achieving the kind of understanding for which rationality plays its constitutive role requires a sensitivity to the specific detail of the subjective stance of others, and an openness to learning from it, that is bound to be falsified if one supposes that explanations involving the constitutive ideal work by locating their explananda in a structure specifiable from outside content. Loar conceives the constitutive ideal as just such a structure—misinterpreting, rather than answering, Davidson's claim about the constitutive concept, as I have insisted. His position seems to be the nearest one could come to accommodating Davidson's point within a picture of mind as an element in objective reality.

<sup>2</sup> See Nagel, "What Is It Like to Be a Bat?". A substantial amount of work has been based on the idea that "qualia" pose a *special* problem for functionalism; see, e.g., Sydney Shoemaker, "Functionalism and Qualia".

<sup>3</sup> See Davidson, "On the Very Idea of a Conceptual Scheme". Nagel seems to me to miss the depth of his own point about subjectivity, not only in suggesting that the point is restricted to the qualitative aspects of the mental, but also in espousing a realism that countenances the idea of subjective facts completely beyond our reach.

## Reading 23.3

### EXERCISE 3

From: Davidson, D. (1980). Mental events. In *Essays on Actions and Events*. Oxford: Oxford University Press, pp. 207–227 (Reprinted in D. Rosenthal (ed.) (1991) *The Nature of Mind*. Oxford: Oxford University Press, pp. 247–256 (Extracts pp. 247–8, 248–9, 250).

### Extract 1

Mental events such as perceivings, rememberings, decisions, and actions resist capture in the nomological net of physical theory. How can this fact be reconciled with the causal role of mental events in the physical world? Reconciling freedom with causal determinism is a special case of the problem if we suppose that causal determinism entails capture in, and freedom requires escape from, the nomological net. But the broader issue can remain alive even for someone who believes a correct analysis of free action reveals no conflict with determinism. *Autonomy* (freedom, self-rule) may or may not clash with determinism; *anomaly* (failure to fall under a law) is, it would seem, another matter.

I start from the assumption that both the causal dependence, and the anomalousness, of mental events are undeniable facts. My aim is therefore to explain, in the face of apparent difficulties, how this can be. I am in sympathy with Kant when he says.

it is as impossible for the subtlest philosophy as for the commonest reasoning to argue freedom away. Philosophy must therefore assume that no true contradiction will be found between freedom and natural necessity in the same human actions, for it cannot give up the idea of nature any more than that of freedom. Hence even if we should never be able to conceive how freedom is possible, at least this apparent contradiction must be convincingly eradicated. For if the thought of freedom contradicts itself or nature . . . it would have to be surrendered in competition with natural necessity.

Generalize human actions to mental events, substitute anomaly for freedom, and this is a description of my problem. And of course the connection is closer, since Kant believed freedom entails anomaly.

Now let me try to formulate a little more carefully the “apparent contradiction” about mental events that I want to discuss and finally dissipate. It may be seen as stemming from three principles.

The first principle asserts that at least some mental events interact causally with physical events. (We could call this the Principle of Causal Interaction.) Thus for example if someone sank the *Bismarck*, then various mental events such as perceivings, notings, calculations, judgments, decisions, intentional actions and changes of belief played a causal role in the sinking of the *Bismarck*. In particular, I would urge that the fact that someone sank the *Bismarck* entails that he moved his body in a way that was caused by mental events of certain sorts, and that this bodily movement in turn caused the *Bismarck* to sink. Perception

illustrates how causality may run from the physical to the mental: if a man perceives that a ship is approaching, then a ship approaching must have caused him to come to believe that a ship is approaching. (Nothing depends on accepting these as examples of causal interaction.)

Though perception and action provide the most obvious cases where mental and physical events interact causally, I think reasons could be given for the view that all mental events ultimately, perhaps through causal relations with other mental events, have causal intercourse with physical events. But if there are mental events that have no physical events as causes or effects, the argument will not touch them.

The second principle is that where there is causality, there must be a law: events related as cause and effect fall under strict deterministic laws. (We may term this the Principle of the Nomological Character of Causality.) This principle, like the first, will be treated here as an assumption, though I shall say something by way of interpretation.

The third principle is that there are no strict deterministic laws on the basis of which mental events can be predicted and explained (the Anomalism of the Mental). (pp. 247–248)

### Extract 2

The three principles will be shown consistent with one another by describing a view of the mental and the physical that contains no inner contradiction and that entails the three principles. According to this view, mental events are identical with physical events. Events are taken to be unrepeatable, dated individuals such as the particular eruption of a volcano, the (first) birth or death of a person, the playing of the 1968 World Series, or the historic utterance of the words, “You may fire when ready, Gridley.” We can easily frame identity statements about individual events; examples (true or false) might be:

The death of Scott = the death of the author of *Waverley*;

The assassination of the Archduke Ferdinand = the event that started the First World War

The eruption of Vesuvius in A.D. 79 = the cause of the destruction of Pompeii.

The theory under discussion is silent about processes, states, and attributes if these differ from individual events.

What does it mean to say that an event is mental or physical? One natural answer is that an event is physical if it is describable in a purely physical vocabulary, mental if describable in mental terms. But if this is taken to suggest that an event is physical, say, if some physical predicate is true of it, then there is the following difficulty. Assume that the predicate ‘*x* took place at Noosa Heads’ belongs to the physical vocabulary; then so also must the predicate ‘*x* did not take place at Noosa Heads’ belong to the physical vocabulary. But the predicate ‘*x* did or did not take place at Noosa Heads’ is true of every event, whether mental or physical. We

might rule out predicates that are tautologically true of every event, but this will not help since every event is truly describable either by 'x took place at Noosa Heads' or by 'x did not take place at Noosa Heads.' A different approach is needed.

We may call those verbs mental that express propositional attitudes like believing, intending desiring, hoping, knowing, perceiving, noticing remembering, and so on. Such verbs are characterized by the fact that they sometimes feature in sentences with subjects that refer to persons, and are completed by embedded sentences in which the usual rules of substitution appear to break down. This criterion is not precise, since I do not want to include these verbs when they occur in contexts that are fully extensional ('He knows Paris,' 'He perceives the moon' may be cases), nor exclude them whenever they are not followed by embedded sentences. An alternative characterization of the desired class of mental verbs might be that they are psychological verbs as used when they create apparently nonextensional contexts.

Let us call a description of the form 'the event that is  $m$ ' or an open sentence of the form 'event  $x$  is  $m$ ' a *mental description* or a *mental open sentence* if and only if the expression that replaces ' $m$ ' contains at least one mental verb essentially. (Essentially, so as to rule out cases where the description or open sentence is logically equivalent to one not containing mental vocabulary.) Now we may say that an event is mental if and only if it has a mental description, or (the description operator not being primitive) if there is a mental open sentence true of that event alone. Physical events are those picked out by descriptions or open sentences that contain only the physical vocabulary essentially. It is less important to characterize a physical vocabulary because relative to the mental it is, so to speak, recessive in determining whether a description is mental or physical. (There will be some comments presently on the nature of a physical vocabulary, but these comments will fall far short of providing a criterion.)

On the proposed test of the mental, the distinguishing feature of the mental is not that it is private, subjective, or immaterial, but that it exhibits what Brentano called intentionality. Thus intentional actions are clearly included in the realm of the mental along with thoughts, hopes, and regrets (or the events tied to these). What may seem doubtful is whether the criterion will include events that have often been considered paradigmatic of the mental. Is it obvious, for example, that feeling a pain or seeing an afterimage will count as mental? Sentences that report such events seem free from taint of nonextensionality, and the same should be true of reports of raw feels, sense data, and other uninterpreted sensations, if there are any.

However, the criterion actually covers not only the havings of pains and afterimages, but much more besides. Take some event one would intuitively accept as physical, let's say the collision of two stars in distant space. There must be a purely physical predicate ' $px$ ' true of this collision, and of others, but true of only this one at the time it occurred. This particular time, though, may be pinpointed as the same time that Jones notices that a pencil starts to

roll across his desk. The distant stellar collision is thus *the event*  $x$  such that  $px$  and  $x$  is simultaneous with Jones' noticing that a pencil starts to roll across his desk. The collision has now been picked out by a mental description and must be counted as a mental event.

This strategy will probably work to show every event to be mental; we have obviously failed to capture the intuitive concept of the mental. It would be instructive to try to mend this trouble, but it is not necessary for present purposes. We can afford Spinozistic extravagance with the mental since accidental inclusions can only strengthen the hypothesis that all mental events are identical with physical events. What would matter would be failure to include bona fide mental events, but of this there seems to be no danger.

I want to describe, and presently to argue for, a version of the identity theory that denies that there can be strict laws connecting the mental and the physical. The very possibility of such a theory is easily obscured by the way in which identity theories are commonly defended and attacked. Charles Taylor, for example, agrees with protagonists of identity theories that the sole "ground" for accepting such theories is the supposition that correlations or laws can be established linking events described as mental with events described as physical. He says, "It is easy to see why this is so: unless a given mental event is invariably accompanied by a given, say, brain process, there is no ground for even mooted a general identity between the two." Taylor goes on (correctly, I think) to allow that there may be identity without correlating laws, but my present interest is in noticing the invitation to confusion in the statement just quoted. What can "a given mental event" mean here? Not a particular, dated, event, for it would not make sense to speak of an individual event being "invariably accompanied" by another. Taylor is evidently thinking of events of a given *kind*. But if the only identities are of kinds of events, the identity theory presupposes correlating laws.

One finds the same tendency to build laws into the statement of the identity theory in these typical remarks:

When I say that a sensation is a brain process or that lightning is an electrical discharge, I am using 'is' in the sense of strict identity . . . there are not two things: a flash of lightning and an electrical discharge. There is one thing, a flash of lightning, which is described scientifically as an electrical discharge to the earth from a cloud of ionized water molecules.

The last sentence of this quotation is perhaps to be understood as saying that for every lightning flash there exists an electrical discharge to the earth from a cloud of ionized water molecules with which it is identical. Here we have an honest ontology of individual events and can make literal sense of identity. We can also see how there could be identities without correlating laws. It is possible, however, to have an ontology of events with the conditions of individuation specified in such a way that any identity implies a correlating law. Kim, for example, suggests that  $fa$  and  $gb$  "describe or refer to the same event" if and only if  $a = b$  and the property of being  $F =$  the property of being  $G$ .

### Extract 3

Anomalous monism resembles materialism in its claim that all events are physical, but rejects the thesis, usually considered essential to materialism, that mental phenomena can be given purely physical explanations. Anomalous monism shows an ontological bias only in that it allows the possibility that not all events are mental, while insisting that all events are physical. Such a bland monism, unbuttressed by correlating laws or conceptual economies, does not seem to merit the term “reductionism”; in any case it is not apt to inspire the nothing-but reflex (“Conceiving the *Art of the Fugue* was nothing but a complex neural event,” and so forth.)

Although the position I describe denies there are psychophysical laws, it is consistent with the view that mental characteristics are in some sense dependent, or supervenient, on physical characteristics. Such supervenience might be taken to mean that there cannot be two events alike in all physical respects but differing in some mental respect, or that an object cannot alter in some mental respect without altering in some physical respect. Dependence or supervenience of this kind does not entail reducibility through law or definition: if it did, we could reduce moral properties to descriptive, and this there is good reason to *believe* cannot be done and we might be able to reduce truth in a formal system to syntactical properties, and this we *know* cannot in general be done.

This last example is in useful analogy with the sort of lawless monism under consideration. Think of the physical vocabulary as the entire vocabulary of some language  $L$  with resources

adequate to express a certain amount of mathematics, and its own syntax.  $L'$  is  $L$  augmented with the truth predicate ‘true-in- $L$ ,’ which is “mental.” In  $L$  (and hence  $L'$ ) it is possible to pick out, with a definite description or open sentence, each sentence in the extension of the truth predicate, but if  $L$  is consistent there exists no predicate of syntax (of the “physical” vocabulary), no matter how complex that applies to all and only the true sentences of  $L$ . There can be no “psychophysical law” in the form of a biconditional, ‘ $(x)$  ( $x$  is true-in- $L$  if and only if  $x$  is  $\phi$ )’ where ‘ $\phi$ ’ is replaced by a “physical” predicate (a predicate of  $L$ ). Similarly, we can pick out each mental event using the physical vocabulary alone, but no purely physical predicate, no matter how complex, has, as a matter of law, the same extension as a mental predicate.

It should now be evident how anomalous monism reconciles the three original principles. Causality and identity are relations between individual events no matter how described. But laws are linguistic; and so events can instantiate laws, and hence be explained or predicted in the light of laws, only as those events are described in one or another way. The principle of causal interaction deals with events in extension and is therefore blind to the mental-physical dichotomy. The principle of the anomalism of the mental concerns events described as mental, for events are mental only as described. The principle of the nomological character of causality must be read carefully it says that when events are related as cause and effect, they have descriptions that instantiate a law. It does not say that every true singular statement of causality instantiates a law.

## Reading 23.4

## EXERCISE 4

From: Kripke, S. (1980). *Naming and Necessity*. Oxford: Blackwell. Extract reprinted in D. Rosenthal (ed.) (1991). *The Nature of Mind*. Oxford: Oxford University Press, pp. 247–256 (Extract pp. 242–3)

I have not given any general paradigm for the appropriate corresponding qualitative contingent statement. Since we are concerned with how things might have turned out otherwise, our general paradigm is to redescribe both the prior evidence and the statement qualitatively and claim that they are only contingently related. In the case of identities, using two rigid designators, such as the Hesperus–Phosphorus case above, there is a simpler paradigm which is often usable to at least approximately the same effect. Let ‘ $R_1$ ’ and ‘ $R_2$ ’ be the two rigid designators which flank the identity sign. Then ‘ $R_1 = R_2$ ’ is necessary if true. The references of ‘ $R_1$ ’ and ‘ $R_2$ ’, respectively, may well be fixed by nonrigid designators ‘ $D_1$ ’ and ‘ $D_2$ ’, in the Hesperus and Phosphorus cases these have the form ‘the heavenly body in such-and-such position in the sky in the evening (morning)’. Then although ‘ $R_1 = R_2$ ’ is necessary, ‘ $D_1 = D_2$ ’ may well be contingent, and this is often what leads to the erroneous view that ‘ $R_1 = R_2$ ’ might have turned out otherwise.

I finally turn to an all too cursory discussion of the application of the foregoing considerations to the identity thesis. Identity theorists have been concerned with several distinct types of identifications: of a person with his body, of a particular sensation (or event or state of having the sensation) with a particular brain state (Jones’s pain at 06:00 was his C-fiber stimulation at that time), and of *types* of mental states with the corresponding *types* of physical states (pain is the stimulation of C-fibers). Each of these, and other types of identifications in the literature, present analytical problem rightly raised by Cartesian critics, which cannot be avoided by a simple appeal to an alleged confusion of synonymy with identity. I should mention that there is of course no obvious bar, at least (I say cautiously) none which should occur to any intelligent thinker on a first reflection just before bed-time, to advocacy of some identity these while doubting or denying others. For example, some philosophers have accepted the identity of particular sensations with particular brain states while denying the possibility of identities between mental and physical *types*.<sup>1</sup> I will concern myself primarily with the type–type identities, and the philosophers in question will thus be immune to much of the discussion; but I will mention the other kinds of identities briefly.

<sup>1</sup> Thomas Nagel and Donald Davidson are notable examples. Their views are very interesting, and I wish I could discuss them in further detail. It is doubtful that such philosophers wish to call themselves ‘materialists’. Davidson, in particular, bases his case for his version of the identity theory on the supposed *impossibility* of correlating psychological properties with physical ones.

The argument against token–token identification and the text *does* apply to these views.

Descartes, and others following him, argued that a person or mind is distinct from his body, since the mind could exist without the body. He might equally well have argued the same conclusion from the premise that the body could have existed without the mind.<sup>2</sup> Now the one response which I regard as plainly inadmissible is the response which cheerfully accepts the Cartesian premise while denying the Cartesian conclusion. Let ‘Descartes’ be a name, or rigid designator, of a certain person, and let ‘ $B$ ’ be a rigid designator of his body. Then if Descartes were indeed identical to  $B$ , the supposed identity, being an identity between two rigid designators, would be necessary, and Descartes could not exist without  $B$  and  $B$  could not exist without Descartes. The case is not at all comparable to the alleged analogue, the identity of the first Postmaster General with the inventor of bifocals. True, this identity obtains despite the fact that there could have been a first Postmaster General even though bifocals had never been invented. The reason is that ‘the inventor of bifocals is not a rigid designator; a world in which no one invented bifocals is not *ipso facto* a world in which Franklin did not exist. The alleged analogy therefore collapses; a philosopher who wishes to refute the Cartesian conclusion must refute the Cartesian premise, and the latter task is not trivial.

Let ‘ $A$ ’ name a particular pain sensation, and let ‘ $B$ ’ name the corresponding brain state, or the brain state some identity theorist wishes to identify with  $A$ . *Prima facie*, it would seem that it is at least logically possible that  $B$  should have existed (Jones’s brain could have been in exactly that state at the time in question) without Jones feeling any pain at all, and thus without the presence of  $A$ . Once again, the identity theorist cannot admit the possibility cheerfully and proceed from there; consistency, and the principle of the necessity of identities using rigid designators, disallows any such course. If  $A$  and  $B$  were identical, the identity would have to be necessary. The difficulty can hardly be evaded by arguing that although  $B$  could not exist without  $A$ , *being a pain* is merely a contingent property of  $A$ , and that therefore the presence of  $B$  without pain does not imply the presence of  $B$  without  $A$ . Can any case of essence be more obvious than the fact that *being a pain* is a necessary property of each pain? The identity theorist who wishes to adopt the strategy in question must even

<sup>2</sup> Of course, the body *does* exist without the mind and presumably without the person, when the body is a corpse. This consideration, if accepted, would already show that a person and his body are distinct. (See David Wiggins, ‘On Being at the Same Place at the Same Time’, *Philosophical Review*, Vol. 77 (1968), pp. 90–5.) Similarly, it can be argued that a statue is not the hunk of matter of which it is composed. In the latter case, however, one might say instead that the former is ‘nothing over and above’ the latter; and the same device might be tried for the relation of the person and the body. The difficulties in the text would not then arise in the same form, but analogous difficulties would appear. A theory that a person is nothing over and above his body in the way that a statue is nothing over and above the matter of which it is composed, would have to hold that (necessarily) a person exists if and only if his body exists and has a certain additional physical organization. Such a thesis would be subject to modal difficulties similar to those besetting the ordinary identity thesis, and the same would apply to suggested analogues replacing the identification of mental states with physical states. A further discussion of this matter must be left for another place. Another view which I will not discuss, although I have little tendency to accept it and am not even certain that it has been set out with genuine clarity, is the so-called functional state view of psychological concepts.

argue that *being a sensation* is a contingent property of *A*, for *prima facie* it would seem logically possible that *B* could exist without any sensation with which it might plausibly be identified. Consider a particular pain, or other sensation, that you once had. Do you find it at all plausible that *that very sensation* could have existed without being a sensation, the way a certain inventor (Franklin) could have existed without being an inventor?

I mention this strategy because it seems to me to be adopted by a large number of identity theorists. These theorists, believing as they do that the supposed identity of a brain state with the corresponding mental state is to be analyzed on the paradigm of the contingent identity of Benjamin Franklin with the inventor of bifocals, realize that just as his contingent activity made Benjamin Franklin into the inventor of bifocals, so some contingent property of the brain state must make it into a pain. Generally they wish this property to be one statable in physical or at least 'topic-neutral' language, so that the materialist cannot be accused of positing irreducible nonphysical properties.

A typical view is that *being a pain*, as a property of a physical state, is to be analyzed in terms of the 'causal role' of the state,<sup>3</sup> in terms of the characteristic stimuli (e.g., pinpricks) which cause it and the characteristic behavior it causes. I will not go into the details of such analyses, even though I usually find them faulty on specific grounds in addition to the general modal considerations I argue here. All I need to observe here is that the 'causal role' of the physical state is regarded by the theorists in question as a contingent property of the state, and thus it is supposed to be a contingent property of the state that it is a mental state at all, let alone that it is something as specific as a pain. To repeat, this notion seems to me self-evidently absurd. It amounts to the view that the *very pain I now have* could have existed without being a mental state at all.

<sup>3</sup> For example, David Armstrong, *A Materialist Theory of the Mind*, London and New York, 1968, see the discussion review by Thomas Nagel, *Philosophical Review* 79, (1970), pp. 394–403; and David Lewis, 'An Argument for the Identity Theory', *The Journal of Philosophy*, pp. 17–25.