

into being under the care of A, is almost numerically the same picture which, on Monday, A discovered on a tablet belonging to B (the converse claim at all promising.) If this is right, then it supports the thought that identity itself is not only all or nothing. It actively excludes (or so it now appears) the ideas of just making it, almost making it, a near miss.

Those who want always to be sure there is some formal demonstration in the offing may see the doubts I have just rehearsed about the possibility of an entity *a*'s being almost identical with an entity *b* (and my reading of Caecilia's claim as altogether innocent of that idea) as furnishing an advance commentary on the following derivation:

- (i) $a = b$ (Hypothesis)
- (ii) a is not almost identical with a and is not almost distinct from a
- (iii) a is not almost identical with b and is not almost distinct from b
- (iv) $(a = b) \rightarrow (a$ is not almost identical with b and is not almost distinct from $b)$
- (v) $(a$ is almost identical with b or a is almost distinct from $b) \rightarrow a \neq b$.

Commentary. (v) is the contraposition of (iv). (iv) is arrived at by conditional proof from (iii) and (i), discharging (i). (iii) comes from (ii) and Leibniz's Law. (ii) is a substantive claim about *a*'s identity with *a*. In (ii) the 'not almost' is read twice over in one natural way. This is the same as the way in which it is read in (iv). (ii) itself reflects one special thing about identity. It reflects the same thing that organizes the personal efforts of all the preceding paragraphs of this section. Moreover, (ii) is true.

12. CONCLUSION

In this chapter there is just one rash claim about identity (and singling out) that I have held back from entering. I have not said that findings of identity are uncontentious. All or nothing, no near misses, determinate and all the rest – that is one thing. *Uncontentious* is another thing altogether. Identity can be all or nothing, determinate, exclude degree, exclude near misses . . . and be contentious. But the theory of identity that allows findings of identity to be contentious – this I have aspired to render less contentious. Once sober conceptualism comes to the aid of ordinary common sense, philosophy has no need (I declare) to supplant the common sense conception according to which identity is utterly special and unique.

CHAPTER SEVEN

Personal identity

*You would not find out the bounds of the soul, though you traversed every path:
so deep is its logos.*

Heraclitus

The human body is the best picture of the human soul.

Wittgenstein, *Philosophical Investigations*, p. 178 (trans. Anscombe).

PART ONE

I. AN EXPEDITIOUS IF PRECIPITATE ANSWER TO THE QUESTION OF PERSONAL IDENTITY

If the thesis of the Sortal Dependency of Identity (**D**) is correct, then a simple expectation arises: namely that, in the case where A and B are persons, the content of the sentence 'A is the same as B' may be elucidated, through the equivalence of 'A = B' and 'A is the same something as B', as requiring for its truth that A should coincide with B under the concept *person*. This can scarcely be wrong. But, since the question that it pushes into immediate prominence, of what sort of thing persons are, leads into a thicket of philosophy, anyone who has been once through the thicket may wonder whether it can be skirted.

One way round seems to be to reflect that, if A and B are the kind of creature that we take ourselves to be – the kind of creatures writing or reading this book, I mean – then an equally good answer to the *same what?* question ought to be 'human beings'. For if A and B are persons and human beings, then A could never be the same person as B and not the same human being as B; or the same human being as B and not the same person. (Cf. Chapter One, §2 following.) So, *given* the human beinghood of A and B, this furnishes a perfectly good covering concept for the identity 'A is the same as B'. 'Person' and 'human being' differ in sense. They may even differ in their extension. But that

is immaterial.¹ What matters is that here, in so far as they assign any the concepts *person* and *human being* assign the same underlying principle of individuation to A and to B, and that that principle, the *human being* principle, is the one that we have to consult in order to move towards the determination of the truth or falsehood of the judgment that A is B.

Human beings, endowed as they are with a distinctive mode of activity of their own, are substances in the rich sense of 'substance' that has been built up in the preceding chapters.² Their mode of activity is nothing unknown to us. So one might well wonder how questions of identity covered by the *human being* concept could fail to find the notion of identity at its most straightforward and unproblematical best. By admitting of answers that are principled, such questions ought to exclude the risk (and even the risk of the risk) of indeterminacy. See Chapter Six. Ought it not to be safe then, in any case where it is not doubtful that A and B are determinately identified, to see the apparent non-determinacy of truth-value residing in a claim 'A = B' as entailing that A is actually not B? On this view, the difficulty of undecided questions of personal identity ought to derive either from faulty formulation or from ignorance. Or else, if *human being* is metaphysically troublesome (which need not be very surprising), then we must attend again, at need, to the question of what human beings are (cf. §16) in order that we be better placed to apply to the problem of their individuation the principled ingenuity and vigilance that ought to be promoted by any adequate dialectic of sameness and otherness.

2. DOUBTS, AND ANSWERS TO DOUBTS: SUBJECTS OF CONSCIOUSNESS

If, as some declare, these expectations (of determinacy and the rest) are not in practice borne out, why is this? It may be said that the reason is that the proposal just made will be deeply controversial so soon as it is applied to sentences of the form 'I am . . .', as in 'I am

¹ That they are not synonymous makes easier the expression of certain truths that it would be a pity for us to dispense with. Or so I aver at pp. 69–70 of Wiggins (1987b) and again at pp. 244 ff of Lovibond and Williams (1996).

How may sortal predicates differ in extension but impose the same principle of individuation? Here is one model: 'tinker', 'tailor', 'soldier', 'sailor' all impose the principle *man/human being*. This is not the only model.

² More precisely — and imprecision is dangerous here — 'substance' has been shown to stand for a determinable whose particular determinations, *human being, ilex, mouse, . . .* must be expected to be rich, copious and exigent in what they require of things that answer to them.

Lohengrin', or 'I am thy father's spirit'. 'Such expectations were bound to fail' it may be said, 'because they have so little regard for the fact that, first and foremost, human beings experience themselves as subjects of consciousness.'

As a diagnosis of why personal identity is an issue that will not easily go away, this is plausible. But it is unclear, I should say, how much further regard, beyond that already accorded to it, a philosophical account of personal identity needs to have for the fact of our experiencing ourselves as subjects of consciousness. In advance of philosophy's magnifying or distorting things, what else is there here but this: that we experience ourselves as embodied and changeable substances among other human or non-human substances; that, even as we perceive things or do things or undergo things done to us, we experience ourselves as human beings interacting with other animate or inanimate beings (and can experience ourselves as experiencing ourselves so)? But, if this is what things come down to, it carries us back to the human being conception, to the philosophically familiar idea of a person as possessor of P-predicates and M-predicates,³ and then to the question of how human beings do in fact get the idea of the distinction between that which is themselves and that which is other than themselves or grasp therewith the moral significance of that distinction.

At this point, various objectors will complain that too much that is bluntly empirical has been built too soon into the claim that we experience ourselves as subjects of consciousness; that the whole tenor of the explication just given falsifies, as if deliberately, the real depth and difficulty of the original finding that we conceive ourselves as subjects of consciousness or as conscious selves.

In the face of such a complaint, there is a temptation for a *human being* theorist such as myself to find fault with the way in which the determinable 'conscious subject' has now been transformed into a fully fledged philosophical substantive of a non-empirical kind. Such a

³ See P. F. Strawson (1959), Chapter Three. In the argument of that chapter, I should like to reconstrue P-predicates as predicates standing for properties of a person that are directly or indirectly consciousness-involving and M-predicates as predicates standing for properties of a person that are directly or indirectly body-involving. Physicalists can say then that, so far from being disjoint, all P-predicates are M-predicates (but not of course vice versa).

Other works that have left their mark on the contentions of this chapter are Chapter One of Stuart Hampshire, *Thought and Action* (London, 1959); Lev Vygotsky, *Thought and Language* (Cambridge, Mass., 1986), revised translation edited by A. Kozulin — see especially the last chapter; Lev Vygotsky, *Mind in Society: The Development of Higher Psychological Processes*, ed. and trans. M. Cole, V. John-Steiner, S. Scribner, E. Souberman (Cambridge, Mass., 1978); David Bakhurst, *Consciousness and Revolution* (Cambridge, 1991), see Chapter Three.

theorist will say that he is reminded of the way in which the determinable 'substance' is misunderstood (and then welcomed or turned away on all the wrong sorts of ground) as purporting to stand for a special and specific kind of thing. He will protest bitterly at the way in which, rather than see the dummy sortal or determinable predicate, 'subject' as the first fix by which one fumbles to subsume X (in this case some conscious thing) under a more specific and informative sortal specification, philosophers have come to think of a conscious subject as a peculiar kind of thing reached by a special kind of abstraction. One can only be amazed (he will say) that a mode of abstraction so dangerously similar to the procedures of creative definition⁴ should come to be seen as legitimating, without further ado, a whole new branch of philosophical thinking.⁵

The matter cannot be so simply wound up, however, by warning about the provenance and credentials of the term 'subject'. For the *human being* theory will still appear to do insufficient justice to a line of reflection still prompted by John Locke's account of these things: what I am in the present ('my present self') always lies under the cognitive and affective influence of what I remember having been or having done or undergone in the past, no less than of that which I intend or am striving to make real in the present or the future. But if it is the nature of persons to be remembering beings whose conception of what they themselves are is all of a piece with their experiential memory, then some constitutive connexion ought to be expected (it will be said) between their experiential memory and their identity. But, in that case (it will be asked), how can one be content in the simple expectation announced at the beginning, namely that personal identity for creatures such as A and B should amount to no more than the identity of the human beings A and B? So soon as the Lockean insight is developed in the direction of the theory of personal identity that it suggests, namely that person A is person B if and only if A's consciousness (or the subject self that A is) corresponds to B's consciousness (or the subject self that B is), it will prove that our judgments of personal identity can diverge in indefinitely many cases from those enforced by the human being account. Or so it will be objected.

⁴ On creative definition, see still Patrick Suppes, *Introduction to Logic* (Princeton, N.J., 1957), Chapter Eight.

⁵ The defensibility of the stand I take here is much enhanced by Quassim Cassam's painstaking and maybe exhaustive enumeration and critique of subject theories of the kind I am opposing. See *Self and World* (Oxford, 1997). See my (1987) and "Replies" in Lovibond and Williams (1996), cited at note 1 above.

3. THE LOCKEAN CONCEPTION; AND BUTLER'S CRITICISMS OF SUCH CONCEPTIONS

Familiar though it is, the reflection just rehearsed needs to be set out more historically and in a little more detail, if its force is to be measured and the response to it is to be fully clear.

Locke thought that it was among the defining marks of a person to be able to consider oneself the same thinking being in different times and places. (See *Essay Concerning Human Understanding* II.xxvii.9: 'A person is a thinking, intelligent being, that has reason and reflection, and can consider itself as itself, the same thing, in different times and places.') In Locke's account, identity of person is distinguished from identity of man. The apparent divergence of Locke's conception of personal identity – which is avowedly both subjective (inner) and (he says) 'forensic' – from the human being conception of personal identity, which I advocate, comes to light most clearly in cases where the human being theorists would say that B had completely forgotten what A had done. (The example that follows is drawn from Reid's *Essays on the Intellectual Powers of Man*.) The general has forgotten the exploits of the boy. He has forgotten stealing apples on some distant occasion that was at the time noteworthy in his life. For a human being theorist, as for one who believes in the unitary persistence of the soul, such forgetting is not constitutively relevant to the question of personal identity. For the Lockean, on the other hand, it is a challenge to reflection that ought not to be impatiently foreclosed.

Among the earliest critics of Locke's account of personal identity was Bishop Butler, the guardian at once of common sense and the substantial soul. In his polemic against the peculiarities of the divided account that Locke had been drawn into offering, Butler continued with a logical objection:

'though consciousness of what is past does thus ascertain our personal identity to ourselves, yet to say it makes personal identity, or is necessary to our being the same persons, is to say that a person has not existed a single moment nor done one action but what he can remember; indeed none but what he reflects upon. And one should really think it self-evident, that consciousness of personal identity presupposes, and therefore cannot constitute, personal identity, any more than knowledge, in any other case, can constitute truth which it presupposes.' (*First Dissertation to the Analogy of Religion*. (Compare Samuel Clarke, *Works*, III, p. 787 – a reference I owe to Mr Dewey Ducharme.))

One might wonder why this passage of Butler has not been permitted to settle the whole matter once and for all in favour of a substance

account, either a human being account or else a soul account. Not only does it seem obvious that the general is the same person and the same subject as the boy. To one content to work within the framework which we owe to miscellaneous insights of P. F. Strawson (see especially Chapter Three of his *Individuals*) and of Donald Davidson, it will appear plainly that the most interesting point in the case sustaining Locke's insight could have been readily accommodated within the human being view. Let us see how.

Par excellence, a person is not simply one of us or a proper object of our reciprocity,⁶ but a subject also of interpretation, a being that both interprets and is interpreted. In making sense of others and of ourselves in the business of interpretation, we cannot, however, make do with the bare idea of 'a subject of interpretation'. We need some further notion, however rough and ready, of what sort of thing is to count, empirically or operationally speaking, as a subject of interpretation. The grammatical substantive 'interpreter' lies too close to being a mere determinable to serve the purposes of individuation. If 'subject' or 'interpreter' is to play any properly sortal rôle at all here, if sense is to be made of how subjects of interpretation find other subjects of interpretation, there is need for some further determination. As we find our feet in the world and learn to adjust to others' expectations, as we begin to understand what we mean by the words we have learned already to utter, as we extend our capacity to interact with others and participate or co-operate with them, what stereotype of personhood do we have to catch onto, clearly if not distinctly, and learn to elaborate?⁷ In its individuating and recognitional non-effectiveness, the conception furnished by John Locke is not very much better than the bare idea of 'interpreter'. Even here, however, we may be able to make a new beginning from the materials that Locke provided. People are *animate beings* (we may revise his declaration to say) – or *palpably substantial souls*, if you prefer – *bearers of P-predicates and M-predicates*, who not only *think things, but do things and undergo things*, beings who reason, reflect, perceive, feel, imagine, desire, intend and have experiential memory . . . are happy or miserable . . . conceive

⁶ For the thought that 'us' has an essential occurrence in this claim, see *op. cit.* at note 1 above.

⁷ Pace Peter Winch's presidential address, 'Eine Einstellung zur Seele', *Proceedings of the Aristotelian Society* (1980-1), the question does *not* presuppose that we find our feet in the world by procedures analogous to those that the philosophers of semantic interpretation have described. It only presupposes that, beginning and proceeding however we do, participatively and interactively no doubt, we reach some point where we know roughly what a person is. In catching on, we grasp the stereotype, clearly though indistinctly. For my use of the word 'stereotype' here, see Chapter Three, note 2, last paragraph. For clear and indistinct knowledge, see Chapter Three, footnote 6.

of themselves thus (as above) . . . who have and conceive of themselves as having a past accessible in experiential memory and a future accessible in intention. . . . Is it not by virtue of our grasping practically and following through some such stereotype, is it not by virtue of our perfecting our application of it in the perceptible world, that the marks of personhood come home to us, and are available to be elaborated in philosophy? More soberly, let us say that in this way the marks are assembled of persons *as we know them* from the only case we shall ever become familiar with, namely that of persons who are human beings. Finally, at the last step, the position of experiential memory comes to the fore. If experiential memory is a mark of personhood, and persons need to conceive of this faculty as essential to their own way of being, then that secures its full significance to our expectation that, when people do or suffer something, this will impress itself on their mind, extend their information, colour their experience and influence their future responses. In the absence of that expectation, few if any of our finer-grained interpretations of people or the practices these expectations support could ever be sustained.

If so much is correct, then, because of its central role in interpretation, experiential memory plays a special part in the full picture of personhood. But that finding (I insist) does not imply that experiential memory will impinge on the necessary (or sufficient) conditions of the *survival* of persons or play any role in the statement of sufficient conditions of survival. (Still less is memory put at the service of a construction of some self that is distinct from the human being.) Locke can be right about remembering as central among the marks of personhood without Locke's or the neo-Lockeans' being right about personal identity.

Neo-Lockeans have ignored such partial accommodations and distracted attention from them. Human being theorists may justly complain of that. But there are at least two other reasons why Butler's criticism of Locke's theory of personal identity has been found inconclusive. The first is that so many of Butler's admirers have appeared to overlook the possibility (the bare theoretical possibility, I mean, but it is one much prized by neo-Lockeans) that the act of experientially remembering (remembering 'from the inside') planting a fig tree (say) might be modelled (logico-grammatically speaking) on such acts as conceiving or imagining or visualizing planting a fig tree.⁸ For the notable point about

⁸ These are acts where the mental content imagined, visualized or conceived does not call for the subject of the act, the planter, to be specified. *A fortiori*, the subject need not be specified as the same as the conceiver, imaginer or visualizer. On these and related matters, see Bernard

acts such as these is that they do not make the kind of reference back to the self which invited Butler's charge against Locke. Taken by themselves, the would-be truisms adduced by the Butlerians are not enough to vindicate Butler's criticisms. And so much must be admitted by human being theorists.

In the second place, friends of Butler and human being theorists have sometimes overlooked how easily the boy whose exploits the general has forgotten can be reunited by the neo-Lockean view with the general himself. Lockeans were well within their rights to be irritated when Butlerians represented that there was an easy objection against Locke to be found in this quarter. The Lockean's point is worth making more explicit – not so much because it is decisive in Locke's or anyone else's favour as for the sake of a fairer recapitulation of some of the saner or more durable preoccupations of the philosophy of personal identity. In the absence of such a recapitulation, Butler's central insight will continue to be obscured, or its real significance will not be appreciated.

4. A NEO-LOCKEAN IDENTITY-CONDITION

Consider the relation that holds between person Q and person P when, for some sufficiency of things actually done, witnessed, experienced . . . by person P, person Q has at a time after the doing, witnessing, experiencing, some sufficient real⁹ recollection of doing, witnessing, experiencing . . . them.¹⁰ Let us call this relation (suitably tidied up with more about 'sufficiency', and glossed somehow to allow for sleep and other forms of inattention) the relation of *strong co-consciousness*. Then anyone bent on grasping the nerve of Locke's conception of person will see that the identity-condition he really has to consider is one that begins with

footnote 8 (cont.)

Williams, 'Imagination and the Self', in Williams (1973), especially p. 43 and thereabouts; David Wiggins, 'Locke, Butler and the Stream of Consciousness: and Men as a Natural Kind', *Philosophy*, 51 (1976), especially the second section (to which the present chapter is, of course, a necessary correction and supplement); David Velleman, 'Self to Self', *Philosophical Review*, 105, 1 (January 1996).

⁹ The question will soon arise whether the words 'or apparent' should be inserted between 'real' and 'recollection'. My view is that it must not. See the end of §5 below.

¹⁰ The notion of sufficiency figures twice in the definition. There is a great deal to be supplied here from the basis of ordinary norms and expectations relating to memory. At any given time one expects a person to remember fairly well his relatively recent past; and remembering one's very recent past involves being able to muster a rather detailed and complete narrative. Where t_1 is much later than t_2 , slightly less may be expected. The objections we shall urge against the Lockean account of personal identity are not very sensitive to how these details are supplied. But 'sufficiently' is to be glossed realistically, not in such a way that anything whatever qualifies as remembering provided the event is distant enough.

the idea that, so long as one persists, one normally remembers at least from one time to the next what one was doing or witnessing or suffering at the earlier time. Making use of that relation, one who wants to devise a Lockean identity-condition could first define a new relation C by disjoining co-consciousness with its converse (that is with the relation P has to Q if P is strongly co-conscious with Q) and then introduce the ancestral *C of that relation. (Compare Locke's own use of the word 'extend' at *Essay* II, xxvii, 9.)

With a view to stating the Lockean identity-condition explicitly, let us say then that, if Q remembers at t' doing, witnessing or experiencing a sufficiency of what P did or witnessed or experienced at t , then the ordered pair $\langle Q, t' \rangle$ stands in the strong co-consciousness relation to $\langle P, t \rangle$. Letting C stand for the disjunction of that relation with its converse, we have the relation

$$\langle P, t \rangle C \langle Q, t' \rangle.$$

Then, as promised, we define a new relation *C that an ordered pair of person and time may have to another ordered pair. Thus

$$\begin{aligned} \langle X, t_j \rangle *C \langle Y, t_m \rangle \text{ is true just in case} \\ \text{either } \langle X, t_j \rangle C \langle Y, t_m \rangle \\ \text{or else, for some } \langle W, t_i \rangle, \\ \langle X, t_j \rangle C \langle W, t_i \rangle \text{ and } \langle W, t_i \rangle C \langle Y, t_m \rangle \\ \text{or else, for some } \langle W, t_j \rangle \text{ and some } \langle Z, t_k \rangle, \\ \langle X, t_j \rangle C \langle W, t_j \rangle \text{ and } \langle W, t_j \rangle C \langle Z, t_k \rangle \text{ and } \langle Z, t_k \rangle C \langle Y, t_m \rangle \\ \text{or else for some } \langle W, t_j \rangle, \langle Z, t_k \rangle, \dots \end{aligned}$$

Note that, in the absence of doctrine about personal identity or the semantics of 'recollect' or 'remember', it is simply not stipulated whether C itself is a transitive relation. But *C, introducing a chain of pairs consisting of persons and times at which they remember, and requiring for its satisfaction only membership of such a chain, gives a simple co-consciousness relation that is transitive by definition. It makes possible the following amended neo-Lockean equivalence:

$$\begin{aligned} I_p: P \text{ is the same person as } Q \\ \text{iff } (\exists t, t') (t < t') \text{ and } \langle P, t \rangle *C \langle Q, t' \rangle. \end{aligned}$$

A Lockean who is not too discontent with this formulation of his intentions may be expected to insist that the 'P' and 'Q' here hold places for designations of persisting things or continuants that 'consider themselves as themselves at different times and places'. I could only applaud

that insistence (however unclear I find the phrase just quoted), The strategy with the boy, his stealing the apples at t , and the general is that, if the general doesn't in 1700, when he's a general, remember stealing apples in 1660, because he doesn't have sufficient recollection of a sufficiency of what he did in 1660, then we must look for someone (don't say someone *else*, say *someone*), a lieutenant in the Hussars, say, and at an intermediate time, 1680, such that (a) that lieutenant's doings, witnessings and experiences as a Hussar are ones the general still remembers in 1700 and (b) the lieutenant remembers well enough in 1680 stealing the apples in 1660.

Let it be clear that there is nothing in these explanations of C and *C that Locke or his followers will want to see as committing them to abandon the ontology of mental continuants.¹¹ Let it be clear too that, if the definition of C is read just as we stated it and that of C* is understood conformably with that, then the full strength of the right-hand side of I_p may be brought out by drawing at need upon the causal conditions of remembering and/or real recollecting that are transferred more or less unexamined into the definition of C. For that reason, I_p is better than it at first appears. But this is no guarantee that it amounts to a faithful elucidation of personal identity or even to a formally satisfactory proposal.

I_p is a theory that the Butlerians will want to criticize. But it will be noted that, once the ancestral of C is deployed and C's full import is noted and enforced, the number of apparent divergences between the neo-Lockean account of personal identity and the *human being* account appears to be immensely reduced. For the general and the boy do figure together within the extent of a *C chain of paired persons and times. On the other hand, no deluded empathizer who has read my diary will become a candidate to stand in such a *C chain with me just by virtue of being in the state of seeming vividly and completely to remember most of what I did when I came home to find the kitchen in flames. The chief disagreements that remain will seem to turn on cases that many philosophers will find problematic in their own right. These are cases such as that of a person knocked down in a road accident, who never recovers consciousness at all even though (as a human being theorist would say) he lies in hospital for many months afterwards, still alive; or the person who recovers consciousness but has entirely forgotten (as I

¹¹ Fixing on the particular times t, t' , of an experience and the recollection of it, nothing prevents the Lockeans from isolating the two-place relation holding for t, t' between the experiencer and the rememberer as follows: $\lambda x y ((x, t) *C (y, t'))$.

should say) who he is and everything else about himself. Such cases remain in dispute. But they are a poor locus for a decisive engagement over personal identity.

Here then is an interim report on the state of the argument. There is the human being theory (in purely logical alliance with Butler and with Reid) and there is the neo-Lockean theory. The fundamental disagreement is over whether experiential memory is in some way constitutive of personal identity over time. The human being theorist denies that it is so constitutive, even though he ranks the capacity for memory as one part of personhood. The Lockean persists in maintaining that experiential memory is constitutive of personal identity over time. This disagreement is important. Yet it is remarkable how closely, once we adjust the Lockean position in all the obvious ways that have issued in the proposal I_p , the two positions can apparently be aligned with regard to normal cases. The scene is now clear for a more decisive attempt to resolve the disagreement.

5. BUTLER'S CENTRAL INSIGHT

It is not hard to reconstruct Butler's comment on I_p . Consider the case where we do not know yet if B is A, but we think we may be able to establish whether now B is co-conscious with A. Suppose that we know that, probably inadvertently, A once (in August 1990) caused a fire in the library-stack in the library of the Chigwell College of Commerce. It may seem we should simply ask: does B remember anything of causing such a fire? It seems a fair test for whether B is A. (It is not the sort of thing people readily forget having done. Nor are many others likely to have done the very same act.) So we arrange for the question to arise somehow, but in a manner that does not make B jump to the conclusion that he is about to be handed over to the police (or yet that he is due to receive a medal). Suppose that it appears in due course that B does remember this, so that (for practical purposes) we seem very close to establishing the identity. What then?

In this case, I_p may appear for one moment to be an effective and plausible principle of identity. But how does it work? Well, B's seeming to remember setting fire to the book stack suggests to the inquirer, in the context of his knowledge that there was this fire in which A was involved, that B can indeed remember causing that fire in the book stack. Does the inquirer's new finding not come close enough to establishing that B is strongly co-conscious with A?

One might agree that the finding narrows the gap. But one can still ask how exactly the investigator argues from B's appearing to remember to B's actually remembering this incendiary activity. After all, B is only the same person as A if his seeming to remember is his really and truly remembering setting fire to the book stack. The answer will probably be that B can rehearse the act of setting fire to the book stack as if from the point of the view of the perpetrator and it seems to be the best explanation of B's ability to rehearse these things/events as if from the point of view of the perpetrator that B was there and did do the act.

All right. But now, on behalf of Butler and his doubt against Locke, it must be inquired what the foregoing argument takes for granted. The answer is that the argument takes for granted some understanding of *what it is* for someone to do such an act at a place and a time, survive, move on to other places and times, and remember later the earlier act, place and time. So it appears to take for granted the possibility of an account of identity other than I_p . Or (as I should say, deploying the terminology of this book) it takes for granted a criterion constitutively determining that for which all sorts of things can count as tests (tests the favourable outcome of which can count as evidence). What is this other account of identity? Even if it is consistent with I_p (which it may be or may not be) or even if it subsumes I_p , this other account can scarcely be I_p itself. For I_p presupposes the concept of real experiential memory, and, once we see that that presupposes the idea of someone's doing something and *then* correctly remembering doing it, we are reminded how exactly experiential memory presupposes identity. *Where someone appears to remember starting that fire, they can't be right unless they were indeed there at the fire.*¹² Pending the production of some other account of the identities of subjects, one that is anterior to I_p , the only visible candidate is the sort of account that is offered by the *human being* theorist.

¹² This way of explicating Butler is indebted to Bernard Williams (1956–7) and to A. J. Ayer's *The Problem of Knowledge* (Harmondsworth, Middx., 1956), p. 196. For an earlier critique of Locke, see A. G. N. Flew (1951) – to which I have come to see that the 1980 chapter, which the present chapter replaces, was culpably unjust.

Here is another fix on the same underlying point that I want to urge against Locke. Locke writes (ii.xxvii.17), 'That with which the *consciousness* of this present thinking thing can join itself, makes the same *Person*, and is one *self* with it and with nothing else; and so attributes to it *self* and owns all the Actions of that thing as its own as far as that consciousness reaches.' The thinking thing that owns Actions as its own (acknowledges them, etc) needs to have got itself somehow some conception of itself *as* the thing that thinks etc. So soon as it follows through its conception of that thing, however, it must know that it is neither necessary nor sufficient for an Action's being that thing's that the thing should be prepared to 'join itself' to that Action. Apart from the case of forgetfulness, the said conception needs to take account of the cases of shame, delusion, confusion. . . .

Here the verdict must go to Bishop Butler, or so I conclude, or to his latter-day followers. It goes to him, not for the sake of his positive account of personal identity,¹³ but for the sake of the thing that is validated as telling in his objection to John Locke's account.

At this point, the friends of John Locke may be tempted to try to escape Butler's point by experimenting with the definition of C and *C. They may weaken the right-hand side, for instance, so that it shall only require that one who is a candidate to be in some relation C to P should have sufficient real *or apparent* recollection of doing, witnessing, experiencing whatever P did witness/experience.¹⁴ Their mistake hitherto was to be insufficiently bold in adjusting the idea of experiential memory, the Lockians might say; perhaps they have been too timid in facing up to the potentially radical ontological consequences of their original insight.

There is no formal difficulty in the weakening manoeuvre, however questionable its philosophical motivation. But it should be noted that the price exacted will be a further diminution in the plausibility of I_p , and an altogether new kind of mismatch, a serious mismatch, with the common sense conception's verdicts on cases – as with the verdicts of the human being conception. The neo-Lockean must do much less or much more than tinker with the definitions of C and *C. (See below, §10.)

Let us call the relation to be aimed at in all future attempts by the neo-Lockians the relation R.¹⁵ As for *C itself, defined as we first defined it in the most ordinary terms of ordinary language, let us not modify it, but leave it just where we first had it, namely as part of one possible explication of the idea (almost certainly identity-involving, in fact, but that is a part of the philosophy of the matter, and nothing to be settled straight off by definition or stipulation) of the philosophical notion of co-consciousness. If C and *C have any interest, it will transcend that of the formulation I_p .

6. A NEO-LOCKEAN CONCEPTION

To the friends of Butler, that might seem to be the end of the matter. But it is notorious that the neo-Lockians have not wanted to give up. No

¹³ Butler's own account of identity is imperfect and at some points confused. In fact, it comes as a relief that he unreasonably refuses to elucidate a notion so simple, so final and so spiritual as that of identity of person. For the best statement of something that is closer to what Butler, given his beliefs and preoccupations, really ought to have said, see Thomas Reid (1785), II, §4.

¹⁴ Of course, *some* strength might be restored to the right-hand side by reimporting (somehow) any causal components that are lost by inserting 'or apparent'. See below, §6 following. But the considerations urged in the next paragraph will still apply.

¹⁵ I call it R because Derek Parfit calls it R. Let not the relation R be confused with the thesis R of the relativity of identity, denied in Chapter One; or C with C, denied at p. 75.

doubt there have been bad reasons for this – blind perseverance allied with the inertia of misplaced ingenuity, phenomenological (would-be phenomenological) overdrive, a hypostatization of the conscious subject (which prepares some of us to reject the conscious subject as so characterized) or the misunderstanding of the traditional idea of a substance (allied, no doubt, with an unacknowledged residual attachment to it or dependence upon it). But perhaps there have been good reasons, as well.

In a book published thirty-seven years ago and devoted in part to the diagnosis of a mistake that was far commoner then than it is now (in 1999/2000) – namely that of supposing that ordinary claims of experiential memory have to represent themselves as claims of identity (rather than as importing the presupposition of it) – Sydney Shoemaker proposed a striking thought experiment:

‘Suppose that medical science has developed a technique whereby a surgeon can completely remove a person’s brain from his head, examine or operate on it, and then put it back in his skull (regrafting the nerves, blood-vessels, and so forth) without causing death or permanent injury. . . . One day a surgeon discovers that an assistant has made a horrible mistake. Two men, a Mr Brown and a Mr Robinson, had been operated on for brain tumours, and brain extractions had been performed on both of them. At the end of the operations, however, the assistant inadvertently put Brown’s brain in Robinson’s head, and Robinson’s brain in Brown’s head. One of these men immediately dies, but the other, the one with Robinson’s body and Brown’s brain, eventually regains consciousness. Let us call the latter ‘Brownson’. . . . He recognizes Brown’s wife and family (whom Robinson had never met), and is able to describe in detail events in Brown’s life, always describing them as events in his own life. Of Robinson’s past life he evidences no knowledge at all. Over a period of time he is observed to display all of the personality traits, mannerisms, interests, likes and dislikes, and so on that had previously characterized Brown, and to act and talk in ways completely alien to the old Robinson.

What would we say if such a thing happened? There is little question that many of us would be inclined, and rather strongly inclined, to say that while Brownson has Robinson’s body he is actually Brown. But if we did say this we certainly would not be using bodily identity as our criterion of personal identity. To be sure, we are supposing Brownson to have *part* of Brown’s body, namely his brain. But it would be absurd to suggest that brain identity is our criterion of personal identity.¹⁶

At the time when Shoemaker’s book appeared, and along with almost everyone else, I was extremely impressed by this example. Yet I see now that we were not all impressed for the same reason. Some were

¹⁶ (1963), pp. 23–4. (See also pp. 20, 30, 193–4, 231, 245–71.)

impressed by the simple thought that, if the brain transfer were performed with preternatural dexterity, then Brownson’s experience would be a subjectively seamless continuation of Brown’s. Others, a smaller group perhaps, but the group to which I belonged and would have avowed loyalty (this is not only hindsight), would have preferred to say that the special thing about Brownson was that he was the functional inheritor and continuator of all of Brown’s vital faculties. *This* was the reason why Brownson counted as the unique inheritor of the title to be Brown, the reason why Brownson was Brown, that very substance. Neither Brown nor Robinson nor Brownson *was* a brain. But the brain, being the seat of memory and consciousness, was not just any old part of the body among others.¹⁷ It was the essential nucleus of a person (of a human being) – or so we were wont to maintain. Moreover, experiential memory was more than a mere potentiality. It was a shaped and developed capacity, conspicuous among all the vital functions of personhood and coeval (in persons as they are known to us) with our developed faculties for sentience, locomotion, desire, cognition, . . . etc.¹⁸ Why then should not the embodied faculty for experiential memory enter into the whole *principle of activity* of persons conceived as we conceive them? (See **D(v)**, Chapter Two, §7.) Why should it not enter into some plausible emendation of the serviceable principle suggested by a misappropriation of Locke’s own words ‘the identity of the same *man* consists in nothing but a participation of the same continued life, by constantly fleeting particles of matter, in succession vitally united to the same organized body’? Why not let the emendation that is needed for the sake of the Brown–Brownson case, considered as a case of personal identity, require for the persistence of a person the operation of the same principle of activity and/or the participation of the same continued life ‘vitally united to the same vehicle’ or ‘vitally united to the same seat of vital functions’?

¹⁷ See my (1967), p. 45, para. 2. (See also p. 246 of the second item cited at note 1.) There was nothing to prevent one from being impressed by both kinds of consideration. I must have had some regard to the other kind when, a little later, I wrote ‘we should scarcely allow a criminal to escape the penalty due to the unique doer of a criminal act by contriving his own fission’. See ‘Locke, Butler and the Stream of Consciousness: and Men as a Natural Kind’, *op. cit.* note 8 above, the precursor of the chapter that the present chapter replaces. But this kind of consideration was secondary.

¹⁸ Even if Brownson had taken some time to settle down after the operation to get used to his body, to exercise its neural connexions and resume Brown’s life – and philosophers are still apt to underestimate the preternatural dexterity and knowledge that the imaginary surgeon and his equally imaginary team of anaesthetists, suturists, radiographers, laser-technicians, physiotherapists, psychotherapists, counsellors and the rest, would have to bring to bear – , the need for such a convalescence might not have made any essential difference to that which then impressed one.

7. UNFINISHED BUSINESS

In 1965–6, when I first came across Shoemaker's example and was drawn into this response, I ought to have been much more troubled by the point that, even if Brownson could talk like Brown (which would not in itself have been easy – if it was required off-stage as well as on-stage, so to speak), he could scarcely have stood and walked and run and jumped and smiled and sulked and earnestly entreated and frowned and laughed like Brown. But the chief thing that then concerned me was the realization that, because of the doubleness of the brain itself, my response to Shoemaker's example threatened to reintroduce the same kind of reduplication problem that had already given me so much trouble with the case, which I had recently disinterred from Hobbes's *De Corpore*, of Theseus' ship.¹⁹ What if Brown's brain was *itself* divided before multiple transplantation? It was with reduplication in mind that I drafted the personal identity condition that I defended in 1967 to require the vital functions of a person who persisted through time and change to be 'subserved by a single organized parcel of matter'.

For several reasons, this was not a stable resting point. About the connexion between accepting the Shoemaker example and being a Lockean in matters of personal identity, the reader was left with the vague (and I should now say wrong and confused) impression that *either* one must reconsider one's positive response to Shoemaker *or else*, persisting in the positive response, one must subscribe to the neo-Lockean idea that memory is constitutive of identity. In the case where Brown's brain was divided and each half was supplied to the body of one or other of two twin Robinsons, the question was whether Brownson (1) and Brownson (2) were the same person as Brown. Admittedly, given the transitivity of identity, the answer was no. They could not be the same because Brownson (1) was not identical with Brownson (2). But could nothing more be said? More was certainly needed. But three years later, the 'one parcel' stipulation and the universal-particular distinction²⁰ accounted for most of what I had in store to block the very different conclusion that Derek Parfit was shortly going to arrive at in his 'Personal Identity' (*Philosophical Review*, 80 (1971)).²¹ Parfit's conclusion was that, since

¹⁹ See Wiggins (1967), pp. 33, 37, 50–2. Cf. Derek Parfit (1984), p. 254.

²⁰ See Williams (1956–7), (1960), (1970), and Mrinal Miri (1973).

²¹ I place on record here one not insignificant reminder that I mentioned in 1967 but did not emphasize. Not only does the full performance of all sorts of ordinary perceptual tasks require both halves of a brain. The two halves are not in fact equipollent. Moreover, one should hesitate before pronouncing it to be an unremarkable contingency that they are not. The halves are

Brownson (1) and Brownson (2) were not the same yet each had 'the relation to Brown that matters in normal cases of personal identity', identity itself was not what mattered in survival. He concluded that that which mattered in survival was the relation *C, or rather the new relation, still to be elucidated, that we are calling R.

From a defender of the 1967 account that I defended, which Parfit was going to upstage, the chief thing that was really needed at that time was for the defender to show how naturally and inevitably the 'one parcel' condition arose from the exigencies of understanding better what kind of thing a person or human being is. Further philosophical reflection was badly needed to explain, and explain on the basis of the sort of thing a person is or a human being is, what distinguished the relation Brownson bore to Brown, if that case was to be allowed as an identity and a true survival, from the relation that the splinters Brownson (1) and Brownson (2) bore to Brown. That is what this chapter must provide.

First, though, before entering into any of that, it will be good to make as plain as possible on what terms an adherent of Bishop Butler's criticism of John Locke might reasonably prepare to accept Sydney Shoemaker's case. It needs to be clear that, on the neo-Butlerian view, the intuitive (however provisional) willingness to allow the Shoemaker case as a case of identity creates no presumption at all in favour of I_p or any criterion like it. In concurring with Shoemaker about the case, we are not undertaking to wait for the idea of co-consciousness to be clarified or replaced by some better notion (of quasi-memory, for instance) which can figure in a condition like I_p . Nor yet are we preparing to dispense with the ontology of persons. Well or badly, we are simply thinking through our *existing* conception of the principle of activity for a human being. It is not to be denied that, if Brownson is Brown, his life story will be in one respect strange. But, on the still provisional terms

unequal because of the specialization of certain functions and most notably perhaps because one side of the brain has been colonized by structures that subserve the singular and extraordinary capacity of human beings to produce and understand syntactically ordered speech. If this capacity is as intimately connected as it appears to be with our capacity to keep track of the past, to direct our attention to our own perceptual and motor activity, to turn from the task in hand in order to think discursively, or in order to reason about what we are doing with a prospect of lighting on something else that needs to be adduced or compared, then the very thing that renders brains asymmetrical is *not* accidentally connected with that which renders the idea of a conscious self so interesting to us – and so productive of illusion. Perhaps it is almost as plausible now as it was when Aristotle wrote the first sentence of *Politics* that Nature (however much she does in vain) does nothing with no purpose. (The metaphor, as it has to be, will need to be unpacked, of course. But I shall exercise the right to be as lazy about disintentionalizing it as numerous present-day evolutionists are about the teleological language they themselves employ.)

now contemplated, it need not count as ontologically or metaphysically or epistemologically strange. For, as interpreted within the scope of our decision about Shoemaker's original case, Brownson's cognitive states need to be fully answerable in the ordinary way to all the norms by which we standardly regulate the ideas of memory, experiential memory, testimony and the rest. If Brownson claims to have planted a fig tree or scaled Vesuvius or seen the Aurora Borealis from a fishing boat off Orkney, he can only count as remembering such acts if he (Brown and Brownson) really did these things. Moreover, if he *thinks* he remembers, then that must create for him the normal if defeasible presumption that he did indeed do this planting or scaling or was indeed there in the fishing boat; and he must take responsibility for his account of what these experiences were and were like, etc.

This picture may usefully be filled out a little more. Suppose that, in a variation of Shoemaker's example, we are keeping track, eagerly and anxiously, of Brown's survival, Brown being a valued friend and colleague, mortally sick. Like every other substance, Brown is a *this such*. (This is Aristotelian jargon, but usefully resonant here.) In our watching over him tenderly and vigilantly, the *such* is taken care of by virtue of our remembering who and what it is, with what principle of activity and what particular determination of that principle, that we are to keep watch over. The *this* is subsumed by the focused specificity of the *such* that trains our concern onto the particular patient who is Brown. Here the mode of activity that is Brown's is one that is marked out, as I have said already, not merely by his possession of the standard human capacities but by his particular adaptation, development and realization of these capacities. The special and recognizable shapes and forms that he will have given at any point to the standard capacities will not of course be indelible or permanent. But, if we seek to cherish Brown and prolong *his* life, then in our concern for his particular realization of such potentialities we cannot help but feel a very special concern for the physiologically vital foundation of these things, namely the brain and/or the nervous system. Suppose then that, in some dreadful turn of events, it is planned that this foundation, the brain, should be somehow transposed to a new body. Then, in so far as we can still make sense of what is happening and keep track of it determinately (these are not conditions to take lightly) and in so far as we can regard the attempt as in any way appropriate or in keeping (a condition to be linked with the first, and not taken lightly, see below, §§15–16), our continuing concern must be for the transfer of the plurality of specialized and refined faculties

and capacities²² that are supported by *that* brain/nervous system. But, in the Shoemaker case, that *was* attended to – or so it seemed to those as impressed by the case as I was. Nor is that all. It is remarkable, once one thinks through the implications of the decision in favour of Brownson and we see the Shoemaker case in the way I am recommending, that absolutely everything that Bishop Butler could want *logically speaking* is already conceded to his objection (leaving on one side, I am saying, Butler's metaphysical conception of the soul itself). In order for the case to be seen like this, I_p is not required, nor is anything else that is very distinctive of neo-Lockeanism.²³

8. THE THESES TO BE ARGUED IN THIS CHAPTER

The conclusion to be defended in the next phase of this chapter is that Shoemaker's case marks a turning point. *If* Shoemaker's case preserves identity (and the interim answer is that it does), then it must seem to be the *ne plus ultra* of identity and survival. In so far as further thought-experiments are to be taken seriously as conceivable thought-experiments (viz. the transplanting of split brains, their carbon copying, their pantographic reproduction, the putative teletransportation of human beings . . .), I claim that the creatures or artefacts that these further developments offer to our speculations will need to be seen as analogous to clones, copies, reproductions, models – as all of a piece with things that approximate to their originals, or simply behave *as if* they were their originals. No doubt clones or models are more impressive in their own way than were any pretenders to a royal throne who were drilled and counselled in the role that they were to feign, then let loose, so prepared, upon the world. But, however that may be, there ought to be no temptation to think that clones, copies, reproductions, etc., *are* their originals – even if those charged with managing the relevant part of reality are bold enough to invite a clone to take the place of the original (just as a bereaved spouse may, without confusion of any kind, be urged in some cultures to marry the brother or sister of their

²² These are 'part of his identity', as ordinary people like to say. Or, as I might say in the language of this book, these things figure in the specific determination of the determinable principle of activity to be invoked in making judgments of identity through time between our colleague and the person the doctors or surgeons will finally confront us with.

²³ In disclaiming this affinity between the neo-Lockean view and the considerations that incline one towards the positive view of the Brown–Brownson case, I want to emphasize that I read the real John Locke's chapter as containing certain insights that every sane view of personhood and every sane view of identity will accept as such. See below, for instance, §§16–17. See also McDowell (1997).

deceased spouse).²⁴ They not only fail the 'same human being' test, by the passing of which it seems Brownson counts as Brown. And they not only fail it for Butler's reason. They also fall short of attaining to that which 'matters' in survival – or so, for reasons that will complement and extend Butler's objection, I shall aver. Finally, though, if so much can be established with any sufficiency of generality, the moment will come at last to return to a lingering doubt and the main topic of Part Two of the chapter, the doubt whether the way in which Shoemaker has Brown ~~put~~ into Brownson really does preserve Brown himself or constitute Brownson the full and proper continuator of Brown. See, in due course, §15.

9. CO-CONSCIOUSNESS AGAIN, AND QUASI-MEMORY

So much for the plan and the question whether there are possibilities that exceed the Brown–Brownson case. Let us go back now to where we were before we registered the *prima facie* persuasive force of Sydney Shoemaker's example. Let us return to the version of Butler's objection that was set out in §5.

Having found nothing helpful in C or C* as they were originally defined for purposes of the formulation of I_p, and seeing no future in a version of C that made do with 'sufficient real or apparent recollection', the defender of Parfit's view of the splinters Brownsons (1) and (2) (contrast the view stated in §7 above) now needs to invoke some mental state or capacity that closely resembles plenary or identity-involving experiential memory, yet is neutral with respect to identity. It will be this that has to make survival conceivable without identity. At the same time, though, the defender's account of this state or capacity will have to resemble the account of ordinary memory (or so it was claimed at §5, penultimate paragraph) in *disallowing* all sorts of specious or apparent memories as not really memories (of any sort).

This task was first begun by Sydney Shoemaker (however unnecessarily, according to my opinion, so far as the Brownson case is concerned)

²⁴ Note in passing the ambiguity of 'clone': it may mean one of a group of almost exactly similar things, but it may mean the whole group or type – a highly unpromising model for those who want to count some unique clone of Brown as Brown himself. And under the first head, one may mean a genetic clone or (as here) a more or less exact copy (of the same maturity and all the rest as the originating being).

I have wondered who first imported the word 'clone' into present-day philosophy. My 1967 use of it was prompted, I think, by that of G. C. Stead. See 'ΟΜΟΟΥΣΙΟΣ ΤΩ ΠΑΤΡΙ', *Studia Patristica*, 3, 78 (1961).

in continuation of the work I have cited.²⁵ But many stages beyond that, the best version of R now on offer would appear to be that of Derek Parfit. His latest version appears in a passage of his *Reasons and Persons* where he is replying to an objection against his kind of account of the phenomena that we normally describe in terms of personal memory and personal identity.

Parfit writes as follows:

The objection would be this: 'It is part of our concept of memory that we remember only *our own* experiences. The continuity of memory therefore presupposes personal identity. The same is therefore true of your Relation R [of weak co-consciousness, which is the relation that our *C first set out to be]. You claim that personal identity just consists in the holding of Relation R. This must be false if Relation R itself presupposes personal identity.'

To answer this objection, we can define a wider concept, *quasi-memory*. I have an accurate quasi-memory of a past experience if (1) I seem to remember having an experience, (2) *someone* did have this experience and (3) my apparent memory is causally dependent, in the right way, on the past experience.

On this definition, ordinary memories are a sub-class of quasi-memories. They are quasi-memories of our own past experiences. . . . [p. 220]

We do not quasi-remember other people's past experiences. But we might begin to do so. . . . Jane has agreed to have created in her brain copies of some of Paul's memory-traces. . . . One apparent memory is very clear. She seems to remember looking across the water to an island where a white Palladian church stood out brilliantly against a dark thundercloud.

Even if [Jane's] apparent memories are presented *in the first-person mode*, Jane need not assume that, if they are not delusions, they must be memories of her *own* experiences. Even if she seems to remember herself seeing forked lightning, she could justifiably conclude that she is quasi-remembering one of Paul's experiences. . . . Jane might have to work out whether it was she or Paul who had some past experience. And this might sometimes be impossible. She might have to say, 'I do vividly seem to remember hearing that tune. But I do not know whether it was I or Paul who heard that tune.' [p. 221]

Confronting the definition furnished at the beginning of this citation,²⁶ reading Parfit's condition (1) as saying that there is an experience (an

²⁵ See Shoemaker (1970), pp. 269–85.

²⁶ I call it a definition not only because Parfit does but because a definition is needed. 'Quasi-memory' is a technical term. Technical terms in philosophy are rarely scrutinized with the same care that is devoted to such questions in mathematics or logic. Yet, as logicians will point out, apparently small faults in definition can give rise in that area – and why only there? – to large and unexpected, or even paradoxical, consequences. Consider the paradoxes of Russell and Grelling – or indeed those of Richard and Berry relating to the idea of definition itself. If unexamined definitions can issue in antinomy within the sphere of logic, is there not a comparable danger that faulty definitions will sever connections in metaphysics or issue in disabled conceptions of our active and cognitive faculties in the theory of knowledge?

action or event consciously participated in) that I seem to remember, and reading his condition (2) as saying that that experience was indeed someone's or happened to someone (for other interpretations seem obviously incorrect), we ought to begin, I think, by entering onto the record the first question it invites. What is it for me, the rememberer, to remember or seem to remember *that* experience which was indeed someone's? Close similarity between how things were with the experience itself, supposing that to be independently identifiable, and how the rememberer rehearses it to himself as having been is obviously not sufficient. Countless similar but distinct events are similarly remembered. Nor yet is it necessary. Some remembering of an experience (perhaps most remembering, really) is impressionistic and inaccurate in the extreme, without being inaccurate enough to be disqualified from counting as really remembering the experience. What is it then for someone to recall *that* very experience? Can the question be answered otherwise than by a condition of the following kind: the rememberer recalls experience *x* if and only if the recollection putatively of *x* recovers to the rememberer some sufficiency of the actual details of *x* and enables the rememberer to place *x* as an event in his, the rememberer's, life and . . . ?²⁷ This sort of suggestion cannot be welcome, however, because it reimports a requirement of identity and depends on our using the rememberer's own life to identify *x*.

10. A SECOND AND THIRD QUESTION ABOUT PARFIT'S DEFINITION OF 'Q-REMEMBER'

The second question we must enter onto the record is this: ordinary memories are stipulated by Parfit to count as quasi-memories, but need quasi-memories as such, defined as Parfit defines them, partake at all of the nature of memory? Either that matters or it doesn't. If it doesn't

²⁷ Well, that is how it needs to be normally (and there needs to be a 'normally'). Abnormally (and necessarily marginally), no doubt there is recollection of events the rememberer *cannot place at all*. The 'and . . .' holds a place for any further condition of a causal or explanatory kind.

For fear of seeming to disagree with absolutely everything, I have acquiesced in the grammatical forms deployed in Parfit's proposal. But I should prefer to see them as condensations of constructions more along these lines: I quasi-remember climbing the spiral stairs of Big Ben on my sixteenth birthday just if (1) I seem to remember climbing Big Ben on my sixteenth birthday; (2) someone did climb Big Ben on their sixteenth birthday; (3) my seeming to remember climbing Big Ben on my sixteenth birthday is dependent in the right way on the event that consisted in someone's climbing Big Ben on their sixteenth birthday. Adapting the first question to this reformulation, one obtains the following: What is it for my remembering, as in (1), to relate to an independently identifiable event *x* in virtue of whose occurrence (2) holds true? What is it for my remembering, as in (1), to rehearse the details of some particular event whose occurrence makes (2) true?

matter, then Parfit loses hold of any non-accidental link between the neo-Lockean conception and the ordinary memory that is conceived in an ordinary factive manner (and whose presence in John Locke's account of these things made it persuasive). If it does matter, though, then we must worry about the following. Surely I can seem to remember some experience that happened to someone, but without actually remembering it at all. For instance, George IV was sincerely convinced, having internalized and embraced imaginatively countless well-founded reports of the battle, that he had been at Waterloo.²⁸ If so, then he could have rehearsed many of the details of the battle accurately enough and as if from a first person or combatant point of view. But George IV had better not count, simply on this basis, as quasi-remembering Waterloo.

Such questions as this are bound to collect the reply that all issues of this kind are taken care of by Parfit's requirement (3) that the apparent memory be dependent *in the right way* on the experience apparently remembered. That reply gives rise to the next doubt, however.

The next doubt is whether quasi-memory, having been vindicated as memory by virtue of being causally explainable 'in the right way', i.e. the way proper to ordinary experiential memory, can then fall short of implying the identity of experiencer and rememberer. If the memory state has to stand to the experience that it is a recollection of as the right sort of effect, then the next question is how to give an explication of 'the right sort of effect', 'the right sort of cause' or dependence in 'the right way' that will place quasi-memory in the required proximity to memory as ordinarily conceived *and* permit that non-identity. That is the third question to be put on the record. (It generates a supplementary question, but that is the subject of §12.)

One way in which a follower of Parfit might try to answer that question and manage all these doubts is to invite us to grasp the right sort of way by examples, examples that he can exhibit. But here there is a challenge. Either the examples are normal cases, in which case there is identity, and we have already seen in §5 how intimately identity is involved in the ordinary point of making a memory-claim; or else they are abnormal. If they are abnormal and there is no identity of experiencer and rememberer, then it will be inherently controversial how much they have in common with normal cases so far as remembering is concerned.²⁹

²⁸ As Anthony Flew points out. See the article cited at note 12.

²⁹ In 'The Causal Theory of Perception', *Proceedings of the Aristotelian Society*, 35 (1961), H. P. Grice gave a new authority to 'by-examples' methods of explication; but this was without any guarantee of the effectiveness of the method in cases where favour was to be gained at the same

A defender of Parfit may say that in the normal case we see an event befall someone and we take note of their capacity to recall the event afterwards. Well, in the abnormal case, this defender will say, it's just like that again *except that*, after the initial transaction that impresses itself on someone's memory, the ability to recall it is *transferred*. For instance, it may be transferred to two splinters of the original experiencer. In what manner then (we ask) is it transferred? Transferred in the right sort of way (the reply must go), and that is a way *sufficiently like* the way in which the capacity to remember is carried forward by the person who is involved in the normal way. Is it not obvious (the reply goes on) that, at least in the Brownson (1) and Brownson (2) case, what happens is sufficiently like? 'Surely, for one *prima facie* disposed to allow the Brown-Brownson case, *that* is not very controversial.'

Here and in response to the last claim, however, rather than dispute whether the Brownsons (1) and (2) case is controversial (for that is itself controversial, or so any follower of Butler will say) or insist again upon the apparent normality (for theory of knowledge purposes) of the Brown-Brownson case, I think one must make the defender stick to the point. A stipulation *de novo* was being offered of 'quasi-remember'. That stipulation brought with it the commitment to say something general about what 'the right way' was. The answer proposed was 'a way sufficiently like the way that is manifest in the normal case'. But, so far, this has only been explicated or justified on the basis of one or two cases. Our difficulty was that the normal case is the identity-involving case. What was needed then were *general* reasons, not reasons relating to one or two (or even three) particular cases, why identity of experiencer and rememberer was irrelevant to normality. Either that was needed or else (much better and more correctly) we needed a general explanation of 'right way' that was both plausible and identity free.

How then is one to move forward? No progress seems to be possible here until theorists of quasi-memory and the R-relation that it is meant

footnote 29 (cont.)

time for a reconstrual of the phenomenon being explicated. Look towards the end of the citation from Parfit given in §9. Ask whether one could pick up from putatively paradigmatic cases of Jane's sort (where memories of an event may seem rather like pieces of shrapnel left over from an explosion) the mode of dependence on its associated event or experience that is required of a mental state that is to count as partaking of the nature of memory. What ideas will be at our disposal for us to focus on the relevant and requisite aspects of such examples if they are included among the paradigms intended to fix the extension and intension of 'causally depends in the right way'? Well, evidently, the ideas of *experience*, *having an experience* and *seeming to remember*. But in availing ourselves of the last we must of course prescind from all elements of *remember* itself that the definition of Q-memory will seek to dispense with. 'By examples' is not the straightforward business it is often taken for. It is certainly not straightforward here.

to define – I shall sometimes call them the Q-theorists – measure up to the requirements of full generality in definition-giving. Equally, though, and by parity, let it be conceded that no progress seems possible until those who are as sceptical as I am about these definitional efforts furnish a better account of our own reading of the semantical and epistemological *point* of attributions of experiential memory. We must clarify the connexion between this point and questions of identity.

II. DIGRESSION: AN ALTERNATIVE METHOD OF DEFINITION, REVEALING BY ITS INADEQUACY THE SEMANTICAL POINT OF THE ATTRIBUTION OF EXPERIENTIAL MEMORY

On the terms that I propose, there are various ways in which the Q-theorists might advance. I should need to shadow all of them. But at this stage in the argument, a digression may prove to be a short-cut. We shall understand better the utility and role of Parfit's definition *and*, as a bonus, approach closer to the semantical point of the attribution of experiential memory if we ask why Parfit did not offer a simpler proposal. The simpler proposal is that X quasi-remembers walking from St Paul's to Trafalgar Square (say) just if it is for X exactly as if he remembered walking so *except that* it may not have been X who did the walking in question. Or (better) X quasi-remembers walking from St Paul's to Trafalgar Square just if there holds, under the headings of causality and whatever else, everything that is required for X to remember walking from St Paul's to Trafalgar Square *except possibly this*: X's being identical with the person who did the walking.

Under this proposal, the idea would be that quasi-remembering ϕ -ing is X's remembering ϕ -ing *less* X's being identical with one who ϕ -ed. Quasi-remembering is remembering with or without identity of experiencer and rememberer. Of course, if (as most people suppose) one's in the ordinary way remembering walking from St Paul's to Trafalgar Square entails that one did walk so oneself, then the remainder that results from the subtraction from ordinary memory must not be reckoned under the scheme of X's remembering ϕ -ing and X's *not* being identical with the one who did this ϕ -ing. But we have already seen an obvious alternative, namely to say that everything is for X as if he remembered walking from St Paul's to Trafalgar Square *except that*, in so far as that remembering is held to require X's identity with the walker, X may fall short of what is required.

To show that this will not do, a critic on my side of the argument will

begin by protesting how easily this kind of 'except in so far as' can have the effect of producing a description that is indeterminate for sense. Let the Q-theorist consider the near vacuity of 'It's as if there were a thunderstorm in progress except, maybe, in so far as that requires something to be happening in the sky', offered as a definition of 'quasi-thunderstorm'.

Here a Q-theorist who saw some future in the simple proposal that Parfit didn't offer would surely retort with the challenge to prove that this was a *fair* parody of his definition of 'quasi-remember'.³⁰ What then? Before rising to the challenge to prove anything, I think the critic should point out that one could never demonstrate in the way by which Parfit's defender insists it be demonstrated that 'quasi-remember' is like 'quasi-thunderstorm' or that it is unlike. For suppose, as is usually the case with real words for non-artificial kinds or acts, there exists no analytic definition at all of 'remember'.³¹ Then any faultiness (vacuity, indeterminacy etc) in 'quasi-remember' might not be demonstrable by a strictly analytical procedure – just as its soundness would not either. But not only that, the critic ought to say. If there is no definition of 'remember', then there is no future in the project of defining 'quasi-remember' by subtraction from 'remember'. For, in the absence of an analytical definition, there is no question of arriving at the thing needed by a procedure of deletion.³²

³⁰ Another thing he might do is to claim that, as defined, a quasi-thunderstorm might be blamed for turning the milk sour. But this scarcely restores 'quasi-thunderstorm' to any parity with 'thunderstorm', e.g. as the name of a causal-cum-explanatory kind.

³¹ In *Tractatus Logico-Philosophicus* (see especially 5.451), Wittgenstein sees it as a precondition of the definiteness of the sense of an expression that it should have some unique exhaustive decomposition into semantically independent indefinables. Wittgenstein assumes that among senses there operates something analogous to the Axiom of Regularity (*Grundierungsaxiom*), which excludes in the world of sets the existence of an endless downward sequence from a given set A in the form: ... (DeC), (CeB), (BeA). But against that one will want to object, with the later Wittgenstein, that grasping the sense of an expression cannot, in the normal or basic case, be the learning of a definition – or the as-if learning of one.

³² This is the only serious thing that could be meant by the project of subtracting the concept of identity from that of experiential memory. If someone starts on the process of definition, someone else can, it is true, start on the process of deletion/subtraction. But how can the person deleting be sure he has finished? How can one ever say the new definition is completed?

Someone might say that 'swift' (the bird) has no definition but, all the same, one could define a new term 'martlet' by saying that a martlet was a swift without any feet. What would be wrong? Well as a recipe for painting heraldic devices, that would be all right. But as the definition of a possibly non-imaginary bird? Even in the case of a bird like the swift, indefinitely many real things about it may presuppose feet. Even if the martlet mated on the wing, could it hatch or feed its young on the wing? If not, how does the martlet perpetuate itself? Is the martlet in the heraldic image a well-defined kind of bird at all then? The grammatical correctness of 'X is a kind of swift that lacks feet' or 'X remembers Y's experience except that X may not be identical with Y' proves nothing.

In his inaugural lecture 'Meaning and Truth', *Logico-Linguistic Papers* (London, 1971), P. F. Strawson spoke of the 'bogus arithmetic of concepts'. For a development of that thought, see §3 of Jennifer Hornsby, 'Feminism in Philosophy of Language', in *The Cambridge Companion to Feminism in Philosophy*, ed. M. Fricker and J. Hornsby (Cambridge, 2000).

Defining a new term by the deleting of all uses of the identity concept from the definition of an old term would be an honest and straightforward enough method of introducing the new term. But there is no question of making such deletions from a definition whose actual text can never be submitted to the blue pencil.

This response to his defensive challenge will encourage the Q-theorist to issue another. If there is no lexical definition of 'remember', then how does 'remember' have its semantic identity and how are we to be assured of its soundness? Will the critic please explain himself?

The critic's reply must be that 'remember' has meaning in the way in which most non-technical words do – not by virtue of introducing a concept built up from simpler or more transparent concepts, but by virtue of use. 'Remember' has its meaning (he may say) by standing for a concept with a distinctive place in a whole skein of coeval concepts that came of age together, developed together and now work together to some present purpose. The work of these concepts in the case of memory is to make it possible for us to undertake the familiar but multiple operations and transactions by which we sustain and extend the economy of knowledge and awareness. What is more, that work progresses. Given this picture and given that there is no question of laying out the definition of 'remember' and scoring out every component that imports '=', the whole question between definer and critic must shift (according to the critic) to a different and more interesting place: *could* a species of memory that was identity free play the role that is played in the epistemic economy by ordinary memory?

One major role of experiential memory, the critic will now say, is that it supplies unmediated (albeit fallible) information that *one can take oneself to have got on the basis of one's own experience in the past*. (Is not the role of experience analogous to that which was assigned to gold or silver in older pictures of the real economy?) If no way is manifest for quasi-memory to do that, then the attempt to define 'quasi-remember' by 'everything is as if ... except possibly in so far as ...' will inevitably subvert the very thing that sustains the signification and point of a concept such as that of memory. Any serviceable or intelligible notion of quasi-memory that depends on the given meaning of 'remember' needs to respect the links that connect experience with knowledge. And these involve identity.³³

Parfit's own proposal is still pending, still held in reserve, remember.

³³ 'Des expressions et formes qui ont un sens dans l'action vraie, sont excitées et employées hors de leur domaine – et perdent pied. Mais le philosophe ne s'en aperçoit pas': Paul Valéry, *Cahiers I*, ed. J. Robinson (Paris, 1973), p. 753.

But, since a real difficulty appears to lurk here for all accounts of quasi-memory, it will be best for the critic's line of argument to be filled out before we return to Parfit's proposal.

Suppose that our use of 'remember' depends on our participation in the living use of the verb phrase 'X remembers (-)ing'. If it is among the expectations conditioning this use of the word that the epistemological role of experiential remembering should be to help provide us with a starting place for the attempt (by further inquiry, in partial reliance on the testimony of others, etc) to extend our knowledge about how things are in the world,³⁴ then the starting place need not be something philosophically indubitable or infallible. The real charm of epistemological foundationalism, once we forget considerations about immunity from error, was rather this: that it saw that, if there is to be knowledge, there have to be some *unmediated* starting points. Among these there must surely be states of perceiving and remembering, including personal or experiential remembering. But if that is right, then the first thing that matters for the theory of knowledge is that perceptions and memories should be accorded by that theory their full presumptive title to count as the place from whence, however fallibly or defeasibly, we begin. For purposes of the pretensions of quasi-memory to play a role analogous to memory, it matters that quasi-memory should be a state of mind or faculty that not merely permits an exercise of it to count as a direct source of knowledge but *constitutes* it as a place to begin. But how is that possible (one wonders) if there are doubts about how Q-memory places the experiences it relates to, or if quasi-memory does not make the rememberer himself responsible either for the correctness of the recollection of the experience that is to be a starting point or for the reliability of the perceptions made at the time of the experience? (Q-theorists may appeal at this point to testimony. But our handling of testimony depends on our proper deployment of our own experience. Nor would

³⁴ The same goes for the role of direct perception and one's memory of what one believes. In the argument of these pages, I follow the argument of my article, 'Remembering Directly' (1992). That argument, I should record, is an attempt to develop in one particular direction one of the points made by Evans, *The Varieties of Reference* (Oxford, 1984), pp. 241ff. (I cannot understand Parfit's answer to Evans, given in his (1984), p. 516.) How dare I, having (in effect) criticized Q-memory theorists for not demonstrating the soundness of their definitions, offer non-conclusive arguments, all founded in the suspicion that the sense of the word 'remember' is at once indefinable and *sound*? I only dare because the dialectical positions are not *exactly* symmetrical. First, we scarcely know of any correct analytical definitions. So the most we have is the *prospect* of a correct definition of Q-memory. Secondly, 'Q-remember' is the neologism. We ought to refuse to use it till it is given a determinate use. In the case of 'remember' it has a determinate use. The conviction is at once defeasible and indispensable to us.

testimony itself be exactly what it now is once quasi-memory was counted among the sources of the testimony of others.)

Suppose it has started to rain and I need to be sure that, if I am going to go out later, then I shall not get wet. Then I need to be able to take it as a perceptual datum that I see that it is raining, and as a memory datum that I have an umbrella. If the question arises in my mind 'Am I sure I still have an umbrella?', then I need to be able to summon another memory, such as the recollection of having seen it very recently hanging on a peg. I need to be able to say from personal memory not only 'there are pegs in the hall and by the back door' but also 'I remember seeing the umbrella there yesterday night.' From perception and memory and memory of coming in last night, etc., I need to be able to say: 'I'm at home. This is the back door through which I entered. So the umbrella must now be hanging in the corridor at the back.'

The trouble with such a train of thoughts in the form in which it would need to be redeployed under the Q-modified conception of memory is that the idea of an umbrella-check effortlessly organized by someone's idea of his own single-tracked unitary life becomes problematic in a quite new way – precisely as we see it doing in the long citation from Parfit given in §9. We can dispense with infallibility. What we cannot dispense with is the organization of such thoughts by starting points consisting of the unmediated presentations of memory and perception and the accreditation of such presentations by their provenance *as* unmediated. That is not all. To know how to start upon any future collation of evidences I shall need to know which findings had that direct provenance. If I am remembering in the normal fashion and I can expect to *place* that which I remember, then I shall know, as and when I need, what explanation to seek of the relevant perceptual and memory findings, both the correct ones and the mistaken. I can weigh how compelling each such finding is and try to put a body of them together into a coherent narrative that accommodates together and explains together their several verisimilitudes or falsisimilitudes. But how else can they be placed so *except in a partial narrative of my life*? For thoughts such as these, identity is *not* eliminable.

The quasi-memory theorist may react to these difficulties by suggesting that the normal truth-condition for remembering be modified in such a way that all I need to be able to claim is 'someone in an R-relation to me saw the umbrella hanging on a peg'. If the Q-theorist is sticking to his brief, however, then he has no title to that explanation. Even if he can whitewash the 'me', he cannot appeal to the relation R in the

explanation of 'quasi-remember'. For quasi-memory is one of the things he needs to use in order to *construct* the definition of R.

12. MORE ABOUT 'DEPENDENT IN THE RIGHT WAY'

That is the end of the digression. I hope it helps to discharge my obligations under the accord proposed at the end of §10. From the considerations rehearsed one has a lively sense of how wise it was for Parfit to define 'quasi-remember' positively and outright rather than by an attempt at subtraction. One sees how wise it was for him to invoke ordinary 'remember' itself (in the existing sense) only obliquely (as in his condition (1)). But one has some sense too, augmenting one's awareness of the open questions still left on the record (§§9–10), of the difficulty to be surmounted in any definition of quasi-memory that will respect the linkage we need to assume between experience, memory and empirical knowledge. At this point then, in advance of drawing any further conclusion, let us draw breath, and recapitulate the state of the argument about Parfit's definition of 'quasi-remember'.

In §§8–9, after it seemed that an identity-free concept of 'remember' was needed and it was furnished by Parfit, two questions were raised about the definition of quasi-memory quoted from *Reasons and Persons*; a third then arose concerning the phrase 'in the right way' as that figures in Parfit's condition (3). Condition (3) reads 'my apparent memory is causally dependent, in the right way, on that past experience'. In §11, the suspicion was voiced that identity of rememberer and experiencer, so far from being cancellable, is in fact integral to experiential remembering as integral as it is to knowing *what* you're remembering. This last is something you normally do by exercising your presumptive right to place the experience apparently remembered within your own life.³⁵ It seems to presuppose the very thing that is still in question.

The time has come to advance to something new, if only to a new sort of question about Parfit's positive proposal. Here is the new question,

³⁵ A Parfitian might respond to these difficulties by allowing that I have given a correct account of how things must be subjectively speaking with the one who remembers. The Parfitian might allow that the rememberer's thoughts must indeed be as I have said. But that does not mean that experiential memory as such requires it. My reply to this would be multiple and would ramify, no doubt, in reply to the Q-theorist's replies to my replies. But the first point is that there is nothing as yet in the definition of quasi-memory that answers to the need for one who is remembering and seeking knowledge to organize his thoughts in the sequence suggested at the end of §11. For it is not obvious that anything in the Q-memorist's conception of memory can furnish the requisite materials for starting points that can be fortified as starting points by being addressed back to the course of the rememberer's life.

Does Parfit's condition (3) seek to discharge a responsibility to say how memory is to be distinguished (as indeed it needs to be distinguished) from the other mental states such as conceiving or fancy or imagination or hope or intention . . . which present objects or events to us? If so, does 'right way' require the operation of a certain distinctive mode of mental functioning? Or does (3) address a normative question, one also needing attention if quasi-memory is to subserve the growth of knowledge in the way in which memory does, a question about the goodness or badness of the fit between how the experience really was and how the recollection presents the experience as having been?

Interpreted exclusively in either of these ways, as concerned either with mode of functioning or with norm, condition (3) appears to be insufficient to constitute quasi-memory with the very same nature, epistemologically speaking, as normal memory. It seems likely then that condition (3) is intended to address both issues at once. Perhaps the thought behind condition (3) is that what quasi-remembering requires is that quasi-rehearsing a thing depend causally on some mode of mental functioning or a mechanism, but a mode or mechanism which *itself* satisfies some sort of normative requirement. Perhaps the normative requirement is this: that, for a wide variety of ϕ , the mode or mechanism should bring it about that the experience is presented afterwards as having been ϕ if and only if it really was ϕ . For instance, the mechanism should bring it about that my memory of climbing the spiral stairs of Big Ben is of a climbing, is of a climbing in the heat, is of a climbing something vertiginous, if and only if the experience itself was one of climbing, in the heat, something vertiginous . . . etc.

Read along these lines (let us not pause to make the choices needed in order to complete the task), Parfit's condition (3) certainly comes to life. But how well does the condition then provide for anything analogous to the distinction between better and worse (normal, identity-involving) experiential memory? There is no escape from the worry that the normative question of how *well or badly* someone's memory mode or mechanism has functioned or functions needs to be distinguished from the question of how satisfactorily the mode or mechanism involved approximates to a *memory* mechanism. Can condition (3) admit *incomplete* or *imperfect* or *partially wrong* or *oddly produced* memories as quasi-memories? If quasi-memory and memory are to be generically related, it needs to accommodate all of these. If the putative relation is to have any of the properties that first motivated the search for something like quasi-memory, then the question is important.

At this point we will turn again to Parfit's definition as it lies on page, and then at last (with a start perhaps) we must wake up. The definition we see that Parfit presents there, once we attend properly to the passage I have quoted in §9, is not a definition of 'quasi-remember' or 'quasi-memory' at all. It's a definition (he himself announces that it is a definition) of 'have an accurate quasi-memory'. Inaccurate quasi-memory is not provided for.

This is not a tiny oversight, but the trace of a major philosophical difficulty. Nor is it hard to guess how the definition in front of us came into being. Putting together the two distinct issues of mode of mental functioning or mechanism and of norms of correctness, but being aware of the extent to which his condition (3) screwed things down too tightly Parfit must have seen that he had to 'balance' his proposed equation (equivalence). Then he found (or so I surmise) that the only way of doing that was to add something to the left-hand side, rather than to remove an unwanted surplus on the right. So the definiendum was changed in the way we see in front of us.

Here one cannot help but sympathize. For it is simply not obvious how to lighten the right-hand side, or what to remove. The trouble is, though, that, as a technical term newly defined, 'accurate-quasi-memory' has now to be seen as a (so to speak) *fused* term. The components appear to have independent meanings, but, when they are taken as combining to mint a new technical term, no separate meaning has been given to them. Under this mode of definition, nothing Q-analogous to inaccurate ordinary memory could be provided for.

To dwell, as we do so often, on examples of highly accurate and otherwise perfectly ordinary remembering tempts us (I conclude) into elision. Such examples distract our attention from the difference between questions of mental mode/mechanism and questions of correctness.³⁶ So soon as these issues are allowed to coalesce, as they seem to have done in Parfit's definition, quasi-memory leaves the epistemological economy with a marked deficit.

In the first place, if quasi-memory excludes inaccurate memory and/or excludes variants on the actual mental modes or mechanisms of memory, there is already a serious problem. For the growth of knowledge by criticism, refinement and correction really cannot dispense with

³⁶ It may seem that I believe I have shown anti-naturalistically that these questions are utterly independent. No. The mechanism is no doubt selected as one for which there can be a norm. The operation of the norm is conditioned by the mechanism for which it is the norm. None of this is denied by one who simply insists as I am doing on the distinctness of distinct questions.

either of these possibilities. The relation R that Q-theorists want to define will not then be built on memory at all.

In the second place, if quasi-memory is to be akin to memory at all or to count as a direct source of knowledge about that to which it pertains, then there needs to be room for the whole purport of a given act of quasi-recollection – room for its pretensions qua memory state *as well as* room for that which it purports to present and room for how it presents that thing as having been – to be determinable independently of questions of the accuracy of the recollection and the goodness of its fit with the experience. Condition (3), taken as it stands, makes this separation problematical for Q-epistemology.

It may seem that the difficulties that arise from the conflation I allege of questions of mode or mechanism and questions of correctness are imported by the contingent particularities of Parfit's definition. But that style of definition arose out of his need to purify the memory concept of identity. It was indispensable to the campaign to show that identity was irrelevant to the thing which mattered constitutively in Brown's survival in the shape of Brownson.³⁷

PART TWO

13. AS IT NOW APPEARS, THE STATE OF THE WHOLE ARGUMENT TO DATE

Here is a summation of what has been claimed so far in this chapter: for personhood as we know it, the identity of persons coincides (I began by

³⁷ There is a special difficulty in disentangling memory from identity. Unlike a perception, whose occurrence ties it to what was there to be perceived and whose correctness can in principle be regulated from other investigations of what was there to be perceived at the time and place of the perceiving, and unlike a portrait whose title and original provenance ties it to some sitter to whose appearance it is answerable, the act of recollecting an experience, the event *e* to which the recollection is answerable for its correctness, does not permit the identification of its content or referent, namely *e*, to be made on the simple basis of the place or time of the occurrence of the act of recollection. The place and time of doing the act of recollection afford no indication whatever.

On the inner view of memory, the importance of this comes out in the need (already mentioned) for the rememberer to place the thing remembered, the experience, and thereby signal the correctness condition of the memory-state in question, by reference to something in the sequence of his own life. He can only do this sort of placing or fixing of *e* by engaging in the kind of thought that the definition of quasi-memory does nothing to make sense of and that is problematical for Brownson's (1) and (2).

On the outer view of memory, the same point comes out first in the need to ascertain from the subject what memory presents to him (see the inner view) and secondly in the need to identify by reference to the outer life of the rememberer the very experience – wherever there is one – to which the recollective act or memory state is answerable for its correctness.

suggesting) with the identity of human beings. Human beings are substances possessed of a specific principle of activity to which, in the course of a life, each one of us gives his own yet more specific, more and more distinctive, determination. Prominent among the specifically human activities is our exercise of the cognitive faculties. Faced with Sydney Shoemaker's Brown-Brownson case, our provisional first finding was in favour of the identity of Brown and Brownson, because Brownson appeared to be the determinately traceable functional continuator of Brown (we said) and Brownson seemed to inherit (in the manner in which any ordinary person who has suffered no such adventures is constantly inheriting from himself) the perfected epistemic and other capacities of Brown. He carries these forward through time, together (we assumed) with Brown's other skills and abilities. Moreover, we found on further examination that the judgment that Brownson is Brown lies well outside the reach of Butler's objection to Locke. If there are difficulties with the Brown-Brownson case, they do not reside here. Neither for purposes of Brownson's thoughts, nor yet for the purposes of describing in third person mode (or in philosophical mode) the set-up that includes those Brownson thoughts, is there the slightest need to try to construct the identity free notion of experiential memory. Nothing appears to prevent us from thinking of Brownson as having full cognitive responsibility for the claims that he makes from direct personal memory.

That summarizes only what happened up to §8. Now for the rest. Consider the case where we have not Shoemaker's Brownson but two splinters, Brownson (1) and Brownson (2), resulting from the transplanting of the two halves of Brownson's brain into bodies of twin Robinsons (each found available, we are to suppose, in debrained condition). Here, we allowed that we may want to say that it is *as if* Brownson (1) and Brownson (2) remember. But we declared that, on the terrain lying beyond the case Shoemaker introduced, no stronger claim ought to be allowed than *it's as if*. For neither Brownson (1) nor Brownson (2) is the same as Brown, and there is no newly minted, properly defined *remembering-of-experiences* relation (or so it was argued) in which Brownsons (1) and (2) can stand to Brown. (That was the conclusion of §§9-12.) No new memory concept that was modelled on the notion of memory that once commended Locke's discussion to us ought to try to make room for personal or experiential remembering in the case of Brownsons (1) and (2). Friends of quasi-remembering seek to improve on the finding that in various ways it is for Brownsons (1) and (2) as if each of them were

Brown. But in so far as quasi-remembering gets us beyond that anodyne judgment into something that is stateable *without* the use of the 'as if', it brings nothing but conceptual disruption.³⁸ 'Quasi-remember' is ill-defined; and in its application to Brownsons (1) and (2), a confluence appears of two things that scarcely mix, the idea of a person as a singular thing with an individual biography and the idea of a person as a quasi-universal, susceptible of multiple instantiation.

14. PARTICIPATION IN THE GROWTH OF KNOWLEDGE

The force of the considerations mustered so far would be blunted if there were some way to contend that Brownsons (1) and (2) might after all participate in the ordinary processes by which empirical knowledge is expanded. Before we go further it needs to be enquired (even at risk of doing again work done already) what prospects there are of its being shown that, despite the fact that their 'earlier lives' are shared, Brownsons (1) and (2) could *in their own way* keep track of things they live through, mark 'memories' by their provenance, and place the experiences that 'memories' record by the position these experiences once occupied in their lives, etc.

I begin on this by saying that the whole case for examining this possibility may seem to depend upon the prospective availability of something like the relation R. If so, let us note how easily and unthinkingly we can relapse into the supposition that R is available already. Perhaps this is because we are bewitched (as I have said) by the putatively sufficient condition that is given such prominence by the hackneyed, familiar but special examples where the norm and mode/mechanism constraints are perfectly and simultaneously satisfied. These happy cases furnish no understanding of the necessary conditions of the presence of the R relation. All we really know about R we learn from purported cases of it and the role that R is meant to play – in philosophy. But that measure of understanding cannot (I have claimed) assure us that there really is a relation that will both perform that office *and* be independent of identity. Indeed, the rationale began to appear some sections ago for insisting upon some residue of the 'one parcel' condition mentioned in §7.

³⁸ At this point in the argument it may be desirable to point out that, for one who agrees to discuss these thought experiments but is drawn to the line I am still taking at this point in the argument, it is not at all necessary to deny that the splinters Brownson (1) and Brownson (2) *have mental states*. It is not even necessary to deny that any of the respect owed to Brown should be extended, where they actually exist, to Brownsons (1) and (2), however pathetic they seem. They are after all repositories of a kind. See §14.

In the second place, it is worth trying to think further about the epistemological condition of splinters such as Brownsons (1) and (2). In so far as they proceed as if they were direct rememberers, and in so far as they take their experiential memories as direct presentations to them of their own past, they must become party (we have said) to indefinite quantities of error. Each of them says that he remembers a momentous one-to-one dialogue with Isabel (say) or a confidential man-to-man transaction with Stephen. But there are two Brownsons and neither of them is the same as the substance that was there in converse with Isabel or Stephen. Each survivor thinks of himself as a single person with cognitive faculties, and a life altogether of his own, but such thoughts are just wrong. Again, each of the Brownsons thinks that he is a person who lives a life, a substance. He cannot be what he thinks he is. Moreover, if experiential memory is to work as it must work for purposes of cognition, there is no thinkable alternative but for it to induce in Brownsons (1) and (2) the sorts of beliefs that unsplintered people have. Nevertheless, the beliefs that Brownsons (1) and (2) need to have are beliefs that *cannot* all be true. For the sake of arranging for all this to seem possible, why on earth should anyone want to pretend that R is well defined when it is not?

Paul Snowdon has helped me to answer this question. The point is that, even after issuing all these denials about Brownsons (1) and (2) and their supposed experiential memories of Brown's life, the human being theorist is still faced with some impressive *apparent* memories. Whatever else one says, it is hard to deny that each Brownson represents a repository of memory traces or traces of memory traces relating to events that took place before either of them existed. So be it. We can and maybe we must treat them as repositories. It is as if they had each been there. There is no more one should say, however. In Parfit's example, Jane might be the sole remaining repository of information about the appearance of the portico of some remote island church destroyed, since Paul's death, by flood. Nevertheless, despite any gratitude we may feel for the existence of this vestige, let us be clear that Jane's state of being such a repository falls far short of the cognitive condition of one who displays her memories in cognitive engagement, takes responsibility for them, places them, and exercises an option to accord them the status of ordinary direct memory. For the same reason, the attainment of repository status by Brownsons (1) and (2) falls short of that which would have mattered in the survival of Brown.

A third remark. By practice, or even without practice, one can fall

readily into the frame of mind where, having been impressed by the Brown-Brownson case, one finds it impossible to think that the Brownsons (1) and (2) case represents the *passing away* of Brown. The explanation is not far to seek. At this point in the supposed history of Brown, one very easily falls into thinking of Brown as a thing that persists in Brownson (1) and Brownson (2). One conceives of Brown as a thing that persists in its/his instantiations, a thing that is wherever they are – just as the sail that lies over you, over me, and over a friend of ours, is where I am, where you are, and where he is. (See Plato, *Parmenides*, 131b.) In short, one thinks of Brown as a concrete universal. Nothing need be wrong with that. Indeed it is very likely that, with a little care, one could invent a contradiction-free way of speaking of such a corporate being, not merely as a thing with a past but as one with a future. One could index the thoughts that are ascribed to the corporate being by reference to the members or constituents that give local hospitality to the thought, even as other members or constituents lack it or have another, possibly conflicting, thought. Under the new convention, the corporate being Brown will be large, but within him he may safely embrace contradictions. Moreover, if we are careful, we shall be able to say this as theorists without ourselves contradicting ourselves. Nevertheless, Brown reconceived as such a concrete universal is not the sort of thing whose survival was to have been described in the case of Brownsons (1) and (2). Nor is this how we conceive of subjects of experience when posing the question of personal identity. The problem relates to subjects who have lives of their own, to potential authors of autobiographies, who can say 'I think. . .', rather than 'In respect of member *m*, I think. . .'

Let it be clear that the difficulty here is not that metaphysics needs to look askance at corporate beings. Perhaps departments of physics or psychology or philosophy could conceive of themselves as such. By an effort of imagination or of memory, one might find even now some way of thinking of such departments as participating not merely in power or glory (high ratings) but in the growth of knowledge or understanding. But there is an important difference here from the sort of thing which Parfitians might think of as surviving or quasi-surviving. For, as we have said, Brown thinks of himself not as corporate and *not* as composed of splinters. If a concrete universal Brown thought of himself in (Shoemaker's) Brown's way, he would be wrong. Departments of physics, on the other hand, considered as cognitive beings that do not need to conceive of themselves as *not* having members or as *not* multiple,

rather thinking of themselves as pursuing the truth about physics along the multiple paths pursued by different individuals or teams, must see it as their business to push physics forward by promoting individuals' and teams' constant effort, in seminars, in the laboratory or the canteen, to explain anomaly and transcend inconsistency. A department of physics has no need to suffer from the delusion that it is itself an individual person seeking after truth. Its corporate thought had better be that it is what it is, namely a corporate thing in search of the truth about physics by dint of the efforts of its members. Not only that. A corporate being such as this can offer a perfectly ordinary account of what sort of thing its members are. Suppose a correspondingly corporate notion of Brown were concocted, what then? Then corporate Brown would be conceived as a person with members which could see itself as a person with members – and could say what sort of thing any one of its members was. Corporate Brown is not then the sort of thing to figure in an autobiography, a story of itself as one individual person. Parfit's account of survival is not, however, designed as an account of the survival of a corporate sort of thing. It is an account of the survival of things on the ontological level not of corporate Brown but of the *members* of corporate Brown.³⁹

Epistemology has long needed to engage with the collective nature of our cognitive labours, with knowledge by hearsay, with the role of testimony, with the conditions for the accumulation and collective criticism and emendation of knowledge. Yes. But no sane enthusiasts for these tasks will seek to dispense with the hard-won insights of the empiricist epistemology of experience, eyewitness and recollection by individuals with lives and histories of their own. Not even an underlabourer can be a splinter like Brownson (1) or (2). The charm of neo-Lockeanism was that it seemed to engage with our given notion of memory – of personal memory, that is. When defended by all the necessary philosophical refinements, it gradually disengages with that notion and puts into question the only account we know how to give of the evidential fabric of our knowledge. Not only does its revised conception of memory denature memory. It has now put itself in danger of disengaging altogether with the idea of an individual 'self' which was its starting point.

³⁹ If a Parfitian person were corporate, what account would it offer of its own members? Does each member itself have members? In case there is a reader who wants to pursue this question and in case that may be useful, he will find an informal statement of the *Fundierungssaxiom* in note 31.

15. THE PENULTIMATE PROBLEM AND A VERDICT UPON IT, ALL LEADING IN DUE COURSE TO A REASSESSMENT OF THE ORIGINAL SHOEMAKER CASE⁴⁰

In the philosophical transition that will carry us slowly back to our starting point, viz. the Shoemaker case in its original state, there remains one other case I have to comment upon, namely the case where the brain of Brown is split but only one of the two portions is successfully transplanted into a Robinson twin. The other dies, we may suppose. Let us call the survivor of these events Brownson Sole. In the circumstances (it will be asked), how can we refuse to treat Brownson Sole as Brown himself? If Brownson Sole claims to have seen the Aurora Borealis from a fishing boat off Orkney on which he was the only US national, if he gets the numerous details of Brown's adventure right and this is no fluke or accident, then how can we help but treat Brownson Sole as one who was indeed there? It may seem that the only thing that is left to ask is how well he remembers seeing the Aurora Borealis.

We can go as far as we like with the *as if* (I reply) and we can take the *as if* as lightly or as seriously as we wish to take it. Indeed we can take the *as if* fully seriously enough to experiment with questions put to a being that we count only as a 'repository' or the carrier of traces of some event they were not present at. When someone devises a whole philosophy of *als ob*, that can be taken seriously too. But consider the judgment that Brownson Sole is the same as Brown. If it is true, it ought to depend only upon Brown and Brownson Sole. It ought not to depend on an assurance about the existence or non-existence of another thing altogether besides Brown and Brownson Sole. It cannot depend *ad hoc* on the accomplished death or destruction or non-viability of the other half-brain that was poised to animate some rival to Brownson Sole. Any verdict about Brown and Brownson Sole issued in such dependence would infringe the *Only a and b* rule, **D(x)** of Chapter Three, §4, which is founded on **D(vi)** of Chapter Two, §7 (*ad fin.*) and the sortalist cum conceptualist argument for **D(vi)**. The argument there was all of a piece, moreover, with our commitment to the distinction between a particular and a universal. (See above, §14, the paragraph beginning with the words, 'A third remark'.)

At this point, it is easy to imagine that someone may say that the *Only a and b* rule looks plausible enough as an abstract logical requirement,

⁴⁰ Readers following the suggestion in the Preface that §13 following should stand as a conclusion for the book will probably be familiar enough with present philosophical trends in thinking about personal identity to guess immediately what has been in progress here. If not, §1 and §6 of this chapter, with or without §8, will supplement §13.

but ask how the normal theory and practice of individuation as we have described it could ever underwrite its satisfaction: 'In practice, in the actual business of making identity judgments, the *Only a and b* rule is unworkable.' My answer to that runs as follows: if Brownson Sole were the same as Brown, where Brown is a human being and thereby a cognitive being, then Brownson would have to be the same human being and the same cognitive being as Brown. But Brownson Sole came into being from Brown *not* in a manner constitutively sufficient to preserve the transitivity of identity or to perpetuate the activity of a human being as a cognitive being. (See §9 following.) On *these* terms, there is no question of Brownson Sole's being the same as Brown. The *Only a and b* rule is not unworkable at all. It asks no more of us than that, when we address questions of identity, we should persist in a way of thinking that preserves the distinctiveness of identity and marks the difference between singular and universal. If identity itself is a matter of indifference to you, of course you can ignore the rule. (Cf. Chapter Three, §4, paragraph 9.) In that case, your judgments will not differentiate identity from resembling, succeeding, going proxy for, being a replica of . . . , and you don't care. If you do care, don't ignore the rule.

16. BROWN-BROWNSON RECONSIDERED

There is a point here which, if it deserves anything at all, deserves to be followed through completely. The principle we have just invoked is that, since our judgments of identity need to be informed by consideration of the kind of things that are in question and thus by the principle of activity of these, our positive judgments of sameness are answerable (by their nature) to a norm that requires anyone who makes a judgment of identity through change to assure himself whether the change is one that preserves this principle of activity – and *equally* requires him to withhold the verdict of identity in any case where the change that is in question would, in *other* cases that exemplified the same process, *fail* to preserve that principle of activity.

Once such a principle is announced, it will seem that in our application of it the question ought constantly and repeatedly to arise: 'what counts here and in this or that connexion as the same process?' One will expect that a constant need to engage with such questions will make us party to an elaborate system of casuistry by which the specifics of each ongoing process can be interrogated in the light of the formal properties of the identity relation and of the character of the thing-kind whose

members are subject to the process. One will expect then that anyone making judgments of identity will have to put himself constantly into the frame of mind of a common law barrister, as if testing his every thought against previous rulings or precedents. But as an expectation relating to the world that we know, doesn't this seem extraordinarily implausible?

If there is a mystery here, I think it will disappear with the reflection that the individuating norms by which we normally proceed so effortlessly (but which it seems we have such difficulty in rendering articulate to ourselves) are incorporated already, en masse, within the grasp of thing-kind conceptions by which we find our way about the world. These conceptions are effective – in Leibniz's sense clear. They are not necessarily explicit or distinct. By reflection and practice we can make them more distinct (cf. Chapter Three, footnote 6). In so far as we have mastered these conceptions at all, however, they not only enforce for us the *Only a and b* rule that is disregarded in the proposal that Brownson Sole be ruled identical with Brown. They mark for us the distinction between singulars/particulars, which are not things instantiated, and universals which are. (Cf. Chapter Two, D(vi), Chapter Three, §4.) Normally we do not need to think about what we do. When we do need to think, we may have to struggle. (Philosophy is no sure ally here. Rather the reverse perhaps.) But none of this need count against the principle recently identified, the principle stated at the end of §15 and resumed in the paragraph before last.

In the light of that principle, the time has come to look once more at Shoemaker's version of the Brown-Brownson case. In pursuit of the enthusiasms of yesteryear, we allowed that case. Then we disallowed all the cases that followed on after it. But as soon as I announced (§8) that I would mark such a frontier, many philosophers will have wondered how much difference I could insist that there is between the process by which a whole brain is transplanted into a body ready and suitable to receive it, that by which one half of a brain is transplanted into a body ready and suitable to receive it and the process by which the second act is done twice over. If the second and third cannot yield identity, can the first?

Taken neat, this question is hard to answer, or even to see straight. But fending off the thought that questions of this sort ought to be reduced without residue to questions about the individuation and differentiation of surgical processes (questions for which I have just allowed we are ill prepared), why do we not remind ourselves yet again of our method and say, 'Well, here as everywhere it all depends on what sort of thing we take ourselves to be'? If we take that question more fully seriously, maybe we

can recover a wiser norm of judgment than any that we exercised in §§6–7. So let us go back to that point and everything that led up to it.

At the beginning we organized our inquiry around the sortal concept *human being* and bracketed the concept *person*. But the aim was not to let go of the idea of personhood, still less to place the question of what we are under the alien direction of physiologists, biologists, evolutionists or others who are expert in matters relating to *organisms*. The aim was only to fight free of the philosophical conceptions criticized in §2 following, then to force upon ourselves and other persons the question of what, in thinking of ourselves as human beings and human persons, we have undertaken to think of ourselves as being. No doubt, the answer must be discursive and draw on our inexplicit knowledge. No doubt it will be essentially contestable too. But what should I say?

The answer I give begins from yet another misappropriation of Locke's thoughts. It proposes that we apply equally to the concepts *human person* and *human being* that which Locke said of *self* and of *person*:

Wherever a man finds what he calls *himself*, there I think another may say is the same person. It is a Forensic Term appropriating Actions and their Merit and so belongs only to intelligent Agents capable of a Law, and Happiness and Misery. This personality extends it *self* beyond present Existence to what is past only by consciousness, whereby it becomes concerned and accountable, owns and imputes to it *self* past Actions, just upon the same ground, and for the same reason, that it does the present. (*Essay II, xxvii, §26*, punctuation modernized.)

If a term is truly forensic, it surely needs a palpable and public use — a use that is intelligible in the *forum*. But then the term's use in connexion with questions of praise and blame needs to be all of a piece with its other uses in the *interpersonal sphere* — and not least with all the uses to which P. F. Strawson drew attention when he showed how intimately our ideas of agency and responsibility depend on human beings' reactive and participative attitudes towards other human beings.⁴¹ Because such attitudes depend on the reality of human presence, that directs us to find out the connexion between these questions and Strawson's earlier insistence on the indispensability to the metaphysics and epistemology of personhood of the idea that a person is the bearer of both P- and M-predicates. How a human being stands or walks or frowns or smiles or laughs or sulks or earnestly entreats, or how he fries an egg, this *is* one part of what he is — no less than is the sort of thing ('more mentally') he chooses to say to this or that conversant, how musically he plays the

⁴¹ See 'Freedom and Resentment', *Proceedings of the British Academy*, 48 (1962), pp. 1–25.

violin (if he does), or how calmly or wildly, how sensibly or recklessly or obsessively, how magnanimously or ignobly (in a manner he is held responsible for) he responds day by day to the predicaments of ordinary life. Human beings' dispositions and capacities are gradually but constantly shaped and reshaped. In the process, the particular ways in which they come in due course to do whatever they do will become distinctive of what they have then become and *can be responded to* as having become.⁴² If so, then P-predicates, properly understood, are a subset of M-predicates. Unless they are, unless physiognomy and the particularities of a person's physical presence and being find their way into the story, Strawson's account of reactive and participative attitudes is hard to follow through. But is enough room left for all these things by the acceptance of Shoemaker's thought experiment? Surely the character of a person is not independent of his or her physiognomy, and this physiognomy can scarcely be independent of the body.

Consider the same matter from another point of view, which is complementary to the foregoing. In the Brown–Brownson case, after all the events that Shoemaker describes, it seems that Robinson still walks. But Robinson *is no more* (or so we were prompted to decide). We must get used to this. It seemed that Brown was dead, but that isn't right either (or so it has been decided). Where Robinson seems to be, there is Brown. We must inhibit then the affective responses we once trained upon Robinson, and redirect our affective responses for Brown onto the person who looks like Robinson but is Brownson. We must look into the face of Robinson and try to see there, try to find there, Brown. Meanwhile Brown himself, wearing Robinson's face, must come to terms with the difficulty that we have in finding him there. These are among the things to be made sense of in the Brown–Brownson experiment. Can they really be made sense of?

In considering the characteristically human modes of activity that appeared to be preserved in the Shoemaker case and not preserved in the cases that go beyond it, there was much to learn about the cognitive

⁴² As one speaks of these things, the phrase that may come to one's lips is 'his doing such and such an act in this or that particular way is a part of his identity'. When we speak at *t*, each of us, of 'our identity', what are we concerned with? Surely we are concerned with the fullest determination *as of t* of that which we are or have by the time *t* become. But this involves us with the so far shaped principle of activity which permits certain judgments of identity over certain imaginable transformations but disallows other judgments of identity over all sorts of other apparently imaginable possibilities. These other possibilities are thought of, however contentiously and by whatever admixture of bluff, as bringing with them the end or virtual end of the subject who undergoes them (if not necessarily of the subject's memory traces or the traces of those memory traces).

faculties and the conditions of their operation. But on the properly forensic view of human being-hood, why did not considerations of physiognomy (etc) that are familiar to us prompt, at the outset, all sorts of other doubts or hesitations about the Brown-Brownson case? How can the present author have been persuaded by the contingencies of current controversy to devote six long sections of this chapter to C, *C and R, but have allowed so fundamental an objection against most or all of these thought-experiments to lie fallow? It was strange, just because Brownson seemed in some ways to perpetuate Brown, for us to ignore the developed faculties of Brown which Brownson does *not* inherit and pay so little attention to all the implications of brain transfers of the kind that were described by Shoemaker. The high quality of the actors and mimics one sees on the stage should not lead one to think that the question of the fit of the brain to the physiognomy of the new body which is to receive it is as relatively simple as the transposition of music from one instrument to another.⁴³ But now that the Brown-Brownson Sole case and the potentially multiple output of the process by which Brownson Sole emerges as Brown alert us at last to further doubts, let us take them more seriously.

17. ONE LAST VARIANT — AND THE PHILOSOPHICAL MORAL OF SAME. FINALLY, HUMAN PERSONS AS ARTEFACTS?

At this point, the position must be considered of a philosopher who was not at all convinced by that which I said about the case of Brown and Brownson Sole (§15) but is suddenly persuaded that there is a serious case to answer concerning physiognomy and the Brown-Brownson case. The difficulty posed by physiognomy impresses him. Nevertheless, after further reflection, we may imagine that this philosopher insists that even the point that he has allowed casts no doubt on a case where Brown's brain is transferred to *the body of a Robinson who is an identical twin of Brown*. We may suppose that the philosopher also insists, with a fervour characteristic of present-day controversy (if not yet of science fiction New Comedy), that the point he accepts about physiognomy casts no doubt at all on a complicated situation where Brownson Sole emerges from the transplant of half of Brown's brain to the body of a Robinson who was one of four quadruplets. Let the first be Brown himself, a long lost son with a name given by foster parents; let the second be the Robinson

⁴³ Cf. Williams (1956: 7).

whose body is animated by half of Brown's brain; let the third be a Robinson in whom it happens that the other half of Brown's brain dies; let the fourth be long since dead.

In the first of these two cases, as newly conceived, the body of Brownson Sole is for all relevant purposes indistinguishable (my adversary will say) from the body of Brown. In the second case too, anything Brown can do with his body Brownson can do with his, etc. As for the process by which Brownson Sole comes to be from Brown [this philosopher may say], let that be characterized as the process by which *a person x is perpetuated in matter qualitatively indistinguishable from x's and as the sole inheritor of x's particular faculties and physiognomy*. If such a thing were demanded [he will add], then a guarantee could certainly accompany the process just characterized. For, *as thus characterized*, this process can be guaranteed not to create multiple candidates for identity with *x*. It can also be guaranteed not to collide with any of the cognitive considerations adduced in the endless chapter on this subject by David Wiggins. It is all very well for the purist of identity to claim that he is coming to terms with the forensic marks of our concept of human being. It is time for the purist to come to terms with the forensic character of our disputations about identity. Provided that we slightly redescribe the Brown-Brownson case and all its successors, these cases can still show everything they were intended to show.⁴

My preliminary reply to this philosopher will be to voice my doubt whether making Brownson's face *very like* Brown's face can fully overcome the disquiet that attaches to the very idea of 'wearing' a face. Off stage, one does not wear a face, only an expression of the face.

My second point of reply is that the guarantee my opponent proffers for Brownson Sole is entirely empty if the guarantee now given for the said process's not delivering multiple candidates only consists in the fact that, if it did do so, then we shouldn't *call* the process in question 'the process by which a person *x* is perpetuated in matter qualitatively indistinguishable from *x*'s own and as the sole inheritor of *x*'s . . . particular faculties and physiognomy'. Within our given ontology of processes, with going standards (however local, inscrutable or challenging to theory) of identity and difference, a separation of the kind that my opponent is attempting simply cannot be achieved by putting a special label upon one part of one variety of one of them. (Compare Aristotle, *Metaphysics*, VII, 1030^a ff.) New nomenclature cannot undo the obvious truth that, by the processes he is deploying and the means he is using to perpetuate Brown in Brownson, a surgeon skilled in such techniques

(namely transplanting brains or half-brains into bodies matching exactly the bodies they come from *or else*, for this involves the same things, into the bodies within which they will be viable) could produce multiples, and might indeed produce all sorts of further outcomes for which we are even worse prepared.⁴⁴

The third point of reply relates to the idea of a guarantee. This is an idea my opponent wants to cut down to size, to mock and belittle. Let us strive harder to make sense of it. A genuine guarantee relating to this or that process must relate to the nature of the process itself rather than a mere description of it. Moreover, genuine guarantees exist. However you describe it, the process of jam-making can be guaranteed not to produce heavy water out of ordinary water. The now standard process by which a stone is removed from the gall-bladder or is broken up there can be guaranteed not to remove the appendix. Another process that comes with a certain guarantee is the natural process, sustained by the operation of numerous laws of biochemistry, physiology and the rest, by which a human being comes into existence, matures and eventually ceases to be, by 'natural death'. That process is *not* of course guaranteed to save a human being from murder or from premature death by asbestosis, say, or irradiation. But it is certainly guaranteed not to produce multiples, not to transplant brains or half-brains, and not (if that were the better way to think of Brownson) to furnish new bodies to living, continuing brains. That is what makes this familiar process and the principle of activity associated with it one part of the basis for the making of judgments of *identity*. It is the lawful dependability of this process that entitles one whose judgments are shaped by that principle of activity to claim that his practice is answerable to the *Only a and b* rule. If the practice is answerable to that rule and it sees itself as reliant on dependable processes of this sort, then the practice will properly differentiate judgments of identity from judgments to the effect that the object *b* is the

⁴⁴ 'When Parties in a State are violent, he offered a wonderful Contrivance to reconcile them. The Method is this. You take a Hundred Leaders of each Party; you dispose them into Couples of such whose Heads are nearest of a Size; then let two nice Operators saw off the *Occiput* of each Couple at the same Time, in such a Manner that the Brain may be equally divided. Let the *Occiputs* thus cut off be interchanged, applying each to the Head of his opposite Party-man. It seems indeed to be a Work that requireth some Exactness; but the Professor assured us, that if it were dextrously performed, the Cure would be infallible. For he argued thus; that the two half Brains being left to debate the Matter between themselves within the Space of one Skull, would soon come to a good Understanding, and produce that Moderation as well as Regularity of Thinking, so much to be wished for in the Heads of those, who imagine they came into the World only to watch and govern its Motion; And as to the Difference of Brains in Quantity or Quality, among those who are Directors in Faction; the Doctor assured us from his own Knowledge, that it was a perfect Trifle' (Jonathan Swift, *Gulliver's Travels*: Book III, *A Voyage to Laputa*).

proper replacement/surrogate/proxy for object *a*, or of *a* ('is one of the *a*'s'). No doubt, as I have already allowed, the distinctiveness of identity may in some connections be unimportant. But those cases are a poor basis on which to engage in rhetoric about the 'forensic character of our actual disputations about identity'.

This threefold reply is likely to stir up at least two counter-objections. The first will consist of a challenge to apply the positive account that I have given of the human principle of activity to the case of the human zygote. Normally, the zygote becomes the embryo, but sometimes it divides and becomes twin embryos. What do I say about that? I am committed to react to this fact with a general ruling to the effect that the human being dates from a time after the zygote finally splits or settles down to develop in unitary fashion. The human being, the human person, dates from the point whenceforth it is nomologically excluded that its zygote will divide. I am happy, however, to be committed to say that.⁴⁵ An advantage of this decision, which entails that the fertilization of the egg is one part of the assemblage and further preparation of the materials from which the fetus develops, is that it has the effect of blocking the unwelcome suggestion that the fetus is the particular egg that the sperm activated. (If we concurred in that unwelcome suggestion, the fetus might even be accounted as just as old as the egg itself. What will explain better than the ruling that I have proposed why nobody wants to say that the fetus is as old as the egg itself?)

The second objection to what I have claimed concerns the guarantees on which the ordinary business of individuation normally relies. It may be expected to arise from the fact that such guarantees fall far short of any *logical* exclusion of the sort of events that figure in cases that reach beyond Shoemaker's example: 'These guarantees are logically insufficient. But it would be foolish [I shall be told], on the part of those who are so enamoured of their theory that they will put their shirt on guarantees of the kind that you mention, to seek to strengthen these guarantees. For then these enthusiasts will find they have excluded the

⁴⁵ I find that, years ago, intuitively and in its own way, by a route not wholly dissimilar from that followed in the text, the Committee of Enquiry into Human Fertilization and Embryology (1982) chaired by Baroness Warnock came to substantially the same conclusion. (See Lords Hansard, 7 December 1989.) On 'become' see p. 73, note 14.

If one is moved by moral arguments to adopt a certain conception of the person and one correctly deploys this conception upon some question of identity and difference, must the decision about identity and difference then qualify automatically as a 'moral decision'? I am uncertain, if only because of uncertainty concerning the terms of the question. Leaving that uncertainty on one side, let us not forget that seeing a moral reality can be the reminder of a metaphysical reality.

interventions of orthodox medicine and dentistry from 'proper' or 'admissible' application to human beings as such.'

To this I should reply that it is not my concern, either here or anywhere else, to represent judgments of strict identity as resting on justificatory premisses that culminate in the deduction of an identity or that bring among their logical consequences the satisfaction of the **D** principles given in Chapters Two and Three. That is not how the strictness of the relation of identity is to be marked or vindicated. We have long since rejected the deductive conception of identity-judgment. See Chapter Three, §5. (See also Chapter Six, §10, where I found myself reminding the reader of the familiar fact that one can see directly, however fallibly, that it is one's brother not one's uncle who has come to call on one.) The thing which really matters, where someone makes a responsible (however fallible) judgment of identity between the individual objects *a* and *b*, is that the various things the person judging has regard for in arriving at it should be the right things for a judgment of identity – that they be things appropriate for a type of judgment which, by its nature, goes well beyond the weaker claim that *b* instantiates *a*, that *b* inherits the role of *a*, or that *b* is interchangeable or interfungible with *a*. One who makes the judgment that *a* is *b* deploys an understanding which is all of a piece with the business of subsuming object *a* under some kind that is associated with a principle of activity (or way of behaving) that things of *a*'s kind exemplify. His understanding must be all of a piece with knowing how to keep track of such things, with grasping things' natures well enough to chronicle what they do or undergo, etc. (See Chapter Two and the various **D** principles that are assembled and paraded there. See also the Preface, last paragraph but six, for the description I see as consequential upon conceiving thus of identity and individuation.) For any genuine object, there is some right way of keeping track of it and this must track it *as a particular*. Cf. Chapter Two, §6. It is from this simple truism that the whole business of making of judgments of identity extrapolates.

What remains then of the difficulty about splitting (etc)? The real difficulty that the theory of individuation must have from cases of splitting, medical intervention, transplantation and the rest (or so I conclude) is not that such occurrences breach the limited and specific guarantees at which the objector is sneering (and which are all that are needed to sustain the relevant way of thinking of something's being securely singled out), but that they *add complications* to the picture we have held out of a substance coasting along (as it were) autonomously or under its own steam. (Which does *not* mean that, even in this picture, it interacts with

nothing else.) In coming to terms with these complications, quasi-causally so to speak, without much generality but in ready appreciation of the universality of the commitments that one creates by deciding in one way rather than another, we need constantly to refer the issues that arise back to our conception of the sort of thing that we are dealing with. Minor interventions (medical, dental, orthopaedic, osteopathic . . .) in a human being's pattern of being and acting confront us with no conceptual difficulties at all. In acquiescing in these, we do not prejudice our understanding of such a pattern. The substance's organic independence can still be conceived as undiminished. There is no problem of persistence. (We do not commit ourselves to allow anything on the scale of the Brown-Brownson case.) At the other extreme, however, in the so far imaginary cases where it seems a human being is simply treated as a template for the production of copies, it is manifest (it ought not to be controversial) that ideas are changed almost out of recognition. That is what I should say about Parfit's teletransportation, a fictional process which can as readily carbon-copy me twice (or thrice, or the number of times it takes to make a regiment) as once. In so far as these cases are taken as amounting to the perpetuation of the person Brown, we have lost hold altogether of the notions we began with of what Brown is. The judgment that the singular being Brown persists thereby is unsustainable.

Everything depends then on that which lies between the unproblematic cases and the cases that are out of the question. In these intermediate cases, where massive transplanting of organs or constant interchange of parts is contemplated, as well as constant fine-tuning of a human being by such expedients as gene therapy, the newly emerging conception under which we subsume a human substance will still be the conception of an individual thing or substance with a destiny of its own. Nevertheless, as we proceed along the road indicated, the conception of a human person will diverge further and further from that of a self-moving, animate living being exercising its capacity to determine, within a framework not of its own choosing and replete with meanings that are larger than it is, its own direct and indirect ends. The conception will converge more and more closely upon the conception of something like an artefact – of something not so much to be encountered in the world as putatively made or produced by us, something that it is really up to us (individually or collectively) not merely to heal or care for or protect but also to repair, to reshape, to reconstruct . . . even to reconceive.

This new or emerging conception of what a person is will perplex us not only with philosophical variants on the problem of artefact identity

but with practical questions. At the beginning, it may seem these questions will be easy enough. Later on, when less and less seems to be excluded by the then prevalent conception of human beinghood, they may bewilder us. I speculate that this bewilderment can only grow as that conception is progressively adjusted to the thought that little or nothing that a human being might more than idly wish needs to be out of the question or excluded by our human limitations. Will not bewilderment then turn to total *aporia* as our conception of human personhood is adjusted to the further thought that it is not merely our destiny that is (in large part) up to us, or our ethical identity, but even the kind of thing it is that we are? In the here and now, at a point well short of this limit, there has, of course, been a huge increment in the sum of human well-being. Why deny that? But before extrapolating this gain mindlessly into the open future, or simply rejoicing in the technological freedom that geneticists and medical scientists have been encouraged to create for us, there are questions to answer.

Here is one of them. If we cannot recognize our own given nature and the natural world as setting any limit at all upon the desires that we contemplate taking seriously; if we will not listen to the anticipations and suspicions of the artefactual conception of human beings that sound in half-forgotten moral denunciations of the impulse to see people or human beings as things, as tools, as bearers of military numerals, as cannon-fodder, or as fungibles; if we are not ready to scrutinize with any hesitation or perplexity at all the conviction (as passionate as it is groundless, surely, for no larger conception is available that could validate it) that everything in the world is in principle ours or there for the taking; then what will befall us? Will a new disquiet assail our desires themselves, in a world no less denuded of meaning by our sense of our own omnipotence than ravaged by our self-righteous insatiability?⁴⁶

I frame the question and, having formed it, I grave it here. But a book such as this is not the place in which to enlarge upon it, to answer it, or to speculate about the mental consequences of our available energies being diverted from the gradual discovery and enhancement, within the limits set by our animate nature, of more sustaining human ends – and the further consequences of these energies being progressively redirected to the endless elaboration of the means to ends that are less and less often explored in thought or feeling. This book is not the place nor

⁴⁶ As the book goes to press, I find a perceptive account of some related misgivings in Harry G. Frankfurt, *Necessity, Volition, and Love* (Cambridge University Press, 1999), pp. 108ff.

is its author the philosopher or moralist to hold up to human beings the image of the insatiability by which they relentlessly simplify the given ecology of the earth and change out of all recognition the given framework of human life. In a treatise on identity and individuation, it is enough and more than enough to point to the conceptual tie that links together these issues of life's framework and meaning, of individuation, and of the way in which we conceive of a human being.

In this work, the chief thing that needs to be made clear is that no freedom that the theory of individuation leaves it open to us to exercise in how we think of ourselves could ever liberate us from the constraints laid down in Chapter One. Whatever it may amount to, our freedom to remake the idea of personhood (to remake it by our pursuits and our choices) could not entail that it would be up to us, once it was determinate what *a* and *b* were, whether a certain identity judgment $a = b$ was true or false. For it is impossible on the level of reference and truth-value for there to be such freedom. As soon as it is determinate what *a* or *b* is, the result of Chapter One, §10, is inescapable. Moreover, once an identity question is construed in a definite way, it must be determinate what the question turns on. Nothing is left to stipulate. Only at the level of sense is there any room for conceptual invention to enter. At that level, I do not deny, there is freedom to reformulate bit by bit the answer that we commit ourselves to give to the *what are we?* question. It is a freedom we cannot really escape. It is cognate with the conceptualist (residually idealist) insight we expounded in Chapter Six and linked there with our miscellaneous capacities to come to terms with the exigency of the identity relation. But we should think much harder how we exercise it. We should think harder than about the choices that commit us to think of personhood in one way rather than another. These questions are essentially contestable. But from that it does not follow that they permit of more than one answer.

Apart from the logic of individuation, what ought to constrain us in how we think of human beinghood? Only the power of discursive reflection. How wise is it then to allow the goal of attaining the perfect state of material and bodily well-being to fly constantly before us like the rainbow or a will o' the wisp? How wise is it for those who have almost everything to project limitlessly into an open future their opportunities to reconstruct and reconceive themselves in order to have yet more? Where once angels feared to tread, how far will we go? If Schopenhauer had conceived how elusive and difficult the question was going to be found of what conception we should make for ourselves of ourselves; if

he had conceived how difficult it would be found by the very substances who exemplify the thing-kind whose nature has come into question; if he had known what kind of commotion the question would occasion or how sensitive the answer would be to the stability or instability of the expectations that still prevail about the world in which we still hope to find meaning and have our engagement; then I wonder whether he would have wanted to persist in his optimistic declaration that 'Just as the boatman sits in his little boat . . . so in the midst of a world of suffering the individual man calmly sits, supported by and trusting the *principium individuationis*' (*World as Will and Idea*, §63).

Representation

Select bibliography

For the basis of selection, see *Preface*, ad fin.

- Anscombe, G. E. M., 1953. 'The Principle of Individuation', *Proceedings of the Aristotelian Society*, Supp. Vol. 27, 83-96.
- Aristotle, *Categories*, especially Chs. 1-5.
Physics, Bk I, Chs. 7-9; Bk II, Chs. 1-3.
De Anima, Bk II, Chs. 1-4.
Metaphysics, e.g. Bk V, Chs. 6, 7; Bks VII, VIII, X.
Topics, 103a6ff. and 103a23-31.
De Generatione et Corruptione, 318bff.
Politics, 1276 a34-b13.
- Ayer, A. J., 1954. *Philosophical Essays* (London: Macmillan), Ch. 2, 'The Identity of Indiscernibles'.
- Ayers, Michael R., 1974. 'Individuals Without Sortals', *Canadian Journal of Philosophy*, 4, 113-48.
1992. *John Locke* (London: Routledge), Vol. II, Parts One and Three.
- Barcan Marcus, Ruth, 1947. 'The Identity of Individuals in a Strict Functional Calculus of First Order', *Journal of Symbolic Logic*, 12, 12-15.
- Black, Max, 1952. 'The Identity of Indiscernibles', *Mind*, 61, 153-64.
- Burge, Tyler, 1975. 'Mass Terms, Count Nouns and Change', *Synthese*, 31, 1977. 'A Theory of Aggregates', *Nous*, 11.
- Butler, Joseph, 1736. *First Dissertation to the Analogy of Religion Natural and Revealed to the Constitution of Nature*.
- Cartwright, Helen M., 1965. 'Heraclitus and the Bath Water', *Philosophical Review*, 74, 466-85.
 1970. 'Quantities', *Philosophical Review*, 79, 25-42.
- Cartwright, Richard, 1971. 'Identity and Substitutivity', in Milton K. Munitz (ed.), *Identity and Individuation* (New York University Press), 119-33.
 1979. 'Indiscernibility Principles', *Midwest Studies in Philosophy*, 4, 293-306.
 1987. 'On the Logical Problem of the Trinity', in *Philosophical Essays* (Cambridge, Mass.: MIT).
- Chappell, V. C., 1960. 'Sameness and Change', *Philosophical Review*, 69, 351-62.
- Chisholm, R. M., 1970. 'Identity Through Time', in Howard E. Kiefer and Milton K. Munitz (eds.), *Language, Belief, and Metaphysics* (Albany: State University of New York Press), 163-82.