

Locke, Butler and the Stream of Consciousness: and Men as a Natural Kind

DAVID WIGGINS

I. John Locke on Persons and One Objection of Joseph Butler

Locke defined a person as 'a thinking intelligent being, that has reason and reflection, and can consider itself as itself, the same thinking thing, in different times and places' (*Essay* II, xxvii, 2). To many who have been excited by the same thought as Locke, continuity of consciousness has seemed to be an integral part of what we mean by a person. The intuitive appeal of the idea that to secure the continuing identity of a person one experience must flow into the next experience in some 'stream of consciousness' is evidenced by the number of attempts in the so-called constructionalist tradition to explain continuity of consciousness in terms of memory, and then build or reconstruct the idea of a person with these materials. The philosophical difficulty of the idea is plain from the failure of these attempts. Hindsight suggests this was as inevitable as the failure of the attempt (if anyone ever made it) to make bricks from straw alone—and as a failure just as uninteresting. Which is not to deny that the memory theorist might get from it a sense that some of the difficulties in his programme have arisen from his leaving flesh and bones, the stuff of persons, out of his construction.

There is a distinction, well founded in the texts of Locke and Descartes on the one hand and Hume, James, Russell and Ayer on the other, which is sometimes put as the distinction between 'subject' and 'no subject' theories of the self. (It is not happily put so, because no-subject theories aim not so much to reject as to reconstruct the subject, and it comes back into discourse as a complex.) But the continuity of persons is to be discussed here on a level at which this division in the traditions of speculation about the nature and persistence of selves is irrelevant. When James writes that 'the continuing identity of each personal consciousness is treated as a name for the practical fact that new experiences look back on old ones, find them warm and appropriate them as mine', and that subsequent experiences must judge that 'these feelings are the nucleus of me' this 'no subject' theory is founded in the same sorts of question as

Locke's theory.¹ Some of these questions are, admittedly, misconceived.² But it does not follow that either tradition is wholly wrong to interest itself in the idea that memory and reflection help to constitute the continuity of the persisting self.

The first aim of what follows is to examine the charges of circularity or manifest absurdity brought against the theory that some such continuity is part of the essence of persons and their identities. Of these charges the theory can be cleared. The second aim is to pursue a potentially conflicting insight suggested by the first defence itself. Even if this insight finally subverts the Lockean identity condition it cannot, I think, refute the whole Lockean conception of a person. There is something so interesting about the idea that a person is an object essentially aware of its progress and persistence through time—a self-recorder so to say—and this notion is so closely related to some of the profoundest contentions of Kant,³ that one should look very critically at any attempt to demonstrate irreducible circularity in the continuity of consciousness condition of personal identity. But one ought to examine even more sceptically the view that adequate experiential memory is no essential part of the concept of a person.

Bishop Butler wrote in criticism of Locke 'And one should really think it self-evident, that consciousness of personal identity presupposes, and cannot therefore constitute, personal identity any more than knowledge, in any other case, can constitute truth, which it presupposes'. (*First Dissertation to the Analogy of Religion*). It must be allowed that if one defines a person in Locke's manner (if one means to be defining persons and not describing them merely), and if one really attaches enough importance to some self-recording capacity *d* to make *active possession of d* a definitive property of persons then, because exercise of the capacity *d* becomes a condition of the very existence and persistence of a person, *d* must indeed register upon the identity conditions for people. This only reflects a wider truth familiar from the teachings of Aristotle, Leibniz and Frege about the intimate relation holding between an account of *what a thing is* and the elucidation of the identity-conditions for members of its kind. If, as some have suggested, a difficulty is involved in stating a Lockean personal identity condition, then this difficulty can only exemplify a more general difficulty attaching to the attempt to build into the definition of any thing-kind *f* a reference to the capacity of a member of *f* to have at *t* some relation or other *to its own states before t*. There has, I suppose, to be (what I doubt there is) some systematic difficulty about this *its*.

¹ Like the role in which James casts 'stream of experience', 'the transition of one experience into another', 'coconsciousness', 'conjunctive relations'. See *Essays in Radical Empiricism* (Longmans, 1912).

² See footnote 4 below.

³ Cf. Bennett *Kant's Analytic* (Cambridge, 1963) p. 117.

There is probably room for disagreement about what exact objection Bishop Butler's famous sentence was meant to make against Locke. It does not seem to recapitulate the point made in Butler's preceding sentence. (This reads 'But 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 . . .'. We shall come in due course to the point Butler is making in this earlier sentence.) The sentence we are concerned with at the moment, about what is self-evident, appears to relate to some point of logical hygiene which, according to Butler, Locke ought to have been able to note before he even embarked on his ill-starred project—something as manifest as 'knowledge presupposes truth'. And the point seems to be that one cannot define A's identity with B in terms of A's *being conscious of having been identical with B* or *remembering being B*. But why, one wonders should a Lockean adherent not, instead of that, make the definition of A's being the same person as B in terms of *A's really or apparently remembering X-ing-at-time-t, just in case X-ing is what B in fact did at time t?*⁴ There will be obvious and extreme difficulties about this, but they are not the difficulties which the particular sentence of Butler's seeks to exploit. For a definition like this one does not even mention A's consciousness of identity with B.

At this point, however, other philosophers, Flew for instance,⁵ have been willing to reinterpret, amplify or extend Butler's argument. Flew writes, 'It is absurd to say that "he can remember that he is the same person". The absurdity is usually slightly masked since expressions such as "I remember doing, feeling, seeing something" do not contain explicit reference to the fact that what is remembered is that the speaker is the same person as did, felt, or saw whatever it was'.⁶ This does Locke the same injustice as Butler did him. It also imports a new claim, that 'A remembers X-ing' is a disguised or 'masked' version of what should be rendered explicitly as 'A remembers A's X-ing'. But so far from 'A remembers X-ing' having 'A remembers A's X-ing' as its canonical and

⁴ A defender of Locke might be momentarily misled into thinking that this definition came to the same thing as one which did mention that, if he failed to see that he must decline what Butler offers to him under the guise of a concession—that 'consciousness of what is past does . . . ascertain our personal identity to ourselves'. The offer is unlikely to have been disingenuous, but it must certainly be refused. See Sydney Shoemaker, *Self Knowledge and Self Identity* (Cornell, 1963).

⁵ See A. G. N. Flew's 'Locke and the Problem of Personal Identity', *Philosophy*, 24 (1951).

⁶ Op. cit., p. 55.

equivalent form (or 'A remembers B's X-ing' implying the identity of A and B), the two locutions are (I submit) importantly different in both construction and sense; though it may well be the idea that 'A remembers X-ing' has as its canonical form 'A remembers A's X-ing' which leads Flew himself into a curious oversight. He says (p. 56) that it is a necessary truth 'that x at time two is the *same person* as y at time one if and only if x and y are both persons, and x can (logical) remember at time two what y did etc. at time one'. To this I object that I still remember Hitler's invading Russia during World War II, and still remember Field Marshal Montgomery reminiscing in some speech or other he once gave when I was at school about how he won the battle of Alamein. If remembering these things makes me into Hitler or Montgomery, it makes me into both. But it doesn't matter exactly why Flew says this extraordinary thing. The important point is that the relation of A's *remembering* X-ing to A's *remembering* A's X-ing is really quite complicated.

If a very cool paratrooper, call him the first parachutist, forgets that he is to jump first; if for some reason, when he does jump, he pays no attention to the fact that he is jumping first; and does not remember afterwards that he jumped first; then he, the first parachutist, may remember jumping without remembering the first parachutist's jumping. (Anyway, if you ask him about it he may say 'I don't remember the first parachutist's jumping'.) And, conversely, he may remember the first parachutist's having his equipment checked—he watched not realizing it was his equipment but realizing that it was the first parachutist's equipment—and remember this without remembering having *his* equipment checked. For, whatever else he remembers, 'having my equipment checked' is not what he remembers about this event. There is much more to be said here⁷ but the difficulties of equating A's remembering X-ing with A's remembering A's X-ing, or remembering his X-ing, are part of a larger phenomenon deserving of more attention than philosophers have given it. Is my imagining being Moses, or an elephant, or Paul Klee's paintbrush equivalent to my imagining (the impossible state of affairs of) me being a paint brush or Moses or an elephant? Compare also foreseeing, visualizing, conceiving. Here, as with imagining, the inability to leave oneself out of the picture precisely disables one from achieving anything much by way of imagination.⁸

That there can be no jumping first or having one's equipment checked, no being Moses or being an elephant or being a paintbrush, without

⁷ The argument needs to be reinforced by examples free of the complications of scope imported by definite descriptions. To go into these and the other important syntactical ambiguities latent in such examples would be diverting but irrelevant, as would the larger task of assessing the putative equivalence of *A remembers X-ing* and *A remembers his X-ing*.

⁸ Cf. B. A. O. Williams 'Imagination and the Self', *Proceedings of the British Academy* (1966).

some subject's doing or suffering or being these things is not the same as to say that the subject must be specified, or must if specified be the same as the subject of the psychological verbs remembering, imagining . . . (jumping/being . . .). What is more, even if one always could rewrite 'A remembers X-ing' as 'A remembers A's X-ing'—which one can't, but even if one could—it would still be right to ask why *should* one rewrite it so? One can rewrite 'C is the same tree as D' as 'C is one and the same tree as D'. Will that refute logicism? If we must rewrite everything that we can rewrite then all attempts at philosophical analysis must be either incorrect or 'circular'.

II. Another Line of Objection and a Restatement of the Requirement of Continuity of Consciousness

Butler remarks in the same passage of the *First Dissertation* that to say that 'memory 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'.

This is a charge not of circularity but of outright absurdity. We might think it a decent first attempt at adjusting Locke's development of his project to evade it, and enough to secure something like the continuity which Locke desired between person P_{t_j} and person Q_{t_k} that, for some *sufficiency* of things actually done, witnessed, experienced . . . at any time by P_{t_j} , Q_{t_k} should later have *sufficient* real or apparent recollection of then doing, witnessing, experiencing . . . them.⁹ Let us call this relation (suitably tidied up, with more about 'sufficiency', and glossed somehow to allow for sleep) the relation C of *strong coconsciousness*, and let us symbolize its holding as $(Q_{t_k}) C (P_{t_j})$. Then anyone bent on grasping the nerve of Locke's conception of person would see, as Leibniz saw but not all of Locke's critics have seen,¹⁰ that the identity-condition he had to refute was one which made the persistence of a person *P* depend only upon *P*'s being related at each successive phase of his biography in this C-relation to *P* at each previous phase. C itself is a non-transitive relation but the *ancestral* *C of the C-relation, the weaker and transitive relation of *x*'s being either C-related to *y* or C-related to some *z* which is C-related to *y* or . . .—I shall call it *coconsciousness* simply—provides us with all we need for the

⁹ 'P' (or 'Q') stands in for a definite description of a continuing person, not of a 'slice'. The subscripts t_j , t_k . . . index the *time at which* the definite description *applies* to the continuing person. The subscript will be omitted where this does not need to be indicated. The notation leaves it an open question whether 'P' and 'Q' stand in for the same or different descriptions.

¹⁰ Even though it is a possible elucidation of 'extend' in Locke's words at *Essay*, II, xxvii, 2 (p. 449 Fraser).

statement of a general neo-Lockean identity-condition—

I_P : P is the same person as Q if and only if $(Q)*C(P)$.

The shortcomings of this will be quickly evident, and they are rehearsed in Section III; but they are not the defects of circularity, logical impropriety or outright absurdity; nor are they symptoms of any inherent defect in the more complex conditions which $*C$ might be used to formulate. What these shortcomings really signify is the unexpected *richness* and *concreteness* of the requirements on any satisfactory realization of Locke's original idea. As will appear, these requirements are radically inconsistent with the immaterialist conception which Locke himself arrived at. But it is not that part of his conception which this paper is to defend. The defects relate to (i) the formal adequacy of $(x)*C(y)$ to define an equivalence relation like identity, and (ii) the need to fortify $*C$ in order to secure what is actually intended by the continuity of consciousness requirement.

III. The Inadequacy of $*C$, as of any Uncontaminated 'Remembering' Condition of Mental Continuity. The Involvement of Remembering with the Physical

Defining $*C$ as we have, is there any positive reason to think that it is symmetrical and reflexive as well as transitive? Unless $*C$ has these properties its obtaining between some x and some y is certainly insufficient for personal identity, and may be insufficient even for what was intended by continuity of consciousness.¹¹

Suppose $(\exists y)((x)*C(y))$. Then does it follow that $(x)*C(x)$, that x sufficiently remembers doing a sufficiency of what x has just done? Even if we scale this down a little to require only that x has the *potentiality* to remember sufficiently at each stage a sufficiency of what x has done just before, there is already a problem. If a person is knocked down in a road accident and never recovers consciousness or memory at all, can he not remain biologically alive in hospital for years afterwards? Is he not an unconscious, or memoriless person? Some will find Locke's disqualification

¹¹ The reader may have believed it to be an objection to I_P that $*C$ is already an asymmetrical relation. Here we must guard against a misconception of what was meant by P_{tj} and Q_{tk} . What such descriptions stand for are not time-slices of people or 'person-moments'. They are people, persisting three-dimensional things which are born, live for some time, and then in one manner or another die. If P_{tj} and Q_{tk} are the same person then, regardless of the fact that these descriptions may pick out the person by reference to predicates which hold of the person at the different times t_k and t_j , their references are one and the same. Everything true of one is true of the other and, with certain merely grammatical adjustments, the designations are everywhere intersubstitutable *salva veritate*.

of this person strained or absurd, or will complain that it smacks of legislation, which it does. But it is not necessarily pointless legislation. It is not with reflexivity that the obvious or immediate difficulty lies. Symmetry and transitivity are the challenge.

Nothing in our definitions so far seems to have ruled out the following possibility: A stream of consciousness divides at some point and flows into a delta of two consciousnesses which are separate thereafter. Here we have, say,

$$(Q_{t3}) *C (P_{t1})$$

and also

$$(R_{t3}) *C (P_{t1})$$

One way of persuading oneself one can conceive of this is by picturing a man split from top to toe, exactly through the *corpus callosum*, artificial replacements being supplied to each half for all missing limbs and organs. Less ghoulishly, one may suppose he can conceive a Lamarckian state of affairs in which children inherit some substantial experiential memories corresponding closely to the life of the mother up to the moment of conception.¹²

If these things are truly conceivable (which for *persons* I doubt, but more of that later), then *C cannot be both transitive *and* symmetrical. If the stream of consciousness of someone who (say) planted the fig tree beside a certain wall divided in the fashion we have described, and if there were later two people who claimed to remember planting the tree (and did indeed remember planting it—all the factual and circumstantial and causal requirements of memory being satisfied in the fantasy), then the criterion I_P would force us to identify the two claimants with one another.¹³ But one may shout and the other may sulk when the dispute breaks out. They cannot be the same people.¹⁴

¹² Cf. Hidé Ishiguro, 'The Person's Future and the Mind-Body Problem' in *Phenomenology and Linguistic Analysis* (ed. Wolfe Mays and Stuart Brown) (London 1971).

¹³ Cf. Flew, p. 67 and B. A. O. Williams, 'Personal Identity and Individuation', *PAS* (1956–57).

¹⁴ Nor would it be right to insist on redescribing the situation as one where a single dissociated or scattered person did X and also, with another part of him, did not X. Such insistence might be based on an analogy with the way in which, without prejudice to his unity, a normal person can fidget with his left foot and not fidget with his right. There may be ideas of what a person is which allow this sort of redescription. But it is in the spirit of the Lockean notion, and of our own ordinary notion of person—the notion of a three-dimensional thing whose only genuine parts are *spatial* parts—to disallow it. When P_{t1} splits into Q_{t3} and R_{t3} , it splits into two *whole* persons, and Q_{t3} does not share R_{t3} 's consciousness of doing X. Even if Q_{t3} and R_{t3} both do X, their consciousnesses of doing X may be as distinct as if they were two people with no common origins at all. And they will communicate *inter-personally*.

Where does this leave us? If we were interested in continuity of consciousness for its own sake without wanting to *define* the identity and difference of persons by its means, we might rely on the fact that remembering is done by people (constituted as they are with the physical limitations they have) in order to underwrite it as an equivalence relation. But it is the other way about. Our only reason for insisting that *C must be an equivalence relation is precisely our desire to use it to define personal identity. In the absence of any commitment to that project, we might be content for *C¹ to be (say) reflexive, transitive and non-symmetrical.

Should we then abandon our interest in continuity of consciousness and scuttle the whole Lockean conception of a person? Not yet. If anything, Ip's insufficiency to define identity properly for the case of a delta in the stream of consciousness draws attention to the actual plausibility of the coconsciousness condition. For, in spite of the difficulty about *identity*, many people feel an almost overwhelming pressure in such thought-experiments to find a way to allow *both* the resulting splinters something like identity with the person who split. In a society where people occasionally divided, but which persevered in its interest in what Locke calls 'that consciousness which draws punishment or reward with it'¹⁵ a malefactor could scarcely evade responsibility by contriving his own fission. Nor, as Chisholm, Parfit and Williams in their very different ways have brought out, can reflection on the transitivity of identity be enough to make me cease to care about the future if I know that I am about to divide and there will shortly be two splinters, both coconscious with me. Such possibilities (or possible possibilities) do not show the mutual irrelevance of coconsciousness and personal identity. What are stirred up here are problems about survival, rationality, altruism, and prudence which are nothing if they are not problems of the self.

I announced that there were two difficulties in the *C conditions as stated. The first, just completed, related to the formal adequacy of *C. The second, to which I now come, is the need to fortify *C to play the role it was always intended to play. As it stands it comes nowhere near defining or capturing that mental continuity which animated the Lockean and constructionalist traditions. This is because it offers no plausible account of *error*.

Suppose that at t_3 I think that I remember locking the back door, though the fact is that some well-disposed neighbour slipped the latch (at t_2) after I had gone off without locking up. Then, unless we are already possessed of a criterion of identity or something else to refute the memory claim somehow, we seem to have

(Wiggins who at t_3 imagines he remembers locking the door) *C
(person who slipped the latch at t_2).

¹⁵ Cf. Leibniz, *Gerhardt*, IV, p. 460.

This is bad enough, but suppose that later at t_4 I began to doubt that I locked the door, remembering that at t_3 I had supposed I remembered doing so. Then undoubtedly

(Wiggins who doubts at t_4 that he locked the door) *C (Wiggins who thought at t_3 he remembered locking the door).

But then, whether I like it or not, *C must by the intended transitivity and symmetry hold between me (even as I doubt) and the man, or woman more likely, who locked the back door. But here I don't even have what Locke intended by continuity of consciousness with the person who locked it. And stiffening the subjective requirements on remembering locking the door (the vividness, stability, richness etc. of the inner representation) will scarcely alter the position. The troublesome example itself could always be correspondingly enriched.

This is a fundamental objection. If we try to answer it by writing 'remembers' instead of 'really or apparently remembers' in the definition of C and *C, then we encounter the difficulty that identity may be already involved in deciding, in some cases, whether someone really remembers planting the fig tree.¹⁶ To avoid the difficulty satisfactorily we must say what really remembering really is, and do so in such a way as to embrace the typical (Lockean) cases of remembering *witnessing*, remembering *perceiving*, remembering *feeling*, remembering *thinking*, remembering *remembering*, remembering *suffering*, remembering *doing*. . . . Nothing less will suffice to make *C the stuff of a correct usable Lockean criterion of identity, with the range of decidable cases we actually ascribe to the relation *is the same person as*. It will not do to define *C *in vacuo* as we did before—as if everything else could be defined at will, and without even relating a person and his actions in such a way as to say what it is for a person to *own actions* (Locke's phrase, *Essay*, Fraser, p. 479). If we do proceed as we ought, however, then no account of memory will suffice which does not arise out of a life-like account of the *whole range* of faculties

¹⁶ This is to suggest that just as the functions of remembering and intending and much else play a crucial role in regulating the individuation of persons, so perhaps identity itself regulates the correct application of the predicate *remembers X-ing*. To the critic who sees in this relationship of reciprocal regulation the chance of reviving Butler's objection to Locke, I offer the following *a priori* refutation of the generally received account of how clockwork functions in a timepiece. It is said that the mainspring unwinds, and in unwinding affects the hairspring. But it is also said that the hairspring affects the speed and manner of unwinding of the mainspring. How can that possibly be? If the normal operation of the mainspring presupposes the normal operation of the hairspring how can the normal operation of the hairspring presuppose the normal operation of the mainspring? Well, it can and it does. Presupposition like mechanical regulation can be reciprocal.

which are distinctive of persons. The objection scarcely shows that it is impossible to adjust the execution of the Lockean project while conferring a special role upon memory *amongst* other faculties. But it shows something about the sort of enterprise of which the analysis of memory must form part. For that reason, I shall pause here for a moment to say a word or two about the account of memory offered by Max Deutscher and C. B. Martin.¹⁷

Deutscher and Martin point out that even if A did (say) plant a fig tree in a certain spot at a certain time, A's thinking he remembers planting it is not enough to establish that A remembers doing so. He may have forgotten the actual planting. If someone has told him about it later, then A may have subconsciously imagined the planting; and even as he imagined it he may have forgotten that he knew about the action only from another person's account. This is a real if remote possibility. What is required then, in addition to some sufficiently vivid and plausible inner representation on A's part, is that there should be a causal relation between A's planting the tree and his subsequent memory-representation. This is a sufficiently familiar point.¹⁸ What is original in Deutscher and Martin's contribution to the subject is to have demonstrated, as a conceptual contention, that it was impossible to define the *right sort* of causal connection between an incident and the memory representation of it without recourse to the notion of the memory trace (which might be identified by reference to the normal neurophysiological connection, whatever it is, between rememberings and the incidents of which they are rememberings).¹⁹ They carefully explore a multiplicity of alternatives to the explicit memory-trace account of the causal connection between incident and experiential memory of incident. They show that none of these accounts can simultaneously allow for the possibility of *prompting* and define the particular sort of *operativeness* we are looking for between incident and representation.

In their physicalistic tendency these arguments about memory have important parallels with what seems to hold of other faculties which are distinctive of persons. If we are to make the distinction we believe we want to make between perception and misperception, for instance, there has to be something independent of what is subjectively given in perception to fix the *position* of the perceiver. What else can fix it but the body of the perceiver? Somewhat similarly, it seems to follow from Deutscher and Martin's analysis that experiential memory is inconceivable without some

¹⁷ 'Remembering', *Philosophical Review* (1966).

¹⁸ Cf. Grice, 'The Causal Theory of Perception', *PAS*, Vol. 35 (1962)—an analogy to which I was already indebted in the analysis of memory at *Identity and Spatio-Temporal Continuity* pp. 45, 55 and footnote 55.

¹⁹ This is not a circular procedure: but even if it were circular that would not matter for my purposes, which relate to the *necessary* conditions of remembering.

matter on which a trace is imprinted by the original incident, and which may carry the trace forward to the time when the person involved recalls it. What Deutscher and Martin lighted upon when they criticized 'if . . . then —' analyses of the conceptually requisite causality was, I think, a completely general difficulty in pinning down by conditionals the character of any mental disposition.²⁰ A disposition is the sort of thing which can rest latent, be revived and refreshed, and is at the disposal of its owner to use under all sorts of different circumstances—or not to use. It is impossible to conceive of memory causality by analogy with action at a distance as a transaction over a matterless gap between the external world at one time and a mind at a later time.²¹

In this sort of reasoning I think we begin to see how *C needed to be tackled. Defining it is just one part of the general exercise of describing a persisting material entity essentially endowed with the biological potentiality for the exercise of *all* the faculties and capacities conceptually constitutive of personhood—sentience, desire, belief, motion, memory, etc.

IV. The Charge of Circularity Rephrased and Reconsidered

That Locke and so many others have concentrated on memory to the virtual exclusion of almost all the heterogeneous other things which are characteristic of persons and which register in the stream of consciousness—the sensory and kinaesthetic aspects, the play of the practical imagination over possible outcomes, the constant surveillance of the fit of means to ends and ends to means and all the other forward and sideways looking elements—this, it may be said, is a lamentable even pathological symptom of much that is wrong with philosophy. Yes. But the charge of blindness or incompetence must not be allowed to obscure the Lockean claim that *x* is a person only if *x* has and exercises some sufficient capacity to remember or record sufficiently well from one time to the next enough of his immediately previous states or actions. That may yet prove to be true. But because this statement gives a part of the purported essence of persons it must impinge on the individuation of persons. And some will say that one circularity was manifest from the outset here. What about the *his* in 'enough of *his* immediately previous states and actions?' I reply that if there were really something illegitimate about the way in which the *his*

²⁰ For a more careful statement of this see my 'Towards a Reasonable Libertarianism' in *Essays on Freedom and Action*, Ed. T. Honderich (Routledge, London, 1973), p. 49.

²¹ It scarcely improves things to think of the memory trace as an immaterial imprint on immaterial stuff. This can only help to the extent that immaterial mind is made intelligible by being *modelled* on the material, and distinguished from it only by the apparently vacuous contention that it is immaterial.

turns up here in the Lockean condition, then there would have to be something equally illegitimate about the way in which the pronoun *it* occurs in the following definition of a black box recorder: *a machine designed to record certain aerobical data useful in case of accident befalling the aircraft which carries it*. I shall not try to add to Geach's deadly criticisms of philosophers who have been unable to understand the occurrence of pronouns as variables.²²

This confused objection does however lead into a better line of objection, which claims that, contrary to what I have implied, the criterion sketched above cannot operate effectively. 'Suppose *P* claims to remember planting the fig tree. We cannot settle the truth of his claim that he remembers planting the fig tree until we can establish by some criterion independent of memory whether *P* is *identical* with the person (if any) who planted it. But then memory is doomed always to bring too little too late to determine any identity-question. For we must already have settled the identity question by other means', it will be said, 'before memory is allowed on to the scene at all. If the memory condition is an extra condition applied after it is already fixed by some criterion *K* *what* we are to trace through space and time, then either it contradicts *K*, or at best it restricts (determines some subset of) the kind determined by *K*. Suppose *K* defines *animal*, then the memory-restriction may define *self-conscious animal*. But this,' the objector will say, 'is going to create a very nasty situation for you. It is going to be possible for the victim of the car crash who loses his memory to be the same animal as the patient who walks out of hospital to start a new life but not the same person—even though both are persons'.²³

So runs the objection: and every difficulty it deduces from the mistaken idea that memory must be an extra condition applied *after* another criterion *C* has fixed an individuating kind is correctly deduced. (What is more we shall revive something from the objection in due course, though not in the form of a charge of circularity.) But, against the objection as it stands, I have to ask what is this other precedent criterion *K*? Bodily continuity, it may be said. But that would make the person the same even as his skeleton and restore Jeremy Bentham (even the Pharaohs, presumably more

²² See e.g. *Analysis* 21, 3 (1961).

²³ An impossibility if we subscribe to Leibniz's Law. See my *Identity and Spatio-Temporal Continuity* (Blackwell, Oxford 1967) p. 3. In rejecting this possibility I distance myself yet further from Locke's own development of his theme. Locke tries to overcome the standard difficulties of *C* (amnesia, sleep etc.) by distinguishing questions of identity of *man* from questions of identity of *person*. This is a thoroughly unsatisfactory part of his discussion. However well one makes the distinction between the concepts *man* and *person*, this can hardly show that nothing falls under both concepts (under which is John Locke?), or that identity can be so relativized as to make the two identity questions independent of one another.

competently preserved) to the number of persons still extant and, so to speak, knocking about the place. Bodily continuity is not enough without life.²⁴ And surely sufficient exercise of the capacity for memory is conceptually connected with a whole cluster of vital functions, of which Locke can consistently argue that memory is just one. Locke's definition does not oblige us to say that memory determines identity questions autonomously, only that it contributes and, in conjunction with other considerations, is *crucially relevant* to our choice of continuity principle for determining the biographies of persons. What the objection really had to confront was a thesis about memory as part and parcel with other vital functions characteristic of persons. But against that thesis it simply begs the question.

The objection says 'We surely cannot determine whether P is remembering the fig tree until we have independently established whether P is indeed one and the same person who planted the fig tree'. It sounds for a moment as if there is something in that. But why not the other way round? Put the other way round the objection says: 'We surely cannot establish the identity or non-identity of P with the person who planted the fig tree until we have established whether (e.g.) this thinking he planted it counts as remembering planting it or not'. Perhaps there are cases in which this might be an equally good thing to say (or the only thing we could say). Shoemaker's brain-transfer case is the obvious example. If the back to front version has less weight than the first version, then we must remember that (on the neo-Lockean view) what is happening here is that we are ranging a *whole battery of components* of the continuity principle for persons against what is (on the neo-Lockean view) *only one component* of the continuity principle.

V. Two Senses in which a Person may be Supposed to Transcend his Body, One Correct, the Other Impossible

Let us pause here and take stock. We have neither formulated exactly—for reasons which will later become plain—nor demonstrated as sound or correct the Lockean condition of continuity of consciousness. But no argument we have examined, nor yet any obvious extension of one, shows that the Lockean notion partakes of outright absurdity or logical flaw.

²⁴ It may be said that we can and do distinguish between the corpse of a man freshly dead—he is still here but dead—and his earthly remains. When ashes (say) are all that is left then he is no longer there, it may be suggested. But the whole distinction which is relied upon by the objector is parasitic upon the point of distinguishing between life and death. Mere material continuity is not sufficient. And if life or its absence gives the *point* of these distinctions, then the *principal* distinction is between being live and being dead, and the best overall view will make existence or non-existence depend upon the principal distinction.

Nor is the revised conception of the memory condition which we have arrived at redundant in the presence of some distinct physicalistic criterion. Whether plausibly or implausibly, the memory condition informs and regulates the continuity condition of personal identity, holds it apart from mere continuity of material body, and leaves its distinctive mark on judgments founded in it.²⁵

The answer to Butler also provides a way to draw together certain other 'phenomena'. I have in mind the discomfort which almost everyone feels about any straightforward equation between himself and his body; the idea to which many of us are prey that the lifeless corpse is not the person; the fact that there is something absurd—so unnatural that the upshot is simply falsity—in the proposition that people's bodies play chess, talk sense, know arithmetic, or even play games or sit down. We bring these findings to the subject before speculation begins. The advantage of a neo-Lockean theory is that we do not need to force the facts to account for them.

A person is material in the sense of being essentially enmattered; but in the strict and different sense of 'material', viz. being definable or properly describable in terms of the concepts of the sciences of matter (physics, chemistry, and biology even) *person* is not necessarily a material concept.²⁶ The neo-Lockean functional-cum-psychological principle of individuation for persons shows how it is possible for *person* to fail of materiality in that sense, and shows it in a manner compatible with the strictest physicalism. For the continuity principle defines a material entity in the first 'enmattered' sense of 'material', while leaving it possible for the concept *person* to be (i) primitive relative to the concepts that pull their weight in the sciences of matter, (ii) primitive relative to the concept *human body*. If we understand what a living person or an animal is then we may define the body of one as that which realizes or constitutes it while it is alive and will be left over when it dies. The instability of the argument from analogy, like the scepticism which it was meant to answer, must testify to the difficulties of reversing this definitional order.

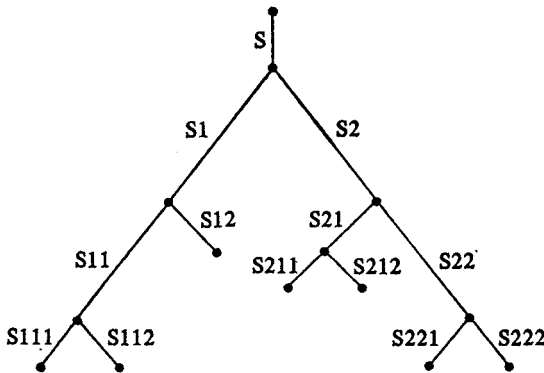
So much for an unexciting and satisfying way in which persons transcend bodies. But the supposed logical possibility of Lamarckian memory and

²⁵ Cf. *Identity and Spatio-Temporal Continuity* p. 45 para. 2.

²⁶ Nor, contrary to the purely physiological view of the question, is the seat of memory and consciousness just one part of the body: although, purely physiologically speaking, I suppose that the difficulties of replacing a brain are simply *more severe* than the difficulties of replacing a kidney or an eye. The physiological view would not be a doctor's view, but this is part of what people have in mind when they say doctors must be men of science and something else besides. Amongst the first to have thought into what thought-experiments involving change of bodily parts really involve is Stefan Themerson. See the last chapter of *Bayamus: a semantic novel* (Gaberbochus, London 1949).

of delta formations in the stream of consciousness (neither of which is manifestly excluded by a good account of memory itself) inevitably suggests another way.

If lines of consciousness could really divide—if the exigent requirements for this to have happened were fully satisfied—then it may be said we should need at least a dual use of ‘person’. There would have to be (1) a new use to denote a group-person or ‘clone’, whose members’ consciousnesses all derived from a point in some common biography;²⁷ and (2) the residue of our present use, to denote individual mental continuants which act separately and communicate interpersonally. Call (2) the ‘splinter’ use. Splinter persons will come into being either by birth or as the outcome of a fission, and will cease to exist if and when they either die or split. Consider the following ‘tree’ of divisions in a consciousness originating from a person S:²⁸



Note that in this diagram the S designations denote not nodes but arcs. The proposal is that if the divisions represented by such a tree come to pass then the arcs S1, S2, S11, S12 . . . are all different splinter persons. In addition, there is the clone or subclone person constituted of S111 and S112, the distinct but related (sub)clone person constituted of that plus S11, the distinct but related (sub)clone person constituted of that plus S12 . . . and so on, enumerating at every point of division all arcs which lie beneath it until we get the whole archetype, the clone-person proper, comprehending S and everything which diverges from S. On this use Smith is flesh and bone, but he has as members both individual (i.e. arc)

²⁷ Cf. *Identity* op. cit. (Appendix) 5.7 and 5.6 and (Part Two) 4.3 p. 54.

²⁸ For simplicity and because they raise special and even graver difficulties, I exclude from consideration here the *fusion* of lines of consciousness.

David Wiggins

members, and subclone members. He is a concrete universal. Each arc or splinter-person bears here the same relation to the archetype clone as all individual Cox's Orange Pippin trees (which have the ancestral of the relation 'produced from a cutting of' to an original hybrid tree) bear to the clone or concrete universal which they perpetuate and jointly constitute, viz. Cox's Orange Pippin.

One quickly comes to think there is yet a third use of 'person' we should need in order to make sense of a mental entity's having personal memory of experiences before the splitting of the consciousness. Conception (3) of a person picks out all the distinct continuous paths which can be traced back to the original arc S , i.e. $S \rightarrow S_{III}$, $S \rightarrow S_{II2}$ Let us say these paths define life-histories. The reason why they are needed is this. Suppose a person R remembers being, say, the first to sail between Scylla and Charybdis single-handed; R did navigate so, and R was the first to do so single-handed. He was there. But a branching may have intervened since this feat was executed. S_I , say, may have been the splinter agent involved, but S_I is no longer. How then does R survive? He is perpetuated by the splinters S_{II} and S_{I2} . They authentically remember the feat, but the actual navigator cannot be identified with the splinters S_{II} or S_{I2} . Neither of these splinters was there at the sailing. Nor can the clone conception help. We want to say of whoever survives the splitting of S_I something we need not want to say of splinter members which branched off before the sailing. Nor again can any subclone, e.g. $(S_{III} + S_{II2} + S_{II} + S_{I2})$, help us. For we may wish to say that the persons whom S_{II} and S_{I2} represent or perpetuate, i.e. the life histories $S \rightarrow S_{III}$, $S \rightarrow S_{II2}$, $S \rightarrow S_{I2}$, think and feel in importantly different ways about the feat. In these different lives the feat may represent different things. Only the life-history gives us a way to say this.

Before we enter into any of the difficulties or absurdities this project involves, it may be well to note how widespread in different cultures is the idea that persons may transcend not only particular bodies but even individual lives. Clifford Geertz has described the lengths to which the Balinese push the hypostasization of social roles, and the strange fusion of role and actor which is involved in their system of naming.²⁹ Perhaps the concrete universal conception of person is an extant conception then. Perhaps we could even discover it was our own conception. Within our own culture, or a neighbouring compartment, J.-P. Sartre has written in his book about Flaubert:

²⁹ The conception of person by which our own everyday life sets such store registers only in a rarely-used nonsense names. The names which matter are birth order names, kinship terms, teknonyms and status titles. Clifford Geertz, *Person, Time and Conduct in Bali* S.E. Asia Studies, 14 (New Haven, Yale, 1966).

Un homme n'est jamais un individu; il vaudrait mieux l'appeler un universel singulier; totalisé et, par la même, universalisé par son époque, il la retotalise en se reproduisant en elle comme singularité.³⁰

But there are serious doubts. If we want from the concrete universal conception only what the Balinese seem to get from having something of that sort—if this is really what they have—then perhaps the clone or the life history will serve to reconstruct it. For if this is all we want then there is no question of building up a coherent historical record of the individual passions, thoughts and actions of any individual person who is (say) Wayan. Nothing in Balinese culture is calculated to highlight the individual or, as it were, perspectival aspect of human experience. The whole ordering of the events of human history is reinterpreted by the Balinese so far as possible in terms of the recurrence of generic types of acting and suffering.³¹ Where there scarcely is such a thing as history the idea of an individual biography loses all purchase. It would be implausible to claim that the resultant conception of a person is utterly foreign to us. But if we take seriously as a thought experiment a world with persons liable to fission, and if we are anxious to hold fast to the very thing which made Locke's conception of a person interesting to us in the first place, then what we cannot abandon without forgetting completely the point of the exercise is our interest in constructing internally consistent, mutually consistent, indefinitely amplifiable, individual biographies. I suppose it might be claimed (implausibly I think) that we could jettison the ideas of childhood, maturity and death if people were perpetuated after the manner of plants in a hedgerow. (Re-duplication by fission would be scarcely very different.) But an interest in the Lockean conception commits us to try to preserve the possibility of an account of the formation of individual character—the path which a man picks through good and bad fortune to be what he in particular is. If so, the question is: Can the three uses of 'person' be deployed somehow, singly or in concert, to salvage this individual conception—even in the description of a world of fissiparous persons? I doubt it.

Consider again the claim to have been the first person to sail *single-handed* between Scylla and Charybdis? Can the claim be reinterpreted to accommodate the fact there are now three equally good claimants, $S \rightarrow S_{12}$, $S \rightarrow S_{111}$ and $S \rightarrow S_{112}$, with differing attitudes to that feat? I see only one way. The boast must be read as the claim that there was once an as yet undivided life history which performed the feat, a life history having the fortified *C relation to all three of the life histories just mentioned,

³⁰ *L'idiote de la famille: Gustave Flaubert de 1821 à 1857* (Gallimard, Paris 1971).

³¹ See especially the vivid and extraordinary description of the Balinese practice of drama in Geertz.

one of them being the claimant. The as yet undivided entity must comprise life histories which *will* be separated as distinct life histories. But at the time of the feat they were unseparated. The question must now be: What sort of a thing is a person if a life history is a person? For there will prove to have been one period, before the separation, during which two or three of them were in the same place at the same time. (An idea harder to accept than that of one place being occupied by two distinct things of different category or kind.)³² Unless our preconceptions about material objects are sadly astray, it follows that people are not only not human bodies but not even material objects. Well, there are those who thought this all along. But how is R, or any of the three life histories which now represent him, and on the importance of whose remembering such things as the sailing single-handed between Scylla and Charybdis the Lockean motivation lays such stress, to conceive (how is he to make a notion for himself) of the person that he is?

There is only one answer I can think of: as a Lesniewskian fusion or aggregate of person-moments or person-stages. The point about such Lesniewskian aggregates is that they may have common members, and that they will coincide as one person-fusion if, and only if, they have *all* their constituent person-moments in common. What is more, two, three or several Lesniewskian fusions *can* be in the same place at the same time; in this respect they are unlike material objects as we normally conceive of them. The trouble is that they are also much too unlike people. Look at the contrast between the predicates of ordinary material objects and the predicates of Lesniewskian fusions. A fusion has its relevant properties derivatively from the properties of its constituent *moments*. But at least half of the things we want to say about persons cannot be even tortuously explained in terms of the states at an instant of person-moments. Consider for example *weak, strong, clever, stupid, brave, cowardly, a good goalkeeper, a bad slip fielder, resolute, opportunistic, erratic, honest, a fair weather friend*, or (Galba's predicate) *capax imperii nisi imperasset*. We might I suppose explain what one or two of these meant in terms of person-moments if we were already possessed of their signification for persons. What seems inconceivable is the reverse procedure.

VI. Real Possibility: the Limits of Personhood

This is the unenviable position into which obsession with the Lockean condition apparently forces us if we accept fission and fusion of persons as logical possibilities. Most of us are committed to thinking that if any

³² Cf. my 'Being in the Same Place at the Same Time', *Philosophical Review* (January 1968).

of our concepts are sound then *person* is. We think that, being people, we understand what a person is. We also have a very exact understanding of the relation of sameness. Together these understandings should suffice to give a sound principle for *same person*—a criterion of individuation for persons. How then can personal identity confront us with such a variegated multitude of apparently insoluble cases? Yet the possibilities of fission and fusion which the recent literature of the subject has been obliged to rehearse are not on the face of it marginal indeterminacies which, however capriciously we find we have to decide them, leave undiminished our confidence in our comprehension of what *is at issue* here. The conceptual possibility of a delta in the stream of consciousness jogs our whole focus on the concept of personhood. So, rather than jump to the conclusion that we have no idea at all what we are about, let us ask: is such a delta really a conceptual possibility? It would be easy to rule out such a thing in a lordly fashion, after the manner of the followers of a certain great philosopher, as ‘undermining the whole application of the concept of person’. But what we really need is a *good pretext* to rule it out. This is to say that we need a reason arising out of a moderately well confirmed bit of semantical theory. My own attempt arises out of Hilary Putnam’s ‘Is Semantics Possible?’³³ and his theory of kind words. This I shall shortly apply to the sortal concept *person*.

If we look at the usual accounts of substance-words in the tradition of Locke’s nominal essence they are all, as Putnam pointed out, curiously unrealistic. Such accounts invariably seek to fix the sense of *sun* or *horse* or *tree* by a description (which always fails either of necessity or of sufficiency) in terms of their properties and relations and/or appearances. Putnam’s counter-proposal is that *x* is an *f* (horse, cypress tree, orange, caddis-fly) if and only if *x* is grouped by the most explanatory and comprehensive true scientific theory with a set of arbitrarily selected normal *exemplars* of the kind *f*. But if the correct articulation of natural kinds ultimately depends on good theory, and good theory is part and parcel of true statement of natural laws, then it follows from Putnam’s proposal that any putative definition of a natural kind *f* will stand or fall with the existence of some set of laws which collect together its actual extension. Unless there are such laws, the putative kind name has no extension, nor even the sense it is required to have. If there are such laws, on the other hand, then their holding is nothing less than constitutive of the existence

³³ *Metaphilosophy* No. 3 (1970). See also *Language, Belief and Metaphysics* (Vol. 1 of the *International Philosophy Year Conference at Brockport*), (Albany, New York, 1970). Similar conclusions have been arrived at independently by Rogers Albritton and Saul Kripke, and related ones by Mr Vernon Pratt of University College, Cardiff in ‘Biological Classification’, *British Journal for the Philosophy of Science* (1972) pp. 305–327. See also W. V. Quine ‘Natural Kinds’ in *Essays in Honour of C. G. Hempel* (Rescher ed.) (Reidel Dordrecht, 1969).

of *fs*: and these laws must define the characteristic development and typical history (or at least the limits of any possible development or history) of individual *fs*. It follows that, *if* person is a natural kind, then when we consider the problem of the identity of persons through change, the whole logic of the situation must exempt us from taking into account any but the class of situations which conform to the actual laws of the actual world. For these serve, and nothing but these can serve, to define the class of persons. And this seems to excuse us from allowing the spontaneous occurrence of delta formations in the consciousness of persons. The condition of persons' *existence* seems, subject to one obvious and important qualification concerning interference, to exclude this.

Before we attend to the antecedent of the conditional just asserted—'*if* persons are a natural kind . . .'¹—it is important to determine the ambit of Putnam's doctrine. It is not intended to cover the whole range of sortal words which we use to answer Aristotelian 'what is it?' questions. Consider artifact-words. There are virtually no natural laws about spades or clocks or tables or writing instruments *as such*. There scarcely could be. Clocks, for instance, may be made of a variety of different kinds of material and may function by radically different kinds of mechanism. Artifacts are things in nature, but they are not collected and classified together *as this or that artifact* by virtue of resemblances of any scientific or nomological import. They are collected up not à la Putnam, but under functional descriptions. A pen is a writing implement, a clock is a time-keeping device and so on. The description gives what it is impossible to specify in other cases, a *nominal essence*. But nominal essence is precisely not what the members of natural kinds have in common. What they have (and Putnam's account at last enables us to see how they can have this) is a scientifically palpable *real essence*. For the theory of individuation this is an important difference, and it results in a related and important difference between natural things and artifacts in respect of conditions of identity through time. I shall touch on this before homing again upon persons.

When there is any dispute concerning an object identified under a natural kind, then one can readily conceive of getting more facts. Remember the nineteenth century discovery that the elvers *Leptocephali* were in fact the young of the species conger eel. Or the humble (but in some sense proto-scientific) discovery that tadpoles become frogs. This observation enlarges the understanding of the concept. On the exemplar theory we can see how these scientific discoveries relate to the semantics of 'eel' or 'frog'. They reveal something which was always fixed by the sense of the kind-word, regardless of whether anyone knew this. It follows that identity questions about members of natural kinds can be expected to find the notion of identity at its best. They are the least plausible possible candidates for conventionalist treatment. Consider by contrast the identity problems of artifacts—the disassemblage of six or seven watches say, and their

eventual reassemblage with some confusion of individual parts.³⁴ Here we do not have natural things pursuing their natural course—there is no such thing as the natural development of a watch or a natural law about watches as such; and, in the last analysis, if parts get muddled there may be no point in arguing about which watch was which. (The one cast-iron indubitable sufficient condition of identity is a condition of limited utility on the lines of a criterion suggested for *quantities*, in the alien category of *stuff*, by Helen Cartwright,³⁵ a condition excluding any addition or subtraction of matter whatever.)

So much for artifacts and natural things. But are people a natural kind? Certainly a pure conventionalist view of the identity of people would fly in the face of the innermost convictions of almost everyone. Nobody thinks of the persons we actually encounter in nature as artifacts, or as having identities which are 'for decision' as artifact-identities are sometimes 'for decision' when there is a changing of parts. Yet *person* is not manifestly equivalent to the real essence *homo sapiens*, and the Lockean project looks much more like the analytical excogitation of a nominal essence than a piece of scientific research into the essence of a natural kind. What is more, it is a highly important fact about *homo sapiens* that in the way he conceives himself he does not allow his imagination to be constrained by that status. Men can readily conceive of having the faculties and powers of other animals. More concretely and practically we sometimes conceive of securing some or all of these and other powers to ourselves by means of tools. We may then start—in philosophy we have started—to

³⁴ Cf. p. 67, Paul Ziff 'The Feelings of Robots', *Analysis*, Vol. 19, No. 3 (January, 1959), and compare Hobbes' problem of Theseus' ship, cited at my *Identity* 2.1, p. 37. There are some telling considerations one can use to try to resolve disputes about that ship. (Pure conventionalism is not a very good story even for artifact identity.) But if doubt or dispute persists this is at least partly because one party is looking for an archaeological relic and the other for a functionally persistent continuant. Is one party wrong and the other right? There is scarcely anything to *discover* to vindicate either conception against the other. I do not mean the decision is arbitrary—only rarely is it entirely arbitrary—but nothing, as it would in the case of a natural kind, *compels* it. The antiquarian who favours the reconstructed ship has a different interest from the priest who favours the continuously repaired continuant. But both are stuck with the identification *ship*. Neither can base his view upon the natural development of a ship, or suggest a programme of research to resolve the question.

³⁵ 'Quantities', *Philosophical Review* (1969). I should make it clear that nothing in my discussion here of natural things *versus* artifacts is meant to rule out the possibility of borderline cases—think of a wasp's nest or (in another category and in relatively undeveloped parts of the world) cheese or bread: and that the discussion is intended to find a distinction with respect to essence between natural kinds and other kinds which will supersede the principle of distinction implicit in the etymology of 'artifact'.

abstract from our own biological condition altogether and imagine applying our technology to ourselves—as if to fashion men anew into the artifacts of men. Man as his own creature or tool. These thought-experiments may culminate in the assimilation of persons to creatures with powers conferred by magical artifacts which are themselves the product of imagination (Excalibur or Nothung or Gyges' ring)—even in the assimilation in thought of ourselves to such artifacts. We cannot assimilate artifacts to natural things, but certainly we are learning to assimilate natural things to artifacts.

What then is the present status of the sortal word *person*? Any answer to this is bound for many reasons to be controversial. One possible opinion is that its sense is still rooted in the sense of the animal species word 'man'; and that as human beings have come to the point where their powers of reason and analogy make it possible for some of them to transcend mere species loyalty, the sense of *person* has been very slightly modified. We have become open to the suggestion that other species may in varying degrees enjoy many of the attributes which we value highly in ourselves (not excluding the capacity to transcend species-loyalty, if the poet Arion is to be believed). As a result, the extension of the sortal word *person* could now be widened to accommodate members of the natural kinds who come near enough to us. To this (fitful) perception of the claims of other species to 'humane' consideration (and more) may also be added a desire to build into the concept *person* certain distinctive aspirations for the human condition; but it is characteristic of the view I have indicated that it insists on treating the essential characteristics of persons, both their capacities and their inherent limitations, as a matter of empirical investigation.

(An investigation which is urgent and important but not, it is consistent to add, so important as to overwhelm all other importance, or to justify any and every method of research.)

Against certain opponents this view of the concept is, I fear, simply question-begging. A second and slightly milder view would be that *person* is a *second order classificatory concept defined jointly in terms of natural kinds and a functional specification* somewhat as follows: *x* is a person if *x* is an animal falling under the extension of some natural kind whose members perceive, feel, remember, imagine, desire, make projects, move themselves at will and carry out projects, acquire a character as they age, are susceptible to concern for members of their own or like species . . . conceive of themselves as perceiving, feeling, remembering, imagining, desiring, making projects etc. . . . have, and conceive of themselves as having, a past accessible in experience-memory and a future accessible in intention . . . etc. On this account *person* is a non-biological qualification of *animal*, and a cross-classification with respect to zoological classifications, against the grain so to speak of an evolution-based taxonomy. Again it is not excluded that the extension of *person* should give hospitality to dolphins, to porpoises, or even (in exchange for suitably amazing behaviour suitably

explained in neurophysiological terms) to a parrot. A person is then an animal who has or is of a kind to have the biological capacity to enjoy the attributes enumerated above.

A third view which dispenses with the *animal* component altogether will be mentioned in due course, but let me concentrate for a moment on the second view. I think it would be worth while to pause longer than I can and ask where, on this view, we get the functional part of the definition of *person* from. It is not presumably intended to recapitulate and simply exhaust the biologically, historically, and culturally determined typical characteristics of *homo sapiens* as we know him. But what else constrains the list? Are we free, conceptually speaking, to shorten or lengthen this component at will?

These doubts relate to the second part of the second definition. The more they worry us, the quicker we recoil from the implausibly total subjectivity in which they ultimately threaten to engulf questions of norm and value, the less I think we shall want to dispense with the natural kind component in the definition. And so long as that component remains, the conviction that personal identity is a bad case for conventionalism is vindicated. What is more, the theory of individuation (whether we prefer the first or the second theory of *person*) suggests the possibility of rational grounds for what has previously looked like fear or prejudice—the conviction that robots and automata do not have any title to civil or personal rights or to any of the consideration due to sentient creatures. For the substance-concept under which we individuate such things is not that of any natural kind. To have feelings or purposes or concerns a thing must, I think we still think, be (at least) an *animal*. (None of this should be confused with the claim I do *not* make—that no higher animal can be artificially synthesized. What seems certain is that this would not be an *automaton*.) Finally, we understand the inherent limitations not only of speculation which goes contrary to the physical laws of the actual universe, but also of some not *physically* inconceivable thought-experiments involving the interchange of brains or their parts, the ‘carbon copy’ replication or reduplication³⁶ of skills and memories, or the implantation of artificial devices to replace, enhance, or supplement the operation of the natural organs. I am not arguing here against such practices. I am only saying that we must not be surprised if the associated thought-experiments—not to say their implementation—begin to subject us to the strain of thinking of ourselves as clones, as concrete universals, or as parts of one another. Perhaps this strain is good for us.³⁷ What we should first understand is

³⁶ See Bernard Williams ‘Are Persons Bodies?’ in S. Spicker (ed.) *The Philosophy of the Body*. (Quadrant Books, Chicago, 1970); also N. Miri, ‘Personal Identity and Memory’, *Mind* (January, 1972).

³⁷ As Derek Parfit has in several ways suggested. Cf. his ‘Personal Identity’, *Philosophical Review*, Vol. LXXX, No. 1 (1971).

what makes this strain. It is the apparent inability of *homo sapiens* to treat the essence he defines for himself as constrained by a natural kind, as the essence of one thing in nature amongst the other things in nature. It is his inability to immunize himself from delusions of technological omnipotence. What can result from this, amongst other things which cause more immediate consternation, is, at the limit, a kind of thought experiment, even a kind of practical experiment, which literally denatures the subject. In place of an animal or organism with a clear principle of individuation one finds an artifact whose identity may be a matter of convention, or even caprice. Certainly we do not, at this limit, find a person, if my account of the status of the concept *person* is correct.

VII. Amnesia Reconsidered and the Abandonment of the *C Condition. A Small Amendment to Locke

At this point it may be asked what difference it makes to any specific identity problem whether the indicated view of person is right or not; and how we shall decide whether the view is right or wrong.

The reader will have noticed that we never quite concluded the problem of the man who suffers total amnesia as a result of some appalling physical or mental shock and then begins life anew after a discontinuity which no meaningful *C condition could tolerate or let pass. The case is not impossible, still less counternomic, and there is no interference which pushes the subject towards the status of an artifact. It holds a part of the answer to the question just raised.

There are at most four ways of appraising such a case.

(1) We may hold that it voids any application of the concept of person. But in practice, when people do lose their memory, we never adopt this view.

(2) We may attempt to do honour to *C by deciding that the man who begins life anew after total amnesia is the same organism or animal as, but a different person from, the man who lost his memory. This is a tempting but incoherent decision. If *y* is the same animal as *x* then he is the person that *x* was. But then he is the same person as *x* (see note 23).

(3) We may decide to say that *y* is a different person from *x*, and a different animal or *homo sapiens*. Here there is no outright logical obstruction, unattractive though the view appears. We may disclaim any suggested equivalence between 'animal' and 'living body'—the suggestion is wrong anyway—and we may even extend the psychologicistic Lockean individuating procedure as far down the evolutionary hierarchy as there remain 'psychologically' interesting differences between different members of any single species of animal. We may hold that *animal*, of which *person* is a restriction, is not really individuating in the same way as *horse*, *cat*, *man* or *person* are. What coincidence under the concept *animal* amounts to, we may say,

differs according to the kind of animal, the genus-sortal being in this sense less fundamental than the species sortal.³⁸ And we may hold that *homo sapiens*, being a natural kind consisting of persons, permits of the individuation of its members in accordance with some *C requirement. What obstructs this view, however, if anything does, is the violence it does to what in real life we actually want to say about amnesic persons. In the last stages of Nijinsky's breakdown and madness, his friends and attendants tried, I believe, to reactivate his memory and recognition of the world about him by playing the music or performing the dances of the Diaghilev ballets he had formerly danced in. But by *C this was already a different person. Surely, though, the *homo sapiens* Nijinsky had neither died nor vanished, even in the last stages of (his) madness.

(4) The last or commonsensical view is that *pace* the *C requirement, it is the same person and same animal throughout. It is wrong to suppose that one can decide between this and decision (3) in isolation from all other cases. The force of analogy and the weight of other cases may make tolerable the element of legislation involved by decision (3). But in truth a case by case examination will not get very far unless it is informed from the start by some overall theoretical concern with the kind of concept that *person* is. If so, we need to be clearer about the theoretical rationale, if any, for decision (3) and for its being preferred to (4).

I think it would be fair to confront an upholder of decision (3) with this question: Since you allow the functional element in the notion of *person* to dominate the *animal* element, and permit a *C condition to invade and reshape the biological principle of individuation of *homo sapiens*, why stop there? What import does it have to insist that a person is an *animal* who . . .? Why not allow a mere artifact to qualify as a person provided only that it is programmed to satisfy the functional description in a substantial manner? And why not retreat to a position where the concept *person* is a nominal not a real essence? It seems to me that the champion of decision (3), like any upholder of the view that *person* is a primarily social concept, is pretty well defenceless against this suggestion. Many, in pursuit of analogies which it is not my present purpose to show how to block, will in any case embrace it and are already anxious to take physicalists, purists, and reactionaries of the identity relation to task for failing to see that the concept of person is a 'social' concept with identity criteria of an adaptability and pliability suited to this role.³⁹ But some day I think some identity purist should try his hand at drawing out the social

³⁸ These psychologically interesting differences between different members of a species will presumably give out well before we reach the oyster, to which Lucian thought it was so comical that Aristotle attributed a 'psychology'.

³⁹ Cf. J. M. Shorter, 'Personal Identity, Personal Relationships, and Criteria', *PAS* (1970-71).

and philosophical implications (in bourgeois, totalitarian and all other situations) of treating the concept of person otherwise than as a particular restriction of the natural kind concept *animal*. Which, I would ask, is really the philosophically more misleading or practically more dangerous conception of a man: a *pretentious imaginative tool using animal who is ϕ* or a *demiurge of animal origin and some residual animal inclinations who is ϕ* (where ϕ is in both cases some Lockean functional specification)?

I am not sure that I know how to take the point further (beyond a rhetorical question that is). But it is not merely *ad hominem*. The 'social concept' theorist is perfectly right to point to the conceptual importance of the role of the notion of a person in social and moral philosophy. His principal error lies in the instant deduction he makes from this, and in his, however unwitting, withdrawal of human nature from that morally directed objective empirical scrutiny which is our sole mode of access to men's real capacities and to the difference between what is variable and what is naturally invariable in their make-up. If we dispense with the *animal* component, and if the functional specification captures everything that is essential to personhood,⁴⁰ then I think there is no specifically *conceptual* strain in supposing that, provided only they continue to satisfy the specification, men can be adapted by whatever means, without the slightest detriment to personhood, to no matter what (politically stable) outcome. But then the focus of any rationally argued inquiry into political science must either be purely descriptive of existing forms of society and existing behaviour: or if possibility enters then the focus of the inquiry can only be technological—the pursuit of a social engineering interest in optimizing the quantity or distribution of want satisfactions. Where the empirically given nature and potentiality of man has no essential relevance or conceptual title beyond that which registers at the level of desire—or (worse) of behaviour, interpreted in terms of some (necessarily atomistic) theory of desiring—there can be no other 'hard' data, nothing else from which an inquiry in political science can rationally proceed. The one discrimination for any time t between the different possible outcomes projected by the managers of a society for the time t will be with respect to the total (and distribution) of satisfactions of wants existing at t , and the temporally discounted net present value at t of satisfactions actually or hypothetically existing at times after t . I do not deny that this view, recently celebrated by Marco Ferreri in his film *La Grande Bouffe*, is actually entertained by intelligent men; indeed I suppose that, with certain saving inconsistencies, this (reduced into mythical order by 'the' social welfare

⁴⁰ I assume that the functional specification is not trivialized by being continued to a point where it coincides with the full set of attributes which human beings (and . . .?) actually have in virtue of their biologically etc. given nature. Cf. the penultimate paragraph of VI above. This would immediately concede to me everything I am pleading for—and more.

function) is more or less the utilitarian theory of public rationality. So (at least in this place) I must rest content simply to remark upon the connection, and to note the distortion which the purely functional view must (in my opinion) induce in every moral or political concept or conceptual distinction that depends upon the idea of human nature. I cite as one instance the distinction (to whose practical and theoretical neglect our ravaged landscapes and the plundered social and architectural fabric of modern cities testify) between the basic needs of human beings—what because of their nature they must have to flourish⁴¹—and all their other appetitive states. This distinction between what depends upon the variable and what depends upon the naturally invariable elements in human life is not foreign to ordinary men, or even to the subjectivity of their most unreflective experience. It is indispensable to practical thinking in a world of scarcity. But if the distinction is conceptually inaccessible to a utilitarian manager he can scarcely interpret the everyday human behaviour he surveys as expressive of the distinction. And when his concept of person has lost any determinate centre of gravity (can be anything at all he makes of it) his methodology scarcely even requires the additional reassurance that ‘there is no such thing as human nature’.

In *Prometheus Vincitus* Prometheus reports how he found men witless, with ears to hear and eyes to see but perceiving nothing. Like shapes in dreams they passed long ignorant lives without purpose or achievement.⁴² When he taught them to tame the beasts of the field and sail the sea, to heal illness, to count and write, and to foretell the future from the entrails of birds, he gave men not only fire and tools, but optimism. He relieved men of foresight of their fate, and in the place of that, he says, he put blind hope. Without hope certainly nothing much could have been done. But without some vestige of the sight which Prometheus says he took away, my theory of individuation would predict that men will face something worse than the disappearance of fossil fuels. Not only will ambiguity have compounded the older difficulties of the ancient question ‘*what shall I do with myself?*’. We may confront a situation where there is no firm answer to what we are or even who we are, and where the limits of the class of persons is no longer supplied from outside us. For me (judging from here and now—where else?) it is difficult to see how the bulwarks, the shallow foundations if there are indeed any, of what Strawson has called the participative attitude⁴³ will continue at that point to sustain that view of the world and other persons.

⁴¹ I owe this view of needing to Miss Sira Dermen. It is prefigured in one form in Aristotle at *Metaphysics* Δ .5. More generally, see e.g. MacPherson in Blackburn (ed.) *Social Science and Ideology* (Fontana, London, 1972) pp. 25 ff.

⁴² Aeschylus, *Prometheus Vincitus*, 440 ff. and 250 ff.

⁴³ ‘Freedom and Resentment’, *Proceedings of the British Academy*, **XLVIII** (1962).

David Wiggins

There is too much guesswork and too much rhetoric here. Rather than add to it, let me summarize my conclusions. Decision (3) is quite plainly the wrong decision, using 'person' in the sense in which we still employ it. By *person* we mean a sort of animal, and for purposes of morality that (I maintain) is the best thing for us to mean. There is more to be said no doubt, and the concept *person* is in some way open-ended. We may (to some little extent) make of it what we will. But it is on pain of madness that we shall try to see ourselves as both *homo sapiens* and something with a different principle of individuation. Not even the cleverest charcutier could slice the world so to accommodate decision (2). In any case decision (4) is, I think, manifestly correct. The hospitalized amnesic, or Nijinsky even at the last stage of madness, are the same man and the same person.

What then about the *C condition? Its critics have been wrong to maintain that the *C proposal was absurd or circular. The only thing wrong with the *C condition was that for purposes of actual decision it was not quite right. Its critics have also overlooked how easily the project out of which it arose can be repaired. Let us amend Locke, and say a person is any animal the physical make-up of whose species constitutes the species' typical members thinking intelligent beings, with reason and reflection, and typically enables them to consider themselves as themselves, the same thinking things, in different times and places . . . Memory is not then irrelevant to personal identity, but the way it is relevant is simply that it is one highly important element amongst others in the account of what it is for a person to be still there, *alive*. It plays its part in determining the continuity principle for persons, as opposed to bodies or cadavera. Its bearing beyond this on individual problem cases is uncertain.⁴⁴

Bedford College, London

⁴⁴ A responsible review of the situation made in the light of the theory of individuation outlined here could not, simply by explaining the source of the perplexities they create, evade the problems posed by surgical and mechanical interventions which fall short of imperilling the application of *person* and related concepts. It would have to distinguish between the (implausible) necessary conditions and the (perhaps less implausible) sufficient conditions of personal identity which *C could be used to frame. I only remark that it should take nothing for granted about how well we really understand brain transfers of the kind described by Shoemaker. How do we fit the brain to the physiognomy of the new body which is to receive it? (Cf. Williams, 'Identity and Individuation', *PAS* 1958–59). How is the existing character expressed in the new body? We are deceived by the quality of the actors and mimics we see on the stage if with the help of greasepaint and props they have made us think this is as (relatively) simple as the transposition of music from one instrument to another.