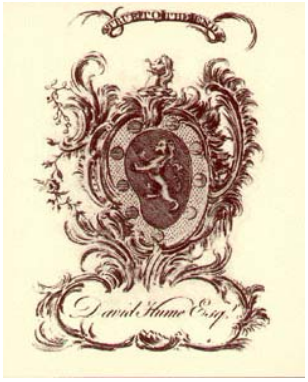

Reid's Criticism of Hume's Theory of Personal Identity
Harry Lesser
Hume Studies Volume IV, Number 2 (November, 1978), 41-63.

REID'S CRITICISM OF HUME'S THEORY OF PERSONAL IDENTITY

One of the most interesting philosophical controversies is that between Reid and Hume, considered as representatives of two different sorts of empiricism. Hume, for these purposes, represents 'radical' empiricism, and the attempt to base knowledge solely on experience and what can be validly inferred from it, regardless of how far this leads one from everyday notions and beliefs. Reid, in contrast, represents 'common sense' empiricism, and the view that the results of experience and deduction must conform to a number of 'first principles or intuitive judgments' (Essays on the intellectual powers of man, VI, 4), which are unprovable, but which no one can avoid accepting.

It is important to be clear about the status of Reid's 'intuitive judgments', or 'principles of common sense'. They are not merely the general beliefs of a particular culture, which could be 'vulgar errors or prejudices'. Nor are they merely beliefs generally held by people uncorrupted by philosophy. They are assumptions which have to be made if theoretical or practical activity is to be possible at all, principles which are always accepted in practice even by those who deny them in theory. For example, although the uniformity of nature is not a logically necessary truth, everybody always thinks and acts in ways that presuppose that it is true: 'Nature is governed by fixed laws, and if it were not so, there could be no such thing as prudence in human conduct .... the whole fabric of natural philosophy is built upon this principle.' (Essays, VI, 5) Again, everyone, even a theoretical determinist, makes moral judgments and deliberates what to do; and these activities presuppose a belief in human free will. (Essays, ibid.)

Hume himself, as Reid was aware, accepted the validity of some principles of this sort, notably of the

uniformity of nature. But Reid was dissatisfied with Hume's account, partly because he thought Hume had rejected far too much, and partly because of the grounds Hume gives for accepting, for example, the validity of induction, and the existence of the external physical world. Hume regards these beliefs as non-rational, and deriving from our psychological make-up; it is nature and not reason which causes the belief in an ordered physical reality. In Hume's eyes this is a sufficient justification: if our nature compels us to believe something, this is as good a ground for believing it as any other. Reid, on the other hand, believed, first, that there were many more principles of this sort than Hume had acknowledged, and secondly, that the justification of them was rational, and not merely psychologically determined, since it was based on the impossibility of thinking and acting coherently without them. Hence, even though Hume builds up with one hand what he pulls down with the other, and destroys the rational basis for belief only to replace it by one based on human nature and psychology, Reid continues to treat him as the arch-sceptic.

In constructing his list of first principles in Essays, VI, 5, Reid explicitly or implicitly attacks Hume's 'scepticism' on five points - the existence of the mind; the existence of the physical world and the reliability of our senses; the reliability of memory; the reliability of reason; and the uniformity of nature. Against this scepticism he deploys three types of argument. The first is the argument from the universal beliefs of mankind, particularly as displayed in the structural features supposedly common to all languages. The second is the argument that the position of the sceptic is, when its consequences are worked out, either self-contradictory or incoherent. And the third is the argument already mentioned, that this position is inconsistent with other beliefs which the sceptic inevitably holds, since he thinks and acts in ways which are

inconsistent with their falsity, and cannot avoid doing so.

To assess the force of Reid's attacks, we have to consider both the strength of arguments of this sort in general and whether they apply in all these five cases. In this article, I want to concentrate on the first issue of the five, that of personal identity. Apart from displaying Reid's arguments at their most developed, this is in at least one respect the most interesting of the five points of controversy. On the other four points Reid and Hume do not disagree about what is the case, but only about what can be known to be the case: they accept the same set of beliefs, at least on a realist interpretation of Hume, and are arguing only about whether the basis for them is rational. But about the 'self' Reid and the Hume of the Treatise of Human Nature are in ontological disagreement.

This disagreement is not so much a dispute about the existence of the soul, mind or self (these three words are, for our purposes, equivalent), as about its nature. Hume identifies the soul with the continuous process of 'perception', i.e. mental activity, whether thought, feeling, sensation, decision, imagination or anything else. It is not a subject of these activities and experiences, distinct from them and unchanging, but is the stream of experience itself: [mankind] *are nothing but a bundle or collection of different perceptions, which succeed each other with an inconceivable rapidity, and are in a perpetual flux and movement* (T252). *I cannot compare the soul more properly to any thing than to a republic or commonwealth, in which the several members are united by the reciprocal ties of government and subordination, and give rise to other persons, who propagate the same republic in the incessant changes of its parts.* (T261). Reid on the other hand, identifies soul or mind precisely with an unchanging subject of activity and experience: 'I am not thought, I am not action, I am not feeling; I am something that

thinks, and acts and suffers. My thoughts and actions, and feelings, change every moment; they have no continued, but a successive, existence; but that self or I, to which they belong, is permanent.' (Essays, III,4)

The first type of argument Reid uses in support of his position and against Hume's, to be found already in Essays, I, 1-2, is, as I have indicated, an argument from language; indeed, it is notably like some of the arguments popular with British and American philosophers in the fifties and early sixties. It is not intended as a conclusive proof, but as evidence for the existence of a belief which Reid claims to be shared by all rational people uncorrupted by philosophy, and therefore to be respected until proved false. There are really two arguments in Reid, one seeking to show that the belief in a self is virtually universal, the other that Hume's 'no self' theory cannot be coherently stated: consideration of this second argument will lead us into considering the coherence of Hume's theory as such.

Of these two arguments the second is, I think, much the stronger. The first is interesting, but highly dubious. Reid argues that 'in all languages' the verbs for 'to think', 'to perceive', etc. are invariably active, and that this shows that in all societies thought, perception, etc. are believed to be activities, and as activities to be the products of an agent, i.e. a mind distinct from the thoughts themselves. But in fact even those languages with which Reid was acquainted use active verbs to denote processes as well as actions: the earth spins, rivers run, stones fall, and the impersonal 'it' rains and snows. The mere use of verbs in the active voice does not seem to prove anything by itself.

On the other hand, Reid is right to draw attention to the linguistic oddity that would result if ordinary speech incorporated Hume's doctrine of personal identity, so that words like 'I' or 'Hume' would refer to particular

successions of related 'ideas and impressions' (i.e. mental events). It would mean, Reid says in Essays, VI,5, that 'this succession of ideas and impressions, not only remembers and is conscious, but ... it judges, reasons, affirms, denies; nay, ... it eats and drinks, and is sometimes merry, sometimes sad. If these things can be ascribed to a succession of ideas and impressions, in a consistency with common sense, I should be very glad to know what is nonsense.'

This argument requires some modification before we can evaluate it. First, we should not be misled by the reference to eating and drinking. It is indeed absurd that a succession of ideas and impressions should eat and drink; but it is equally absurd that a 'Cartesian ego' or non-physical self should eat and drink. The absurdity results from defining a person without reference to his body, so that he can be said to perform physical actions only indirectly. This is a weakness both of theories like Reid's which posit a spiritual self and of those like Hume's which identify the person with a set of mental events; it does not tell against one view more than the other. Moreover, both views can be adapted to take account of the physical; and if anything this is easier for a follower of Hume, who needs only to include physical as well as mental events in the chain of events that constitutes a 'self'.

The problem now is whether a succession of mental events can intelligibly be said to remember and be conscious: 'I would wish to be further instructed, whether the impressions remember and are conscious of the ideas, or the ideas remember and are conscious of the impressions, or if both are conscious of both.' (Essays, ibid.) Here Hume could object that to say 'the succession remembers' is indeed incoherent, but only because it mixes two kinds of discourse. We may say 'Yesterday Hume remembered his boyhood', using ordinary speech, which Hume regards as perfectly correct and unobjectionable once we are clear what we mean

by it. Or we may say 'Yesterday there was in the succession of events we call Hume a memory of certain earlier events', which is more accurate, but also more awkward, and unsuitable for ordinary purposes. But we cannot mix the two ways of speaking, and say 'Yesterday the succession of events remembered its boyhood', without absurdity.

But, one may still ask, is not the impersonal mode of speech, even when correctly used, absurd in itself? Is it not replacing sense by nonsense to say, instead of 'I thought it was dinner-time, decided to break off studying, and went to eat,', 'At such-and-such a place and time there was a thought about dinner, a decision that study should cease at that place and time, and a movement towards food'?! To see whether or not this is nonsense, let us consider why we would not want to talk in this way. Let us suppose, for the moment, that Hume's analysis is correct, and ask why it is that we do not talk in what would then be the most accurate way, and whether anything in the answer makes the supposition that Hume is right implausible.

First of all, the impersonal mode would be intolerably long and convoluted for ordinary purposes, as we can see from the above example. Secondly, if we did not habitually give names, such as 'Hume' or 'I', to these 'chains of events', and ascribe activities and processes to them, as in 'Hume remembered his boyhood', but instead talked only about individual events and their relations to each other, we would often have difficulty in discovering in what chain a particular event belonged - for example, whether a particular action belonged in the chain of events we call 'Locke' or the one we call 'Hume', i.e. to revert to normal speech, whether Locke or Hume did this particular thing. To name chains of events, thereby treating them for practical purposes as entities, and then to ask which events should be ascribed to these quasi-entities, is much easier than to name individual events and ask which of them belong together.

Thirdly, the impersonal way of talking tends to be both emotionally neutral and focussed on particular events rather than on the connections between them. This is sometimes desirable: there is sometimes an obvious need to adopt a state of mind which views dispassionately the other states that constitute the person of which it is part and the states that constitute other people. There is also a need for states of concentration on the immediate environment, without the distractions of memory or prognosis. But there are other times when what is needed is an awareness of the continuing influence of one state of mind or body on another, whether within the same person or between two people. This awareness is helped by the ordinary way of talking, which in this respect performs a function for which a more accurate mode of speech would be less satisfactory.

There are thus some very good reasons for the way we talk, even if it is, in its implications, not entirely accurate, as Hume would have us believe. But the reasons are essentially pragmatic: they relate to the convenience of the ordinary way of speaking, to its suitability for the sort of things we want to know and find out, and sometimes, though by no means always, to its practical and ethical merits. They show that to speak impersonally would be inconvenient and impractical for ordinary purposes, but not that it is either unintelligible or uninformative. Moreover, the fact that a mode of speech is undesirable on the everyday level does not show that it is not nevertheless more accurate than everyday speech and also more useful for other purposes. The language of particle physics is no use for ordinary talk about physical objects, and the various technical languages of psychology are not much use in everyday human communications; but this does not affect the usefulness of either for scientific purposes. Similarly, if Hume is right, the impersonal way of speaking about mental and physical states might be extremely useful for the purpose of accurate and dispassionate self-analysis.

So it appears that Reid's argument, if we interpret it as a linguistic one, is in the end unsuccessful. But this is not the only way of taking it: Reid's words suggest that he is talking both about the unintelligibility of the way of speaking and the incoherence of the theory itself. He could accept all that I have just said, and still maintain that to try to talk or think about mental activity impersonally is incoherent simply because it tries to be impersonal. To say by implication, as Hume does, that there are thoughts, emotions and actions, but no thinkers, subjects of emotions or agents, is, in Reid's eyes, to say something nonsensical and incoherent: 'that every act or operation ... supposes an agent ... I do not attempt to prove, but take for granted.' (Essays, I,2)

This last point can be made true by definition. But the question then will be: which events, if any, are acts and which are processes? (We have seen already that verbs in the active voice refer to both.) Our ordinary beliefs recognize three sorts of 'event', in the neutral sense of 'something which happens'. There are events which simply take place, such as storms, rises in temperature, and dawns. There are events which have to happen to something or someone, in order to take place; a fall or a movement or a colour change must be the falling or moving or change in colour of some one, or more, particular objects. And there are 'events' which must be performed by some agent, such as murders, the making of promises, and looking and listening; murder, for example, is defined, in part, as consciously bringing about a death, and not merely happening to be the cause of one. In ordinary speech we call 'events' of this last sort 'actions', and distinguish them from states or processes (or 'events' in the usual sense of the word).

Hume and Reid agree, as we shall see later on, that a distinction of this sort has to be made. But Reid sees it as a self-evident truth that the difference between

actions and states or processes is simply that 'every act ... supposes an agent', whereas Hume denies that any sort of event presupposes agency, i.e. that there are 'acts', in Reid's sense: all events, for Hume, are of the first type, and none of the second or third.

By saying, in effect, that there are no acts, but only events (if we may now revert to the usual sense of 'event'), Hume avoids the incoherence, attributed to him by Reid, of positing acts but no agents. But there is another problem here, that of whether there can be 'pure' events, which do not happen to anything. For it can be argued that the difference between the first and second types of events is merely verbal: talk of storms, rises in temperature, etc. is really shorthand for talk about sets of events all of which must happen to particular things in order to take place at all. A state, it may be said, can exist only if it is a state of something, and a process only if it is undergone by something.

I am myself uncertain whether this is a fatal objection to Hume or whether we are being led astray by the requirements of grammar. If the objection is valid, it shows the impossibility of a view of the world which admits only states and processes, and not entities, objects or 'things'. This impossibility would be conceptual, not grammatical: it is almost certainly possible to construct a language in which there are no nouns, but only verbs, and it has been argued that one or two Canadian Indian languages are actually of this type, and represent, for example, 'the stone fell', by a locution which would be more literally rendered as 'it stoned downwards'.[1] The grammar of these languages thus recognizes no stones or houses or people, but only the events of 'stoning', 'house-ing', and 'being-a-person'. But if, as has been suggested, the taking place of such events presupposes, like the taking place of a storm, that there are entities to which the 'sub-events' that make up these larger events are happening, then the

getting rid of objects or entities has taken place only on the verbal level, and their existence is still being presupposed.

Yet even admitting this objection would not require Hume to recognize the existence of 'selves' in Reid's sense. He would have to posit only momentarily existing entities; and he could still describe the world, at any one point in time, as consisting of a set of states of these entities, causally connected in various complex ways with states of other equally momentary objects. This would not be very different from describing the world in terms of events; and it would give us a world quite different from Reid's 'common sense' world of objects and persons persisting through time. These momentary objects would lack not only persistence but also agency; they would take part in events, but not do anything.

Hence, provided we give up talking about acts, we can describe the world quite coherently without positing 'selves': the incoherence which Reid pointed out arises simply from positing acts but not agents. Since it is coherent, at any rate, to talk only about events, Hume's theory seems to be internally coherent. But the question remains whether it is consistent with other fundamental beliefs which we need to maintain. On this question Reid indicates a number of such beliefs which entail the existence of persisting and acting selves.

The first of these beliefs is one which Reid does not mention in this context, but which I take the liberty of supposing that he would be sure to mention if he read the above revision of Hume's theory. This is the belief in free will. I think that if one could convince Reid that active verbs can denote states and processes as well as actions, and then go on to maintain that mental events are states or processes, he would regard this as destroying the possibility of human free will, since persons can be free only if they can act. Now the conviction that we have

some degree of free will 'is necessarily implied in many operations of the mind [such as volition, deliberation and moral judgment], which are familiar to every man, and without which no man can act the part of a reasonable being'. (Essays, VI,5) So no belief inconsistent with human free will can be seriously maintained in practice, but only in theory; and if the 'no self' theory is inconsistent with belief in free will it must be rejected.

But the 'no self' theory, whether in Hume's version or any other, seems in fact to be neutral as between determinism and indeterminism. If we identify a person with a set of events, we must admit that earlier events in the chain influence later ones, but we need not hold that they totally determine their nature. However, this kind of indeterminism would probably not satisfy Reid. The principle which he claims to be a principle of common sense is 'that we have some degree of power over our actions, and the determinations of our will'. This seems to entail not indeterminism, but rather the maintaining of the distinction between acts and events; the important point for Reid is not that some things we do are undetermined, but that we have some control over them, i.e. that they are genuine actions.

Reid is, I think, right to maintain that we must make a distinction between what a person does and what happens to him, between, for example, my raising my arm and my arm's going up, or between jumping and falling into a swimming-pool. But is there no way of making this distinction other than by positing persisting agents? A Humean, even though he classes both actions and processes as events, can still maintain that they are events of different sorts, and that a sufficiently detailed account of the physical and mental circumstances of each event will enable us to distinguish between them. A full account of what happens when 'I raise my arm' and when, in contrast, 'my arm goes up' as a reflex action, will show the two

events to be different. To make the distinction it is very convenient to introduce an 'I', but we do not need to do so: we need only to describe the event, and what precedes and follows it, in sufficient detail.

We may now turn to Reid's other argument, or set of arguments, against Hume. These appear in Essays, III,4, in a passage from which I have already quoted, but which I would now like to quote at greater length:

> My personal identity ... implies the continued existence of that indivisible thing which I call myself .... I am not thought, I am not action, I am not feeling; I am something that thinks, and acts, and suffers. My thoughts, and actions, and feelings, change every moment; they have no continued, but a successive existence; but that self, or I, to which they belong, is permanent.... [T]he proper evidence I have of all this is remembrance. I remember that twenty years ago I conversed with a certain person .... my memory testifies not only that this was done, but that it was done by me who now remember it. If it was done by me, I must have existed at that time, and continued to exist from that time to the present .... Every man in his senses believes what he distinctly remembers, and everything he remembers convinces him that he existed at the time remembered .... The identity of a person is a perfect identity; wherever it is real, it admits of no degrees; and it is impossible that a person should be in part the same, and in part different; because a person is a monad, and is not divisible into parts. The evidence of identity in other persons besides ourselves does indeed admit of all degrees .... But ... the same person is perfectly the same, and cannot be so in part, or in some degree only .... [Identity] is the foundation of all rights and obligations, and of all accountableness; and the notion of it is fixed and precise.

In this passage there are at least four arguments stated or implied. The first is the argument that thoughts require a thinker, which I have already discussed. The second is the argument from memory. To this Hume can reply that memory does not give us evidence of a persisting self, but merely evidence that an event took place in the past which is causally connected with the present memory. If I

remember that I took part in the event - for example, that I conversed - this is of course different from remembering only that a conversation took place. But the difference, Hume could say, is only that in the first case there is an awareness of the existence of a particular chain of continuous and causally connected events between the original conversation and the memory; whereas in the second case the chain of events extends between merely witnessing, or being told about, the conversation through to the memory. If I remember correctly that I had a particular conversation twenty years ago, it follows that the conversation and the memory are both parts of my life history, and therefore that they are parts of the same chain of events. It does not follow, however, that a permanent 'self' has persisted all through this chain of events, any more than it follows that because the Mersey at Stockport and the Mersey at Liverpool are parts of the same river there must be a single molecule of water stretching from Stockport to Liverpool.

The third argument to be extracted from this passage is the one implied by the words 'The identity of a person is a perfect identity, etc.'. The argument here is that the only intelligible view of personal identity is that it is either a total identity or it is no identity at all. A person changes vastly from the cradle to the grave, both physically and mentally; this is obvious enough. The physical changes, according to Reid, do not present a problem: provided that the process of change is sufficiently gradual, it does not prevent us from ascribing identity through the changes to a body, organic or inorganic. But the mental sphere is different: for it to be true that the same person has had two experiences at different times, it must be the case that there was an unchanging subject of experience that persisted from the one time to the other.

The pre-philosophical 'intuitive' notion behind what Reid says has been very clearly formulated by Derek Parfit in an Open University radio programme.[2] It is that 'what-

ever happens between now and some future time, either I shall still exist or I shan't. And any future experience will either be my experience or it won't.' To take a common illustration[3], suppose someone tomorrow is going to feel intense pain. It seems intuitively clear that either it is I who will feel the pain, in which case I have cause to be afraid, or it is someone else, in which case I have nothing to fear, but good reason to feel pity. Either I feel the pain or I do not: it seems absurd to suppose that someone will feel the pain, but that it is an arbitrary decision whether that someone is me.

Yet this is precisely the conclusion to which we are led by a Humean view of the self. If a person is merely a chain of causally connected events, whether physical, mental or both, it follows that there is no absolute distinction between one person and another. In practice, the obvious thing to do is to regard the bodily continuity from birth to death as constituting one person; but on Hume's analysis it still is in the case that, however sensible this may be, it is not logically necessary - we would, if we wished, regard the child and the adult as two altogether separate people, or the person under 35 and over 35 as two people.

This is further brought out if we consider the various theoretical and practical puzzle cases. For example, Buddhists believe both that the self is a chain of events and that there is 'reincarnation', i.e. that the chain continues through death to a new birth. But there is no 'right' answer to the question whether the new-born baby is the same person as the man who died: he is the same in the sense that he is part of the same causal chain, but he is not the same if we wish to regard the events before and after death as belonging to separate persons. (Similarly, there is no 'right' answer to the question whether the Tame is part of the Mersey or is a separate river which eventually flows into the Mersey.)

Again, there are some celebrated cases of multiple personality, in which there is bodily continuity, but the body appears to be inhabited by two or more personalities, with different skills and character traits, and, most important of all, different sets of memories.[4] On Reid's view this conflict of evidence - the evidence from memory points one way, and the evidence from physical continuity the other - makes it very hard to decide whether we are dealing with one person or with two people sharing one body. But there must be a right answer: either there is one person, whose memory is functioning abnormally, or there really are two people. On Hume's view, there is no correct answer: we must simply decide whether it is more just, or more useful, to consider ourselves as dealing with one person or two.

It is very difficult to decide who is right. If we take a subjective view, we are likely to agree with Reid, on the ground that 'it will either be my experience or it won't'. If we take an objective view we may well agree with Hume that there is no perfect identity either psychologically or physically: there will be future states which will be causally connected with the present stream of events which I call 'myself', some of which will include memories of these events; but there is no clear distinction between those which are sufficiently closely connected to be regarded as part of 'myself' and those which are not. In other words, for a Humean the belief that the future will contain a 'me' is an illusion, fostered by the nature of language, and also, according to the Buddhists, by the psychological desire for survival. The future 'contains' only events, some of which have a close connection with the events that constitute 'me' and some of which have a distant one; and the difference between 'the same man' at different periods of his life is the same kind of difference, except in degree, as the difference between one man and another.

The problem now is how to decide whether or not the

subjective view is an illusion. For even if the Buddhists are right in drawing attention to the emotional reasons for believing in a 'self', this does not prove that the belief is false. Nor do the 'puzzle cases' show that there is no such thing as perfect identity: they only show that, as Reid himself says, neither the presence nor absence of memory nor the presence or absence of bodily continuity can give us more than evidence as to whether A and B are or are not the same person.

There is, however, one sort of puzzle case, discussed by Parfit in the articles referred to above, which seems to give support to Hume against Reid, although it is a case which mercifully belongs to science fiction. It is possible to separate the two halves of the brain, as is sometimes actually done for medical reasons; and experiments have shown that under certain circumstances this creates two centres of consciousness. If one, or both, halves of the brain were then transplanted into another body, the two centres could function independently, and there would be two people where formerly there was one: this is logically quite conceivable, though likely to be impossible in practice! Moreover, there is evidence that memory is coded in several parts of the brain, so that the memories of these two people up to the time of brain bisection would be identical. Person A has now 'split' into two people, B and C.[5]

The fact that all this is logically possible is, on Humean principles, what one would expect: there is nothing unreasonable in the 'division' of a chain of events, which means only that two sets of events have the same causal relationship to a third. But if Reid is right, there is a problem, since if one assumes the perfect identity of the self one has no reasonable way of describing the situation. One cannot say that B and C are <u>both</u> identical with A, since this is self-contradictory; one cannot say they are parts of A, which is excluded by the hypothesis. To say that A

is only one of B or C is not a contradiction, but is 'wildly implausible ... because my relation to each of the resulting people is exactly similar.' (Parfit, op. cit., p. 56) Finally, to say that A has ceased to exist is 'grossly misleading' because what has happened is quite different from death: 'Supposing ... only one of the transplants succeeds and only one of the resulting people ever comes to consciousness again ... we would all think that I would survive as that one person .... [W]here both operations succeed ... it's clearly absurd to suppose that a double success is a failure.' (ibid., p. 57) It seems that to maintain the commonsense position one must show either that what has been described is logically, and not merely physically, impossible or that there will always be some difference between the relation of B to A and that of C to A. Unless this can be done, the Humean position can be defended against Reid's attack by maintaining that close examination of the objective facts and possibilities shows that the subjective feeling of 'either it will be me or it won't' is illusory. Nevertheless, I think that this argument is the most powerful of Reid's objections; and the last word has probably not yet been said.

The last of Reid's criticisms is given in the final sentence of the quoted passage. It is the argument that some of our essential moral notions - those of right, duty and responsibility - depend on the supposition that there are moral agents who persist through time. A person can only claim a right if he is the same as the person on whom the right was conferred; he only has an obligation to perform a duty, - e.g. to keep a promise - if he is the same as the person who undertook that duty; he can only be justly punished if he is the same as the person who committed the offence for which punishment is being exacted. Unless there is such a thing as strict personal identity, these notions have no application; and some of our most basic moral concepts are misconceived.

The best answer to this, for the Humean, is probably to treat rights and obligations as 'inheritable'. Rights and obligations can after all be transferred from one person to another, as happens when someone acquires both someone else's property and his debts. Similarly, although a 'person' at any one time is not identical with 'himself' at any earlier time, he is the result of 'his' earlier life, because present events are the result of past events. This is so whether the thesis of determinism is true or false; while the earlier events may or may not be sufficient conditions for the later ones, they are certainly necessary conditions for their taking place in the precise way that they do. Because of this close connection between the states that make up a person, it does not seem in any way unjust that later states should 'inherit' the rights and obligations of earlier ones. If only the rights, or only the obligations, were inherited, injustice would be done; but if both are passed on, this seems as fair as the inheritance by a son of both his father's assets and his father's debts.

To this it may be replied that rights and obligations are not in this respect on the same level. Rights in general arise either through a person's status, such as being human, being a British subject, etc., or through being conferred by, for example, the making of a promise or the transfer of property: arguments about rights are typically about what precisely follows from particular sorts of status or acts of conferring. Both these can be passed on from one stage of a person to later stages, provided that, as is often the case, the later stages have the same relevant status as the earlier ones, and/or that the person conferring the right intends to confer it for a period of time, i.e. intends it to be 'inherited'.

But obligations, it may be said, can only be justly created by the voluntary act of the individual. One person cannot bind another: if the later person is not the same

as the earlier person, he cannot be bound by the earlier person's responsibilities. It may be true that if he wishes to inherit the rights he is obliged to inherit the responsibilities as well; but he must, as a matter of justice, have the option of rejecting both. It is one thing to say that if the son inherits his father's money he must also take on his father's debts; it is another matter to tell him that he must have both the money and the debts, and is not permitted to refuse both.

One may answer this point by denying that obligations are necessarily based on voluntary binding of oneself. If we follow our ordinary moral intuitions, we find two obvious exceptions - the obligation towards members of one's family (one chooses a spouse, and chooses to have children, but one does not choose one's parents or siblings), and the obligations to help people in trouble (for example, to rescue a drowning person, if one can swim), which is based simply on one's being, without prior intention, the only person in a certain place at a certain time. If simply being in a certain situation can create a duty, why not simply being the 'successor' to someone who took on a responsibility?

Alternatively, one might argue that these obligations are the result of a voluntary but tacit reciprocal agreement among people to give each other help in times of crisis. To this it might be objected that, apart from the difficulty of the notion of a tacit agreement, someone could with perfect justice, opt neither to give nor to receive help. Both these points can be countered by Kant's argument in his discussion of the fourth example of the categorical imperative, in which he argues that the maxim never to give or receive help is one that a rational being cannot adopt.[6] Kant presupposes here that a rational being necessarily has purposes of some sort, whatever they may be. Now an imperfect rational being, such as a man, cannot guarantee that he can always achieve his purposes on his

own; he cannot be sure that he will not one day need the help of others, and indeed he can be almost sure, given his imperfection, that he will. Hence he cannot rationally adopt a principle that he should never receive help: this would be to will the end (whatever this end may be) without willing the means. Somewhat analogously one might say that a world in which people did not 'inherit' rights and duties from their previous selves would be a world in which their behaviour towards each other would be so unpredictable that the very unpredictability would interfere across the board with purposive behaviour; and hence no purposive being could wish this to be adopted. Obligations must be inherited if purposive behaviour, and hence morality, is to be possible: this is a sufficient justification, and gives no one a ground for complaint - one can only 'opt out' if one refuses to act morally at all.

So much for rights and obligations, but what about the problem of punishment? Even if one rejects the notion that the purpose of punishment is retribution - as I believe we should - it is notorious that our 'moral intuitions' still incline us to say that only the guilty should be punished, or at least that punishment of the innocent is always unjust, even though it may sometimes be necessary. But if the one being punished is never identical with the criminal, as follows from Hume's thesis, all punishment is unjust. This seems an extremely odd conclusion: one could maintain, rightly or wrongly, that all punishment was cruel, or useless, or harmful, but, except on determinist grounds, it is strange to maintain that it is always unjust or undeserved.

To answer this problem we need to look more closely at the notion of guilt; and a possible solution can be reached by revising our ideas about it. We suppose that a person becomes morally guilty by having done something wrong; and we may add to this the belief that he remains guilty, simply because of what he has done, until either

punishment or an act of contrition 'wipes out' the offence. But one may suggest that a person is guilty, not because of what he has done, but because, and as long as, he remains the kind of person who would do it again under similar circumstances - it is only in this sense that one can speak, metaphorically, of a 'stain' of guilt, a phrase which seems intelligible only if it refers to a persisting fault of character. While this fault lasts, the person who has committed an offence, even though he has changed in other ways, can be said both to 'need' punishment, if its aim is reform, and to have no ground for complaint if he is punished in order to deter others: 'his' state now may be different from 'his' earlier state, in which the offence was committed, but the difference is not relevant to his guilt.

If we take this view of guilt, and we believe, rightly or wrongly, that punishment can reform or deter, we need not regard the institution of punishment as being in principle unjust. There will be particular injustices, since the law can take into account only the fact that an offence has been committed, and can use the apparent sincere repentance of the offender only as grounds for reducing the penalty. But these injustices must arise in any legal system: 'the thoughts of men are not triable', and the law, even given that this definition of guilt is the right one, would have for its purposes to define guilt as breach of the law.

Moreover, I think that this account accords with a number of moral intuitions. It accords with the idea that 'real' repentance consists in reformation of character, and that if and when this takes place the guilt is removed; even in the eyes of the law the fact that a person is extremely unlikely to commit an offence again is taken as grounds for reducing, though not eliminating, his punishment. It also accords with the belief that the justice of reward and punishment diminishes with time: to punish an offence, say, twenty or thirty years after it was committed looks like

vindictiveness rather than justice, and to reward a benefit after so long a time is to go beyond fitting gratitude to positive generosity. (Parfit, in the article already cited, lays a particular stress on this consequence of a no-self theory; it is worth making the point that it is not as much at variance with ordinary moral ideas as it at first appears, since we do, I think, feel that to punish after a long lapse of time could be, though it need not be, almost to punish a different person from the offender.)

If we now review all these various arguments of Reid's, I think we shall find that three of them are crucial. There is the argument that we cannot describe human behaviour intelligibly without the notions of an act and an agent. There is the argument that the identity of persons is a perfect identity, i.e., in Parfit's words, that 'any future experience will either be my experience or it won't'. Thirdly, there is the argument that there are basic moral concepts - shared by all moral systems, and not peculiar to some only - which do not make sense unless we suppose that there are persisting moral agents. I have suggested that a Humean can make a reasonable reply to all these objections, and to some other less weighty ones; but whether these replies totally dispose of Reid's arguments I am still uncertain. What can be safely asserted is that both Reid's case and Hume's are stronger than writers on personal identity have sometimes recognized; that the last word on this topic has not been said; and that therefore more hard thinking is needed!

Harry Lesser
The University of
Manchester.

1. See I. Hacking, 'A language without particulars', Mind, vol. 77, 1968, which cites the relevant work of Boas, Whorf and Sapir.

2. 'Brain transplants and personal identity', ch. 5 of

Philosophy in the Open, ed. G. Vesey, Open University Press, Milton Keynes, 1974, pp. 54-65. See also D. Parfit, 'Personal identity', Philosophical Review, vol. 80, 1971, pp. 3-27.

3. For the development of this example, see Bernard Williams, 'The self and the future', Philosophical Review, vol. 79, 1970, reprinted in Williams, Problems of the Self (Cambridge University Press, 1973) ch. 4.

4. See, for example, M. Prince, The Dissociation of a Personality (New York: Longmans, Green & Co. 1905).

5. For the neurophysiological background, Vesey cites M. S. Gazzaniga, The Bisected Brain (New York: Appleton-Century-Crofts, 1970) and S. Rose, The Conscious Brain (Weidenfeld and Nicolson, 1972).

6. Groundwork of the Metaphysic of Morals, 423. For this interpretation of Kant, see P. Dietrichson, 'Kant's criteria of universalizability', in R. P. Wolff (ed.), Kant: Foundations of the Metaphysic of Morals: text and critical essays (New York: Bobbs-Merrill, 1969), pp. 189-98.