

Chapter 6: Locke's account of our personal identity

In a chapter on identity and diversity, Locke makes many ingenious and sound observations, and some that I think can't be defended. I shall confine my discussion to his account of our own personal identity. His doctrine on this subject has been criticized by Butler in a short essay appended to his *The Analogy of Religion*, an essay with which I complete agree.

As I remarked in chapter 4, identity presupposes the continued existence of the being whose identity is affirmed, and therefore it can be applied only to things that have a continuous existence. For as long as any being continues to exist, it is the same being; but two beings that have different beginnings or different endings of their existence can't possibly be the same. I think Locke agrees with this.

He is absolutely right in his thesis that to know what is meant by 'same person' we must consider what 'person' stands for. He defines 'person' as *a thinking being endowed with reason and with consciousness*—and he thinks that consciousness is inseparable from thought.

From this definition it follows that while the thinking being continues to exist, and continues thinking, it must be the same person. To say that

- the thinking being is the person,
- and yet that
- the person ceases to exist while the thinking being continues,
- or that
- the person continues while the thinking being ceases to exist,
- strikes me as a manifest contradiction.

One would think that the definition of 'person' would completely settle the question of what the nature of personal identity *is*, or what personal identity *consists in*, though there might still remain a question about how we come to *know* and *be assured of* our personal identity. But Locke tells us:

Personal identity, i.e. the sameness of a rational being, consists in consciousness alone; and as far as this consciousness can be extended backwards to any past action or thought, so far reaches the identity of that person. So that whatever has the consciousness of present and past actions is the same person to whom they belong. [Adapted by Reid from II.xxvii.9; the main difference is that Locke wrote 'is the same self' etc.]

This doctrine has some strange consequences that the author was aware of. For example: if the same consciousness could be transferred from one thinking being to another (which Locke thinks we can't show to be impossible), then **two or twenty thinking beings could be the same person**. And if a thinking being were to lose the consciousness of the actions he had done (which surely is possible), then he is not the person who performed those actions; so that **one thinking being could be two or twenty different persons** if he lost the consciousness of his former actions two or twenty times.

Another consequence of this doctrine (which follows just as necessarily, though Locke probably didn't see it) is this: A man may *be* and at the same time *not be* the person that performed a particular action. Suppose that a brave officer

- was beaten when a boy at school, for robbing an orchard,

- captures an enemy standard in his first battle, and
- is made a general in advanced life.

Suppose also (and you have to agree that this is possible) that when he took the standard he was conscious of his having been beaten at school, and that when he became a general he was conscious of his taking the standard but had absolutely lost the consciousness of his beating.

Given these suppositions, it follows from Locke's doctrine that he who was beaten at school is the same person who captured the standard, and that he who captured the standard is the same person who was made a general. From which it follows—if there is any truth in logic!—that

- the general is the same person as him who was beaten at school.

But the general's consciousness does not reach so far back as his beating, and therefore according to Locke's doctrine

- the general is not the person who was beaten.

So the general *is* and at the same time *is not* the person who was beaten at school.

Leaving the consequences of this doctrine to those who have leisure to trace them, I shall offer four observations on the doctrine itself.

[Before Reid does that, the preparer of these texts ventures to intrude as a commentator. Reid offers the foregoing argument so confidently, and to so much applause from others in later centuries, that its lack of charity towards Locke should be pointed out. The tone and tenor of Locke's Identity chapter are compatible with this possibility:

- (1) Locke was thinking in terms of *sufficient* conditions for personal identity.
- (2) He silently assumed that elementary logic was to be built into his account, so that *one* sufficient condition for the truth of 'x is z' would be 'x is y & y is z'.
- (3) His 'same consciousness' account was meant to present not *everything* that could make a statement of personal identity true but everything that (a) could make such a statement true and (b) *didn't itself presuppose the truth of any such statement*.

This interpretation would be ruled out only if he explicitly and persistently said that his 'same consciousness' stuff was meant, on its own, to give sufficient *and necessary* conditions for the truth of 'x is z'. But he did no such thing. Why didn't he *say* that elementary logic—i.e. the transitivity of identity—was being assumed? He shouldn't have needed to; in a fair-minded world it would go without saying.]

(1) Locke attributes to *consciousness* the conviction we have of our past actions, as if a man could now be *conscious* of what he did twenty years ago. It is impossible to make sense of this unless 'consciousness' means *memory*, the only faculty by which we have an immediate knowledge of our past actions.

Sometimes in informal conversation a man says he is 'conscious' that he did such-and-such, meaning that he distinctly remembers that he did it. In ordinary everyday talk we don't need to fix precisely the borderline between consciousness and memory. . . . But this 'imprecision' ought to be avoided in philosophy—otherwise we run together different powers of the mind, ascribing to one what really belongs to another. If a man can be 'strictly and literally' *conscious of* what he did twenty years or twenty minutes ago, then there is nothing for memory to do, and we oughtn't to allow that there is any such faculty. The faculties of •consciousness and •memory are chiefly distinguished by this: •consciousness is an immediate knowledge of the present, •memory is an immediate knowledge of the past.

So Locke's notion of personal identity, stated properly, is that personal identity consists in clear remembering. . . .

(2) In this doctrine, not only is •consciousness run together with •memory, but (even more strange) •personal identity is run together with •the evidence we have of our personal identity.

It is very true that my remembering that I did such-and-such is the evidence I have that I am the identical person

who did it. And I'm inclined to think that this what this is what Locke meant. But to say that my remembering that I did such-and-such, or my consciousness that I did it, *makes me* the person who did—that strikes me as an absurdity too crude to be entertained by anyone who attends to the meaning of it. For it credits memory or consciousness with having a strange magical power to produce its object, though that object must have existed before the memory or consciousness that ·supposedly· produced it.

Consciousness is the testimony of one faculty; memory is the testimony of another faculty. To say that •the testimony is the cause of •the thing testified is surely absurd if *anything* is absurd, and Locke couldn't have said it if he hadn't confused the testimony with the thing testified. . . .

(3) Isn't it strange that the sameness or identity of a person should consist in something that is continually changing, and is never the same for two minutes?

Our consciousness, our memory, and every operation of the mind are still flowing like the water of a river, or like time itself. The consciousness I have this moment can't be the same consciousness that I had a moment ago, any more than this moment can be that earlier moment. Identity can only be affirmed of things that have a continuous existence. Consciousness and every kind of thought is passing and momentary, and has no continuous existence; so if personal identity consisted in consciousness it would certainly follow that no •man is the same •person any two moments of his life; and as the right and justice of reward and punishment is based on personal identity, no man would be responsible for his actions! But though I take this to be the unavoidable consequence of Locke's theory of personal identity, and though some people may have liked the doctrine the better on this account, I am far from imputing anything of this kind to Locke himself. He was too good a man not to have

rejected in horror a doctrine that he thought would bring this consequence with it.

(4) In his discussion of personal identity, Locke uses many expressions that I find unintelligible unless he wasn't distinguishing •the sameness or identity that we ascribe to an individual from •the identity which in everyday talk we ascribe to many individuals of the same species.

When we say that pain and pleasure, consciousness and memory, are *the same* in all men, this 'same'ness can only mean *similarity*, i.e. sameness of kind. ·If it meant individual identity, i.e. identity properly and strictly so-called, it would be implying· that the pain of one man could be the same individual pain that another man also felt, and this is no more possible than that one man should be another man; the pain I felt yesterday can no more be the pain I feel to-day than yesterday can be today; and the same thing holds for •every operation of the mind and •every episode of the mind's undergoing something. The same kind or species of operation may occur in different men or in the same man at different times, but it is impossible for the *the same individual operation* to occur in different men or in the same man at different times.

So when Locke speaks of 'the same consciousness being continued through a succession of different substances', of 'repeating the idea of a past action with the same consciousness we had of it at the first' and 'the same consciousness extending to past and future actions', these expressions are unintelligible to me unless he means not the same individual consciousness but a consciousness that is of the same kind.

If our personal identity consists in consciousness, given that consciousness can't be •the same individually for any two moments but only •of the same kind, it would follow that we are not for any two moments the same *individual* persons but the same *kind* of persons.

As our consciousness sometimes ceases to exist—as in sound sleep—our personal identity must cease with it, according to Locke's theory. He allows that a single thing

can't have two beginnings of existence; so our identity would be irrecoverably lost every time we stopped thinking, even if only for a moment.