

Locke and Personal Identity

Locke denied that we have any knowledge of the real essence of substances. He had no doubt as to the existence of material substance and he believed that Descartes and others were right in their claim that the mind is a thinking, non-physical substance. He believed that substances have powers to produce ideas in our minds and that these ideas represent substances to us. These ideas of substances inform us of the existence of substances and give us some information as to some of their characteristics but they do not tell us the real essence of these substances. If they did, there would be no need for science to invent theories about the real essences of material substances. Of material substances we only know their bulk, figure, texture, weight, motion, etc.. We cannot know what it is that has such qualities. That the mind is a thing that thinks there is no doubt and we have immediate knowledge of our existence as thinking beings, just as Descartes said we do. We may believe that the mind is an immaterial substance as Descartes did, but it is impossible in this life, according to Locke, to know that it is an immaterial substance. This result follows from his general claim that we cannot know the real essences of any substances. He scandalized many by being prepared to admit that it was not self-contradictory to suppose that God might have given a material thing the power of thought. Whether Descartes could have admitted such a possibility is a good question. When the mind turns its attention to observing itself, attending to its own operations of thinking, perceiving, feeling, willing, etc. it receives ideas of these operations, called ideas of reflection. In sharp contrast to Descartes, Locke says the mind has no idea of its own real essence. For these and other reasons, Locke abandons numerical identity of immaterial substance as the criterion of personal identity.

One issue that Locke thinks is important to deal with in finding a better criterion of personal identity is to distinguish between the concept of a human being (man) and the concept of a person. It is not necessary that all human beings are persons or that all persons are human beings. The conditions of the identity of a human being, Locke says, are much like the conditions of the identity of any animal or plant.

“That being then one plant which has such an organization of parts in one coherent body, partaking of one common life, it continues to be the same plant as long as it partakes of the same life, though that life be communicated to new particles of matter vitally united to the living plant in a like continued organization, conformable to the sort of plants.”

“This also shows wherein the identity of the same man consists, viz in nothing but a participation of the same continued life, by constantly fleeting particles of matter, in succession vitally united to the same organized body for if the soul alone makes the same man and there be nothing in the nature of matter why the same individual spirit may not be united to different bodies ... it will be possible that the souls of men may be detruded into the bodies of beasts..... But yet I think nobody, could he be sure that the soul of Heliogabalus were in one of his hogs would yet say that hog were a man or Heliogabalus”

“ ...whoever should see a creature of his own shape or make, though it had no more reason all its life than a cat or a parrot would call him still a man. For I presume it is not the idea of a thinking or rational being alone that makes the idea of a man in most people’s sense, but of a body so and so shaped joined to it...

“... what person stands for ... is a thinking intelligent being, that has reason and reflection and can consider itself as itself, the same thinking thing , in different times and places. Which it does only by that consciousness which is inseparable from thinking and ... essential to it. For since consciousness always accompanies thinking and it is that which makes everyone to be what he calls self... in this alone consists personal identity ... and as far as this consciousness can be extended backwards to any past action or thought, so far reaches the identity of that person.”

“But as to the question of whether , if the same thinking substance (supposing immaterial substances only to think) be changed, it can be the same person, I answer that cannot be resolved but by those who know what kind of substances they are that do think ... If the same consciousness can be transferred from one thinking substance to another, it will be possible that two thinking substances may make but one person. For the same consciousness being preserved, whether in the same or different substances, the personal identity is preserved. ... whatever has the same consciousness of present and past actions is the same person to whom they both belong”

“This may show us wherein personal identity consists: not in the identity of substance, but ... in the identity of consciousness. If it be possible for the same man to have distinct incommunicable consciousness at different times, it is past doubt the same man would at different times make different persons”

Locke is saying that a monkey or a bird could conceivably be a person, as long as it was a thinking, intelligent being with reason and reflection. Still, it would not be a human being but an intelligent monkey or bird; persons of different species. Similarly, if a creature with the same shape as our own, though it had no more reason than a monkey or a parrot would still be a human being but a very dull one and, if dull enough, not a person at all.

Another sort of case we might imagine in which most of a man’s body was destroyed and replaced by bionic parts might be one in which we were justified in saying that the same person existed but no longer as a human being. According to Locke, this would be true of the person if he remembered most of what he was conscious of before the accident.

What follows from this is that, if he had permanently lost all memory of everything that happened to him in his life before the accident, the same person would no longer exist. In this case, the same man no longer exists either. It would be a bionically embodied new person, metaphysically just like a newborn baby. It also follows, according to Locke, that

this person would not be legally or morally responsible for anything that the person he succeeded had done.

To borrow a phrase of Epicurus, we could say that, in this case, “the chain of identity has been snapped”. The crucial change, according to Locke, is that the sorts of psychological states that are normally continuous in a person from moment to moment have been dramatically altered by the total permanent amnesia. Even if some of the same general personality features and dispositions remained, it would be incorrect, according to Locke, to say that the same person still exists. There is no longer the sort of strong psychological connectedness between the person who once existed and the person who has taken his place. Following Carter (and Parfit), let us define this sort of *strong psychological connectedness* as follows:

A, who exists now, is strongly connected psychologically to B, who existed before, if and only if A remembers doing things only B did.

This seems to mean that, for Locke, reincarnation is not metaphysically impossible but one would have to remember most of what happened in a past life to be the person who lived once in another body.

We should be cautious as to exactly what claim we are attributing to Locke. If Locke were only saying that strong psychological connectedness is a sufficient condition for personal identity then he would be making a much weaker claim than if he were saying that strong psychological connectedness is a necessary condition of personal identity. If Locke were to agree that identity of consciousness is the same as our above definition of strong psychological connectedness then he would be saying that this connectedness is both necessary and sufficient for personal identity.

The importance of this distinction comes out when we consider an important criticism made of Locke by Thomas Reid. Suppose that strong psychological connectedness is necessary for personal identity. Reid interprets Locke as saying this. Now, Identity is a transitive relation. This means that if $A=B$ and $B=C$ then it necessarily follows that $A=C$. Suppose then that B is the same person as A and B is the same person as C. It must follow that A is the same person as C. But is this guaranteed for all cases in which A is the same person as B and B is the same person as C on Locke’s criterion of personal identity?

The story of the brave officer seems to provide an example in which logically sufficient conditions obtain for deducing the conclusion that the boy and the general are the same person, assuming Locke’s criterion of personal identity. But we know there are cases in which a person would not remember enough of the right things about what they once did or experienced to satisfy Locke’s criterion for being the same person.

Suppose we were to read Locke’s criterion of personal identity as saying that sameness of memories (consciousness) is sufficient but not necessary for personal identity. Perhaps

that would save him from Reid's criticism. He seems to allow the possibility that the same consciousness might be transferred to a different substance. If I lost all of my present memories and they were somehow transferred to a new mind (freshly created by God, perhaps) then that new mind would be me, even if the substance that formerly had those memories still existed but all the memories it once had had been destroyed. In fact, Locke even allows the possibility that two numerically different substances might have the same memories and then both of those thinking substances would be one person. There would be, in Locke's terms, two men but each would be the same person. Does this make sense?

To many people this seems just wrong. Even if there could be an exact duplicate of a person, they say, only one of them would be the person in question. Suppose this objection is right. We may suppose it possible that Jane and her twin Joan have all the same memories but it would still seem that if Jane existed first then she is one person and Joan is another, similar, person. Yet how, according to Locke, do Jane's memories guarantee that she is not the same person as Joan? Locke seems to admit that sameness of memories would be sufficient for personal identity. Though they might have different names and different bodies, Jane and Joan would be one person if they had all the same memories. Or, so Locke's theory seems to imply.

One philosopher named Joseph Butler, who lived about the same time as Locke, thought that Locke's theory of personal identity contained a serious fallacy. He was sympathetic with Locke's theory up to a point, saying

"... by reflecting upon that which is myself now and that which was myself twenty years ago, I discern that they are not two but one and the same self." Butler then continues, "But though consciousness of what is past does ascertain our personal identity to ourselves, yet to say that it makes personal identity, or is necessary to our being the same persons, is to say that a person has not existed a single moment nor done one action but what he can remember; indeed none but what he reflects upon. And one should really think it self-evident that consciousness of personal identity presupposes and therefore cannot constitute personal identity..."

What Butler meant by this last remark is that, if Locke is defining the sort of remembering that is the basis of personal identity so that such an act of remembering is one in which I correctly remember myself having done something, then Locke's criterion of personal identity is circular. If Jane's memories are what make her the person she is but these memories guarantee her identity because they are memories of her past self, then Locke's criterion of personal identity presupposes personal identity and cannot define it. Clearly, we must distinguish between seeming to remember and really remembering. Seeming to remember being present at the battle of Waterloo does not guarantee that one was actually there. So, seeming to remember doing things that only Napoleon did is no guarantee that one is Napoleon.

All of this is related to Locke's claim that personal identity is not a matter of the numerical identity of a substance but the sameness of consciousness. This seems to

mean that the notion of a substance plays no role whatever in Locke's theory of personal identity. A person might just as well not be a substance at all, for Locke. One philosopher who noticed this feature of Locke's philosophy was David Hume. Hume applied this point in an interesting way. He challenged the notion that the mind is any sort of substantial thing. In his *Treatise of Human Nature*, Hume wrote

... we may observe that what we call a mind is nothing but a heap or collection of different perceptions, united together by certain relations and suppos'd, tho' falsely, to be endowed with a perfect simplicity and identity.

For my part, when I enter most intimately into what I call myself, I always stumble on some particular perception or other ... I never catch myself at any time without a perception and never can observe anything but the perception. When my perceptions are removed for any time, as by sound sleep, so long I am insensible of myself and may truly be said not to exist. And were all my perceptions removed by death I should be entirely annihilated.

Hume is saying that when we attempt to examine our own minds, to observe ourselves as thinking things, all that we encounter is a collection or bundle of psychological states: ideas thoughts, emotions, memories and sensations. We do not encounter any substantial thing, a continuant which persists through time identically. The mind or self is not an immaterial substance, as Descartes claimed, nor a material substance (such as the brain) because it is not a substantial thing of any kind. We ought to think of a mind as being a temporally extended process like a game or a concert or a storm, only it is a mental process as opposed to these physical sorts of processes. Like other processes, we can speak of its temporal parts as well as of its momentary parts and, as with any process, the sum of the mind's temporal parts cannot exist all at once. As a result, Hume argued that what Butler called the "strict and philosophical sense" of identity cannot reasonably be applied to minds any more than it can be applied to trees or ships. We speak in a "loose and popular" sense of 'same' when we say that the tree now before us is the same one that stood here twenty years ago, Butler reminds us. There is, he said, nothing illegitimate about the loose and popular sort of identity. The criteria for identity in the loose and popular sense are different from one sort of thing to another and we should not fall into the mistake of confusing the one sort of identity with the other. Butler thought that the strict and philosophical sense of identity did apply to persons as opposed to their bodies. Hume denied that the strict and philosophical sense of identity applied to minds or selves. Since they are only bundles of perceptions, the criteria for their identity must be as loose and popular as the criteria for any complex particular such as a ship, a city or a thunderstorm.

One important objection to this conception of the mind or self has been made by R. M. Chisholm. In his article "On the Observability of the Self", Chisholm writes

Our idea of a mind or self... is not an idea only of particular perceptions. It is not an idea of love or hate and of heat or cold. It is an idea of that which loves or hates and of that which feels warm or cold (and of course of much more besides). That is to say, it is

an idea of an x such that x loves or x hates and such that x feels cold or x feels warm, and so forth.

Just as there cannot be a smile or a frown without someone who smiles or who frowns, so there cannot be love, hate, pleasure or pain unless there is someone who loves, someone who hates, someone who feels pleasure or pain. Psychological states cannot exist without a subject; so, the subject cannot be identified with any collection of psychological states. Whenever we observe a perception we are, by that very act, observing the one who is in that perceptual state. To observe hunger, for example, is to observe oneself feeling hungry.

There may be a way of replying to Chisholm that would acknowledge the strength of his argument but preserve the essence of Hume's view. Locke and Descartes identify the self or person with something other than the human body. Suppose that Hume were to say that a person is something that has two parts, one being a human body, the other a mind. Let Hume's conception of the mind remain as he describes it, a collection or bundle of psychological states. Then there will be something to serve as the subject of psychological states, namely a certain human body. This body, permanently without psychological states, would no longer be a person, though its basic life processes might go on for a while. To observe hunger will be, as Chisholm insists, to observe oneself feeling hungry but then the hungry self that one is observing is *one's own body* in a state of hunger.

Perhaps we still ought to have reservations about this conception of the mind as a collection of psychological states, even if we allow that it is the body that is the subject of these states. Even though we may not be satisfied with Locke's or Descartes' theories of personal identity, it is hard to deny their basic premise that who one is as a person is much more closely related to one's psychological makeup than to one's physical makeup. Even if a person is both a mind and a body we should not assume that just any conception of a mind will do the job needed for a good theory of personal identity.

To see how a problem might develop with Hume's conception of a mind, consider the sort of case that Locke thought was possible, namely a transfer of consciousness from one subject to another. We are supposing certain material substances, human bodies, to have minds, i.e., collections of psychological states. When they have such psychological states, human bodies are persons. Suppose that God transfers ownership of the psychological states that constitute Jean's mind to Jill and conversely. On Hume's criteria of personal identity, even if we use a loose and popular sense of identity for tracking the identity of a single mind over time, it seems that the human body that was Jean's body is now Jill's body and the human body that was Jill's body is now Jean's body. Without changing any of their bodily attributes, it seems, personal identity has been switched, simply by switching minds. This might not seem to be any more of a problem for Hume than for any other philosopher but perhaps Hume's conception of the mind does generate problems of its own.

Hume considers the mind to be a collection or bundle of psychological states rather than a substance. Like other complex particulars, such as ships or baseball teams, their identity is affected by changes in their parts. On a strict criterion of identity, changes in the parts of a complex particular terminate the existence of that particular and mark the beginning of another particular. Other changes may not affect its identity. This was Locke's view.

“That being then one plant which has such an organization of parts in one coherent body, partaking of one common life, it continues to be the same plant as long as it partakes of the same life, though that life be communicated to new particles of matter vitally united to the living plant in a like continued organization, conformable to the sort of plants.”(Locke 1)

“This also shows wherein the identity of the same man consists, viz in nothing but participation of the same continued life, by constantly fleeting particles of matter, in succession vitally united to the same organized body for, if the soul alone makes the same man and there be nothing in the nature of matter why the same individual spirit may not be united to different bodies ... it will be possible that the souls of men may be detrued into the bodies of beasts..... But yet I think nobody, could he be sure that the soul of Heliogabalus were in one of his hogs would yet say that hog were a man or Heliogabalus” (Locke 2)

“whoever should see a creature of his own shape or make, though it had no more reason all its life than a cat or a parrot would call him still a man. For I presume it is not the idea of a thinking or rational being alone that makes the idea of a man in most people’s sense, but of a body so and so shaped joined to it...” (Locke 3)

“... what person stands for ... is a thinking intelligent being, that has reason and reflection and can consider itself as itself, the same thinking thing , in different times and places. Which it does only by that consciousness which is inseparable from thinking and ... essential to it. For since consciousness always accompanies thinking and it is that which makes everyone to be what he calls self... in this alone consists personal identity ... and as far as this consciousness can be extended backwards to any past action or thought, so far reaches the identity of that person.”
(Locke 4)

“But as to the question of whether , if the same thinking substance (supposing immaterial substances only to think) be changed, it can be the same person, I answer that cannot be resolved but by those who know what kind of substances they are that do think ... If the same consciousness can be transferred from one thinking substance to another, it will be possible that two thinking substances may make but one person. For the same consciousness being preserved, whether in the same or different substances, the personal identity is preserved. ... whatever has the same consciousness of present and past actions is the same person to whom they both belong” (Locke 5)

“This may show us wherein personal identity consists: not in the identity of substance, but ... in the identity of consciousness. If it be possible for the same man to have distinct incommunicable consciousness at different times, it is past doubt the same man would at different times make different persons” (Locke 6)

"Strong Psychological connectedness"

A, who exists now, is strongly connected psychologically to B, who existed before, if and only if A remembers doing things only B did.

"... by reflecting upon that which is myself now and that which was myself twenty years ago, I discern that they are not two but one and the same self." Butler then continues, "But though consciousness of what is past does ascertain our personal identity to ourselves, yet to say that it makes personal identity, or is necessary to our being the same persons, is to say that a person has not existed a single moment nor done one action but what he can remember; indeed none but what he reflects upon. And one should really think it self-evident that consciousness of personal identity presupposes and therefore cannot constitute personal identity..." (Butler)

"... we may observe that what we call a mind is nothing but a heap or collection of different perceptions, united together by certain relations and supposed, tho' falsely, to be endowed with a perfect simplicity and identity." (Hume 1)

"For my part, when I enter most intimately into what I call myself, I always stumble on some particular perception or other ... I never catch myself at any time without a perception and never can observe anything but the perception. When my perceptions are removed for any time, as by sound sleep, so long I am insensible of myself and may truly be said not to exist. And were all my perceptions removed by death I should be entirely annihilated." (Hume 2)

"Our idea of a mind or self... is not an idea only of particular perceptions. It is not an idea of love or hate and of heat or cold. It is an idea of that which loves or hates and of that which feels warm or cold (and of course of much more besides). That is to say, it is an idea of an x such that x loves or x hates and such that x feels cold or x feels warm, and so forth.

Just as there cannot be a smile or a frown without someone who smiles or who frowns, so there cannot be love, hate, pleasure or pain unless there is someone who loves, someone who hates, someone who feels pleasure or pain. Psychological states cannot exist without a subject; so, the subject cannot be identified with any collection of psychological states. Whenever we observe a perception we are, by that very act, observing the one who is in that perceptual state. To observe hunger, for example, is to observe oneself feeling hungry. " (Chisholm)

