

LOCKE ON CONSCIOUSNESS

Vere Chappell

***** DRAFT *****

upd 11.17.04

Locke took consciousness to be a property of the human mind. In this he was in agreement with other 17th-century philosophers, Descartes and Cudworth most notably. But Locke also used the word 'consciousness' to designate a kind of individual thing which persists through time and which can be distinguished from other individuals of its kind; and in this he was unique in the 17th century. This other consciousness is important in Locke's philosophy, because it is the identity of such consciousnesses that determines the identity of persons. The familiar consciousness is puzzling enough; but Locke's new one seems downright baffling.

In this paper I propose to examine both varieties of consciousness in Locke's philosophy, in an effort to dispel some of the mystification and puzzlement that often attend them.

I. Consciousness as a property.

If consciousness is a property, it is appropriate to ask what its subject is, what thing or things it is attributed to. I have already said that for Locke consciousness is a property of the human mind. But in the *Essay* we find Locke attributing consciousness not only to human minds but to souls, spirits, human beings (men), and immaterial substances. This is not such a disparate list as might appear. Locke often uses the terms 'mind' and 'soul' interchangeably; minds and souls constitute a species of spirit; and spirits are classified, ontologically, as immaterial substances. The most frequently mentioned subjects of consciousness in the *Essay*, are human beings or men, and the connection of minds or souls to persons, but the reason I think is simple. Persons are so obviously conscious that it is pointless ever to say that they are.

men is figurative, a matter of synecdoche, of which Locke makes abundant use. In one passage, he invokes the literal truth: criticizing the Cartesians, he says that "they who tell us, that the Soul always thinks, do never, that I remember, say, That a man always thinks. [But] can the Soul think, and not the Man?" (E II.i.19: 115).

It is noteworthy that Locke does not attribute consciousness to persons. In his chapter on personal identity, in which he more often speaks of consciousness than in any other chapter of the *Essay*, he makes a point of distinguishing persons from human beings, and he makes consciousness a defining feature of persons: a person, he says, "is a thinking, intelligent Being, that has reason and reflection, and can consider it self as it self, the same thinking thing in different times and places; which it does only by that consciousness, which is inseparable from thinking, and as it seems to me essential to it:" (E II.xxvii.9: 335). A human being, by contrast, is defined by its physical form or shape. A rational parrot, he says, would not be a man; Locke doesn't rule out its being a person. He does say on two occasions that selves are conscious, and in other passages he equates selves with persons "*Person* is the name for this *self* [sc. that a Man finds wherever he finds what he calls *himself*" (E II.xxvii.26: 346). But he refrains from ascribing consciousness to persons directly. "The more probable opinion", he says, is that "consciousness is annexed to, and the Affection of one individual immaterial Substance" (E II. xxvii. 25: 345). He is careful to avoid saying, however, that persons are such substances.

It is curious that Locke doesn't attribute consciousne

Another sort of thing that Locke does not attribute consciousness to is mental states and actions. Some recent

philosophers have distinguished between what they call 'creature consciousness', which is the consciousness possessed by men and minds and (perhaps) animals and robots, and 'state consciousness', which is the consciousness possessed by mental states and operations. In every passage in the *Essay* in which Locke specifies a subject of consciousness, the subject is a creature; in no passage does he attribute consciousness to a state or action of the mind. This is not a fact with much significance, however, since it seems that consciousness for states could be defined in terms of creature consciousness: a conscious state is merely one that some creature has or could have (or whose mind has or could have) and is or could be conscious of .

Of course it is a different question whether all creature consciousness is consciousness of mental states, or indeed is consciousness of anything at all. Another distinction made by recent philosophers is that between "transitive" and "intransitive" consciousness, which is a distinction between two forms of creature consciousness. A creature has transitive consciousness if its consciousness is intentional, or has an object; its consciousness is intransitive if it does not have an object. Consciousness without an object is what the OED describes as "the normal condition of healthy waking life" - as in "after several months in a coma, she suddenly regained consciousness". Locke appears not to have used the terms 'conscious' and 'consciousness' in this sense - which is not surprising if, as the OED suggests, this use did not appear until the nineteenth century. The first OED citation is a novel by Dickens that was published in 1837.

So for Locke all consciousness is transitive. The next question is, what sorts of things does he include among the objects of consciousness? He speaks of subjects both as being conscious *that* something - that they are thinking, for example - and of their being conscious *of* something - of their thoughts or perceptions, for example. Locke's grammar suggests, then, that consciousness

has objects of different kinds: on the one hand, those that are propositions or have a propositional structure, and on the other those that do not have such a structure. Objects of the one kind are true or false, obtain or do not obtain; those of the other exist or occur or else do not do so.

More significant, however, is the characterization contained in the closest thing Locke gives of a definition of consciousness in the *Essay*: "Consciousness is the perception of what passes in a Man's own mind" ((E II.i.19: 115). What are these things that pass in the mind? Locke's favorite general term for them is 'operations' of the mind. He means not only that everything that passes in a man's mind is an object of consciousness but also that only such things are (cf. Thomas Reid: "We know of nothing that is in the mind but by consciousness, and we are conscious of nothing but various modes of thinking" (Reid 2002: 322)). But he uses this term, as he says, "in a large sense, as comprehending not barely the actions of the Mind about its Ideas, but some sort of Passions arising sometimes from them, such as is the satisfaction or uneasiness arising from any thought" (E II.i.4: 105-6). In one well-known passage, he distinguishes two genera of mental operation: "*Perception*, or *Thinking* [on the one hand], and *Volition*, or *Willing*" on the other; the suggestion is that these two are mutually exclusive, and also that the disjoined terms in each pair are equivalent. But later we find Locke using 'thinking' to cover 'willing' (E II.nnn.nn: 000) and reporting that, though 'perception ... is by some called Thinking in general, ... in the propriety of the *English Tongue*, ['thinking'] signifies that sort of operation of the Mind about its *Ideas*, wherein the Mind is active; where it with degree of voluntary attention, considers any thing. For in bare naked *Perception*, the Mind is, for the most part, only passive; and what it perceives, it cannot avoid perceiving" (E II.ix.1: 143). Locke was not always concerned to use words consistently, or with precision.

Most mental operations for Locke, like consciousness itself, are intentional: they have objects. Some such operations, indeed, have two distinct objects, one outside the mind, one inside. In veridical sense perception, for example, the subject perceives an external, sensible thing, and also, concurrently, something internal. Locke's general term for the internal object of a mental operation is 'idea': to distinguish ideas from physically instantiated sensible qualities, for example, he often says that ideas are the "immediate object[s] of Perception, Thought, or Understanding" (E II.viii.8: 134). Locke holds that a few types of mental operation - itching and feeling pain (having an itch and being in pain), for example - never have objects at all, internal or external. Of those that do, some, such as sense perception, sometimes have external objects and sometimes do not - that is, there are tokens of these types that do and tokens that do not have objects - and others never do. But every token of every type of mental operation that has an object at all has an internal object which is an idea. Indeed, the connection between an operation and its object is essential for Locke, on both sides. No (token or particular) idea is capable of existing apart from the (particular) operation it is linked to, nor is any (token or particular) operation able to exist apart from the (particular) idea that is its object. It is natural, therefore, for Locke to think, as he does, that to be conscious of any mental operation, active or passive, is necessarily to be conscious as well of the object of that operation. Thus it is that we find him including among the objects of consciousness not only mental activities such as thinking and perceiving, and mental passions such as feeling warmth and solidity, but also ideas such as that of red, sweet, warmth, and solidity. And since we have ideas of mental operations, and these ideas, too, are things that pass in our minds, we must on occasion be conscious of the ideas of thinking and feeling, as well as of thinking and feeling themselves, and of their objects.

I noted earlier that Locke includes propositions among the objects of consciousness. If I am conscious that P then the thing that I am conscious of is the proposition that P. But does the converse hold as well? If I am conscious of the proposition that snow is white, does it follow that I am conscious that snow is white? Locke cannot have intended that. In this case too, consciousness must be restricted to things that happen or obtain in our own minds. That snow is white is not itself such a thing, but the proposition that snow is white is included in this category to the extent that I merely entertain it, or hold it in my mind without affirming it. Some propositions, however, are so included in virtue of their subject matter. These are propositions about the actions and the contents of our minds: that I am now perceiving something, for example, and also that I am now smelling, not burning tires but the smell of burning tires, that is, the characteristic idea or sensation that burning tires normally produces.

Locke often links consciousness with knowledge. If someone is conscious of x or is conscious that P, then there is something that he knows. But what? Suppose a man is conscious of hearing a sound of the kind that is normally made by a coach passing in the street. One thing he knows is that he is hearing something; not merely that the proposition that he is hearing something is before his mind but that the proposition is true. It follows that he knows what sort of thinking or perceiving is occurring in his mind: that he is hearing something as opposed to smelling or seeing it. It also follows that he knows what he is hearing, not in the sense that he knows what is actually producing the sound that he hears, but in the sense of its phenomenal content, "what it is like".

Does Locke have any basis for claiming, as he does, that our hearing man really knows what sort of thinking or perceiving he is engaged in and what its phenomenal content is: that he really is hearing as opposed to seeing or smelling something, and that it

really is the sound as it were of a passing coach he hears and not some other sound? Could not a conscious subject, a thinker or perceiver, be mistaken with respect one or both of these? Locke apparently thinks not. In his view, the mind is perfectly transparent, and all its contents - both its operations and their internal objects - are unavoidably present to it. In one passage he states, as a general principle, that no idea - and by implication nothing we are conscious of - "can be other than such as the Mind perceives it to be; [nor can it] "be other, *i.e.* different, without being perceived to be so" (E II.xxix.5: 364). In other passages he suggests how speakers learn the meanings of words for mental activities: "*What Perception is*", he writes, "every one will know better by reflecting on what he does himself, when he sees, hears, feels, *etc.* or thinks, than by any discourse of mine. Whoever reflects on what passes in his own Mind, cannot miss it: And if he does not reflect, all the Words in the World, cannot make him have any notion of it" (E II.ix.2: 143). 'Perception', as Locke is using the word here, stands for the generic action of the mind (as does 'thought' or 'thinking'); but he holds the same view with respect to specific varieties or forms of perception and thought. Of willing, for example, he says that "whosoever desires to understand what it is, will better find it by reflecting on his own mind, and observing what it does, when it *wills*, than by any variety of articulate sounds whatsoever" (E II.xxi.30: 249-50). But these are merely pronouncements on his part: he does not give or even hint at any reason for thinking that they are true. Nor does he explain just how it is, that if these things were true, they would support the doctrine that what appears to consciousness really is the way it seems to be: an instance of willing, say, as opposed to desiring something, or a case of smelling as it were roses as opposed to hearing as it were a coach passing in the street.

Notice that in two of the passages just quoted Locke speaks of reflecting on a mental operation. He often uses 'reflection' as a

technical term in the *Essay*. So used, it stands for one of the two modes of "experience", whereby we acquire all the simple ideas we eventually come to possess. The other is sensation, and Locke differentiates between the two modes by means of their different objects. By sensation we receive the ideas of "external, sensible things" via our corporeal senses; by reflection we get the ideas of "*the Operations of our own Minds* within us" (E II.i.4: 105). So reflection is directed to the same objects that consciousness is, namely mental operations; indeed, in the same passage Locke says that reflection "is the *Perception*" of these operations, which is exactly what he says of consciousness in a passage, quoted earlier, a few pages later. Are 'consciousness' and 'reflection', then, different words for the same (type of) mental action for Locke? (I mean 'reflection' in Locke's technical sense). It appears that they are, although he does not use them interchangeably or otherwise treat them as synonyms.

In any case, the verb 'reflect' seems not to be used in this technical sense in the two passages quoted a moment ago. Reflection there seems to be something one can do more or less carefully, with greater or less attention, something that is, therefore, to some degree under one's control. (Cf. Reid 2002: 58-59). Reflection in the technical sense, by contrast, is a species of "bare naked perception", something that merely happens in one's mental life. The same, evidently, is true of consciousness. "The Soul of a waking Man", Locke declares, "is never without thought, because it is the condition of being awake"; and it is "hard to conceive, that any thing should think, and not be conscious of it" (E II.i.11: 109-110). That is why it is wrong to equate either consciousness or reflection, as Locke conceives them, with introspection, with inner observation, as some commentators have proposed doing. For observation too is a mental operation that is to some degree under its subject's control. (See Mishori 2004 for further discussion of this point.)

In recent accounts, philosophers of mind often make the word 'consciousness' synonymous with, or at least define it in terms of, 'awareness', as do the authors of our most authoritative current dictionaries: the OED, the Merriam Webster's, and the American Heritage. And I think it is right to say that Locke's consciousness is or is a form of awareness, in our sense of the word. But Locke himself never says any such thing: the words 'aware' and 'awareness' never occur in the *Essay*. It is not that these words were not in use in the seventeenth century: 'aware' certainly was, according to the OED, and there is no reason why 'awareness' should not have been too, although the earliest citations are to texts from the nineteenth century.

I noted earlier that, for Locke, everything that passes in the mind is an object of consciousness. Locke often puts the point more strongly: "'tis altogether as intelligible say", he asserts in one passage, "that a body is extended without parts as that any thing *thinks without being conscious of it*" (E II.i. 19: 115); and again, later on: "consciousness is inseparable from thinking, and as it seems to me essential to it" (E II.xxvii.9: 335). He even goes so far as to declare, on one occasion, that "thinking consists in being conscious that one thinks" (E II.i.19: 115). This last statement is surely an exaggeration. Locke does not think that thinking and consciousness are the same thing, nor even that the latter is an essential ingredient or property of the former. His view seems to be, rather, that consciousness is a necessary concomitant of thought, something that is different from it but nevertheless always accompanies it.

Even this position, however, leads to a problem for Locke. If consciousness always accompanies thinking, in the broad, generic sense that Locke apparently intends in these passages, then it always accompanies every form of thinking, and of perception too, which in its broad, generic sense he equates with thinking. But consciousness itself is a form of perception and hence of thinking.

So if a man is thinking and hence is conscious of his thinking, he must also be conscious of his consciousness of his thinking, and also conscious of being so conscious, and so on. Of course, not all regresses are vicious, and the one entailed here is not sequential or causal as is, for example, the regress that Locke attributes to the libertarian, according to which no action can be free unless it is preceded and caused by a volition that is itself free: that regress, Locke clearly thinks, is vicious (E II.xxi. 25: 247). But even if the regress of consciousness is benign, the doctrine that consciousness always accompanies thinking is problematic. For Locke holds that a man is conscious of everything that passes in his mind. Hence if there are infinitely many different acts of consciousness occurring at any given time, the man must be conscious all of them; and it seems obvious, if only on empirical grounds, that no one is conscious of any such thing. I do not know if this problem can be solved, by Locke or by anyone.

Discussions of consciousness often make mention of self-consciousness. Locke does not use this term in the *Essay*: he does use it in one passage in his second Reply to Stillingfleet, but only because Stillingfleet had used it in the letter to which Locke was responding (W IV: 325). Still, it is reasonable to take Locke as holding that there is such a thing as self-consciousness. In defining 'person' he says that a person can "consider it self as it self, ... which it does only by that consciousness, which is inseparable from thinking" (E II.xxvii.9: 335). And a page later: "an intelligent Being ... is *self* to it *self* now ... by the consciousness it has of its present Thoughts and Actions" (loc cit. 10: 336). What Locke is suggesting, it seems, is that a conscious being is or can be conscious of itself. But how would Locke understand being conscious of oneself? Not, surely, in the Cartesian way that is disparaged by Hume, by being aware of a single, constant, simple substance that is distinct from all of one's perceptions. "For my part," Hume declares, "when I enter most intimately into what I

call *myself*, I always stumble on some particular perception or other" (T I.iv.6: 165). Locke wouldn't quite put it this way, but he would certainly agree with Hume's sentiment here. What Locke does say, in his chapter on "our Knowledge of Existence" in Book IV of the *Essay*, is that we perceive "*our own Existence*, and we do so "so plainly, and so certainly, that it neither needs, nor is capable of any proof" (E IV.ix.3: 618). I perceive my own existence whenever I perform, or suffer, any sort of mental action. When I feel pain, for example, I know that I do, and "if I know *I feel pain*, it is evident I have as certain a Perception of my own Existence as of the Existence of the Pain I feel:) (ibid.) And, in sum, "in every Act of Sensation, Reasoning, or Thinking, we are conscious to our selves of our own Being" (E IV.ix.3: 619). To be conscious of myself is to be conscious that I exist. Consciousness of oneself is thus for Locke a kind of by-product of one's consciousness of one's thoughts and perceptions, that one is having them, what kind of thoughts or perceptions they are, and what they are thoughts or perceptions of. There is nothing in this position either to please Descartes or to displease Hume.

II. Consciousness as an Individual Persisting Thing

The consciousness that I have been discussing so far is a property, whether of minds, men, or persons. But Locke also uses the word 'consciousness' in a quite different way, not as an abstract term with the grammar of a mass noun, but as a count noun with a plural form and taking the indefinite article and numerical modifiers. In this use, 'consciousness' stands for an individual thing which can persist and maintain its identity through a period of time. In the *Essay* this use occurs when Locke distinguishes one consciousness from another, or speaks of two or more consciousnesses, or says that the same consciousness exists at different times. It is true that the phrase 'the same consciousness' can be used to mean 'consciousness of the same quality or character' or 'consciousness

that is exactly similar', and in this use the consciousness meant could well be the property. But there are plenty of places in the *Essay* where Locke means by 'the same' "numerically one and the same" and not "qualitatively" or "specifically the same"; and in these places he is certainly speaking not of a property of a subject but of a subject in its own right.

But what kind of thing is this consciousness? It is not, Locke makes it plain, a substance; it belongs rather to the category that he calls "modes": an individual consciousness depends for its existence upon an existing substance, whether that substance is simple or compounded. Consciousnesses are, in fact, modes of persons for Locke, and persons are compounded substances, or so I believe, although Locke does not explicitly say either of these things: he was anxious to remain agnostic about the nature of persons, and took pains to avoid committing himself in any way with respect to them - as he did also with respect to the nature of human beings. Whatever the nature of persons, however, Locke certainly takes consciousnesses to be modes, just as he takes the property of consciousness to be a mode. Regarding my consciousness, Locke says - and here it is not clear whether he means the individual consciousness or the property of consciousness that belongs to me - that the "more probable Opinion is, that [it] is annexed to, and the Affection of one individual immaterial Substance" (E II.xxvii.25: 345). But he does not think it follows from this that the person I am is probably an individual immaterial substance, though he grants that one or more such substances, may be among my components.

At the beginning of Chapter xxvii of Book II of the *Essay*, Locke, having maintained that the identities of things depend upon the kinds of things they are, gives a list of several such kinds. The list is not complete, nor are the items on it sharply delineated, but still it can give us some clues. Locke distinguishes substances from modes and relations, simple substances from compounded

ones, and, among modes and relations, those "whose existence is in succession, such as are the Actions of finite Beings, v.g. *Motion* and *Thought*" from "permanent Beings" (E II.xxvii.2: 329). What Locke describes as successive beings contemporary metaphysicians would call processes or (temporally extended) events. The parts of such beings, Locke says, are "perpetually perishing" and are constantly being replaced by new parts. He thinks this entails that the processes and events as wholes are perpetually perishing and hence cannot "exist in different times, or in different places, as permanent Beings can at different times exist in distant places" (ibid.). But this is a confusion on his part. Both processes and permanent modes (such as the "powers" of substances) are persisting entities which remain the same entities through extended periods of time. The difference is that they are differently composed and persist in different ways, the one by "perduring", the other by "enduring", as some metaphysicians would say nowadays. Another way of putting the point is to say that events and processes have temporal parts, whereas the parts of the permanent beings, at least of permanent substances, have only spatial parts.

Of course some recent metaphysics have argued that, despite appearances, persisting substances really have temporal parts as well as spatial ones, and hence are themselves "four-dimensional" beings. But this "four-dimensionalist" view of persisting substances is controversial, and its proponents themselves acknowledge that the common and "intuitive" view of such substances is three-dimensionalist. Even so, some Locke scholars have argued recently that Locke too is a four-dimensionalist regarding persons and living creatures, and regarding artificial bodies as well. Christopher Conn, for example, has published an impressive book defending the attribution of a four-dimensionalist view of substances to Locke (Conn 2003). Whether or not Conn is right about Locke's view of substances, however, there is no doubt

that Locke holds a four-dimensional view of successive beings, "things whose existence is in succession" and whose parts, therefore, are temporal parts.

So consciousnesses, as Locke conceives them, are successive beings, processes or events that take time to occur. In this they are like the lives of individual plants and animals. The identity of a living creature is determined by the identity of its life, just as that of a person is determined by that of its consciousness. But there is difference between lives and consciousnesses. If a life existing at a time t_1 and a life existing at a later time t_2 are one and the same life, then that same life must also exist at every time between t_1 and t_2 : it must exist continuously through all the time from the one to the other, without breaks or gaps. But the same requirement does not hold for consciousnesses. Locke holds that it is at least possible for a consciousness that exists at t_1 and also exists at a later time t_2 , not to exist continuously at every time between t_1 and t_2 and yet be the same consciousness. Thus a single consciousness may have a "gappy" career and yet exist, and be the same consciousness, on both sides of the gap, or even of multiple gaps. Thus Locke thinks it is possible for both Socrates and the present Mayor of Quinborough, or for both someone who witnessed "the Ark and *Noah's Flood*" and himself while writing the *Essay* in the 17th century (E II. xxvii. 16: 341 and 19: 342), to "partake of the same consciousness", as he often puts it, and so be the same person in each case, even though there is a long time between the death of the one and the birth of the other, a time during which nothing is partaking of the consciousness in question, during which, indeed, that consciousness does not exist. This is possible, he says, in the sense that "there is no contradiction in it". But not only is it possible: something like this actually happens, he says, whenever an ordinary person falls asleep and sleeps for a while without dreaming. For there is no consciousness during dreamless sleep. Thus when a person regains consciousness some

morning, having slept dreamlessly during all of the last night, and his current consciousness is the same consciousness that he had yesterday, he is still the same person that he was yesterday, despite the interruption in his consciousness during the night.

So consciousnesses for Locke are successive beings, and they are something like the lives of animals and plants. Beyond this, however, Locke does not say much about them. What I propose to do in the rest of this paper is to work out a more elaborate and more detailed account of these consciousnesses than the one that Locke provides. I do not claim that in doing so I shall be reconstructing a position that Locke actually holds. The textual evidence is too scant to allow for that. My aim is rather to construct a position on Locke's behalf, an account of consciousnesses that is consistent with his *ipsissima verba* in the text of the *Essay* and is at least suggested by these *verba* from time to time, hence one that he could have held and might have been willing to embrace had it been presented to him.

My specific goal is to provide diachronic identity-conditions for consciousnesses, since it is the identity of consciousnesses that determines the identity of persons. This means not only that a person existing at a time t_1 and one existing at a different time t_2 are one and the same person if and only if the consciousness belonging to the one is one and the same consciousness with that of the other. It also means that the identity of a consciousness is logically prior, and can be determined independently, of that of a person, whereas the converse does not hold.

A consciousness is a process or temporally extended event, with temporal parts, so one way of proceeding would be to ask under what conditions a part of a consciousness which exists now, and another part of a consciousness which exists at an earlier time, are parts of one and the same consciousness. But before we can answer that question we have to settle what a (temporal) part of a consciousness is. And to do that we have to settle what the

constituents of a part of a consciousness are, for a consciousness, even at a single time, is not a simple thing: it is compounded of parts, though not of temporal parts. (A temporal part of a process or extended event is temporally simple, and so does not itself have temporal parts.)

Typically, when a subject is conscious at any given time, several different things are going on its mind - that is, several different mental actions are being done or suffered at once. Suppose, for example, that a man at some time is simultaneously seeing and smelling a rose, and also wondering how many roses there are in the garden before him and worrying that he won't have enough time to see them all. Locke would say that the man is performing several different mental operations at that time, operations that are different because they are of different "modes" or types of perceiving or thinking: seeing, smelling, wondering, and so on, and also because they have different (internal) objects: a visual idea of a color and shape, or an olfactory one of a smell, or else different propositions: that there are how many roses here or that there won't be enough time. According to Locke our man is conscious of each of these things: that he is seeing and smelling and also what each of these types of perception is; that it is the visual idea of a rose and the smell of a rose he is perceiving; that he is wondering and worrying and what each of these is; and what it is that he is wondering and worrying about, that is, what propositions these actions are directed towards. In addition, in Locke's view, our man is conscious that he exists at that time, and in that sense is conscious of himself. Let us call each of these different things, together with the man's consciousness of it, a *unit of consciousness*.

Among all the units of consciousness that are present in our man's mind at the time in question, there are several that are linked to one another internally or essentially, and others that are not. Our man is smelling a rose, so he is conscious of his smelling and

that he is smelling, as opposed to hearing, and he also is conscious of a distinctive smell and conscious that what he is smelling is a smell of that sort, and a smell of some other sort or another kind of idea, such as a sound, altogether. He also is conscious of his own being, that is, that he exists and that he is the one who is doing the smelling. All these are distinct units of consciousness, but each also is necessarily related to the others in certain ways: no one of them can exist unless each of the others exists. So in addition to the various units of consciousness that occur in or belong to a mind or person at any given time, there are integrated compounds of subsets of these, in which each member of the subset is internally or necessarily related to all of the others. Such integrated compounds I propose to call *experiences*. Locke himself does not use the term 'experience' in this way (nor does he directly attribute consciousness to any mental states or operations, as I noted earlier), though contemporary philosophers - and, I believe, ordinary speakers as well - do use the word in this manner.

The consciousness belonging to a subject at a particular time, then, is a compound consisting of all and only the conscious experiences that belong to that subject at that time. Since each of these experiences is itself a compound, the components of which are units of consciousness, the consciousness belonging to the subject at that time is a compound of units of consciousness as well, all those indeed, and only those, which belong to that subject at that time. (The point of dividing consciousnesses into conscious experiences as well as into units of consciousness will become apparent in a moment when I come to consider memory.) From now on I shall refer to such compounds as *conscious episodes* or *episodes of consciousness*. It is such episodes of consciousness that I take to be the temporal parts of whole consciousnesses, the persisting consciousnesses that take time to occur or unfold and

whose temporal boundaries coincide with those of whole persons, from their beginnings to their ends.

The important thing about episodes of a consciousness, as I conceive them, is that they are momentary beings. By 'momentary' here I do not mean 'instantaneous'. The temporal parts that combine to make up a temporally extended event or process, such as Locke conceives a normal consciousness to be, must themselves have some temporal extent. But if the extent is that of a moment it is relatively small and is not divided with respect to some measure or other, though the measure may not be explicit and may be inconstant.

How big is a temporal moment in the realm of conscious experience? I do not believe that Locke ever considers this question, at least directly. But he may have had views about it, and some of these may be implicit in his discussion of duration in Book II of the *Essay*, which Gideon Yaffe examines in the paper he is about to present at this conference.

A question might be raised about the assumption, which Locke clearly makes, that the attribution of a number of different experiences occurring at the same time to one and the same subject is unproblematic. He flatly declares, at the beginning of his chapter on identity, that the only legitimate question we can have about identity concerns diachronic identity, the identity of something through time, and not the identity of something existing at a time. "When we see any thing to be in any place in any instant of time, we are sure ... that it is that very thing, and not another, which at that same time exists in another place, how like and indistinguishable it may be in all other respects" (E II.xxvii.1: 328). The thing that Locke has in mind here is one that exists in the physical world, one that exists in a place as well as a time. How then, a Cartesian might demand, can this doctrine be extended to experiences, which consist of mental states and actions, and to the subjects thereof, since mental entities do not

exist in space? But Locke does not accept the Cartesian view of mental entities. He says quite explicitly that everything that exists or occurs, including souls or minds, does do so in some place or other (E II.xxiii.19-20: 306-7). So the same *principium individuationis* which holds for bodies and their modes holds also for minds and their modes: x and y are the same body or the same mind if and only if x and y exist in the same place at every time at which they both exist. By this principle Locke could claim - whether justifiably or not is another matter - that all and only the experiences belonging to a single individual subject - whether mind, man, or person - exist when and where that subject does. And from this he could conclude that having identified a set of experiences that share a common spatio-temporal location, these experiences all belong to the one and only subject of experience that occupies that same location at that same time.

But this is not conclusive. How does Locke know that there is only one such subject occupying any given spatio-temporal location? This follows, he might say, from his *principium individuationis* as it applies to minds. But perhaps that principle doesn't hold as neatly for minds as it seems to do for bodies. Had he tried, Locke might well have been able to devise some improbable but possible scenario, like those he invokes to test various hypotheses about the identities of men and persons, according to which two or more distinct subjects of experience, and hence two distinct consciousnesses, do occupy the same place at the same time. Indeed, as a matter not of fantasy but of fact, some such thing may actually occur in human beings who have had their brains bisected - a procedure that was of course undreamed of in Locke's day. A good account of this procedure and of the issues it raises is given in a well-known paper by Tom Nagel (Nagel 1975).

It did not, however, occur to Locke to worry about the synchronic unity either of consciousnesses or of the subjects

thereof. His concern was confined to their diachronic identity. The question he needs to answer about the diachronic identity of consciousnesses, as I propose to formulate it, is: under what conditions are an episode e1 of consciousness occurring at a time t1 and another episode e2 of a consciousness occurring at a later time t2 episodes of, that is temporal parts belonging to, one and the same consciousness c? Locke's answer to this question, as everybody knows, appeals to memory: the subjects of consciousnesses, whatever they are, have the ability to remember in the present things that happened or existed in the past.

Specifically, according to the account I am presenting, Locke's answer is this: e1 and e2 are parts of the same consciousness c if and only if some experience x1 that is included in e1 is also included, by way of memory, that is, by being remembered, as an experience x2, in e2. The experiences here, x1 and x2, are momentary beings, as are the episodes, e1 and e2, in which they are included, and, furthermore, x1 and x2 are numerically identical for Locke: they are one and the same experience, except, as he says, that its occurrence in the later episode is accompanied by the "consciousness of its having been in the mind before" (E I.ii.20: 97). This accompanying consciousness is distinct from the experience x2, in the sense that the latter could occur without it (and did so occur on its initial appearance), although this could not occur without that.

The objects of memory here are whole experiences, not merely mental operations or the ideas or propositions which are the objects such mental operations. What is remembered, in the case mentioned earlier, is not that someone smelled something on a certain occasion, or even that he smelled a smell of the kind that roses normally emit; what he remembers is smelling that smell on that occasion. Or rather, this is what must be the case must be if memory is indeed able, as Locke supposes it is, to "unite" a past consciousness with one occurring in the present.

It is thus by the power of memory, Locke maintains, that an experience belonging to a consciousness that occurs at one time is somehow included in a consciousness that occurs at a later time. By memory a present consciousness "communicates" with a previous one, as Locke says, or the latter is made accessible to the former, or the former "extends to" or "reaches" the latter.

It is, however, a strange theory of memory that makes the object of a present memory to be the same experience, numerically the same, with an experience that occurred in the past, as Locke's theory appears to do. One would have thought that individual experiences, like individual performances or occurrences of mental actions, are at least in part individuated by their times of occurrence. If *x* and *y* are experiences or actions occurring at different times then they must be distinct experiences or actions, that is, different tokens, though they may of course be tokens of the same type. But Locke's texts are explicit, both in the first and in the second editions of the *Essay*. For in the second edition, having thought through the implications of some of his original statements about memory, Locke made some corrections in or added some qualifications to these statements. But in the second edition as well as the first Locke says that in memory, what once was "an actual perception, is so in the mind, [that it is] an actual perception again". Thus memory "brings into actual view" an idea - or by extension a whole experience - that was in the mind before (E I.iv.20: 97). An object of memory, whether idea or experience, was once actually present to consciousness, and that very thing, the same individual, is now actually present to consciousness again. Its status may have shifted: it was actual, became non-actual, and then became actual again. But it is one and the same thing throughout. Unfortunately, Locke does not tell us where that object existed, or how it was able to exist, while it was in its non-actual state.

Note that the memory condition I have stated for the sameness of consciousnesses requires that *some* experience included in the earlier episode be included in the later one, not that *every* such experience be included. This feature may help to diminish the force of the threat posed to Locke's theory by the counterexample devised by Thomas Reid (Reid 2002 III.vi: 276). A boy is flogged for robbing an orchard; the boy grows up and becomes a brave officer, who captures the enemy's standard in his first battle; then in later life the brave officer is made a general. The brave officer remembers the flogging, and the general remembers the officer's taking the standard, but the general does not remember the flogging. So on Locke's view, as Reid interprets it, the general both is and is not the boy, given that identity is transitive, which nobody (including Locke) would deny.

But suppose that in the same case, the general's consciousness does include some other experience which was included in the boy's consciousness at the time of the flogging, an experience either of some contemporary event or, by short term memory, of an event recently past. Then the result claimed by Reid would not ensue. But this move defends Locke only against some versions of the brave officer story. For it is at least possible that the general should have no memories of any of the boy's experiences. What then? Reid might still claim that the general and the boy are the same person. To that claim Locke might respond that Reid is confusing the person (or persons) we call the boy and the general with the human being we use the same terms to refer to. For although Locke distinguishes persons from men, and believes it is possible for different persons to be associated with the same man in the course of the person's duration, and vice versa, he thinks "the more probable opinion", the situation that obtains as a matter of fact, is that every person is associated with one and the same man, and every man with one and the same person, as long as either exists. Locke also thinks that personal names and pronouns

are applied to both men and persons by ordinary speakers, so that we must be careful, when we apply the name 'Socrates' or the pronoun 'I' to something, to note just what we are applying it to: the man Socrates is or I am, or the person each of us also is (E II.xxvii.20: 342).

But suppose Reid could convince us that he is not so confused, that he is observing the distinction between man and person, and yet still insists that the general and the boy are indeed the same *person*, not merely the same *human being*, despite the absence of any connection by memory between them or between their respective consciousnesses. In that case I think, if we wished to save the account that I have been developing on Locke's behalf, we would have no choice but to add a clause to the identity condition I stated earlier. That condition is, that e1 and e2 are parts of the same consciousness c if and only if some experience x1 that is included in e1 is also included, by way of memory, in e2. The revised condition is, that e1 and e2 are parts of the same consciousness c if and only if *either* (1) some experience x1 that is included in e1 is also included, by way of memory, in e2, *or* (2) there is some other conscious episode e3, that exists after e1 and before e2 does, such that there is some experience x1 that is included in e1 and also, by way of memory, in e3, *and* also some experience x3 that is included in e3 and also, by way, of memory, in e2. This addition might seem to be ad hoc, but it also does the job of defending Locke's theory against what is itself a challenge based on a quite unusual scenario.

Another 18th-century objection to Locke's theory was made by Joseph Butler (Butler 1975). This is that Locke's theory of personal identity is circular, or question-begging, in that it makes personal identity depend on the identity of consciousnesses, but the identity of consciousnesses in turn depends on the identity of the person whose consciousnesses are in question. That is, there is no way to establish, or even conceive, that a consciousness c1 and a

consciousness c2 are the same consciousness c without making c1 and c2 be consciousnesses of the same person. I have no time to consider this objection in full, but my response to it is simply that I have tried to develop my account in such a way that this circularity is avoided, and I hope that I have succeeded.

I believe that the account of the identity of consciousnesses that I have presented on Locke's behalf, is coherent, and that it provides an adequate basis for his theory of personal identity. But I cannot in good conscience claim conclusively to have established that it does. There is also the historical question, whether Locke himself would or could have accepted the account, given the things he explicitly says in the *Essay*, and I cannot claim to have established that either. More work, both philosophical and historical, remains to be done. I am hoping that the members of this audience can help me to see, better than I now see, how to go about doing it. Thank you.

7 November 2004

REFERENCES

- Armstrong, David. 1997. What is Consciousness? in NBC 721-728. Originally published in *The Nature of Mind*, by David Armstrong (1981) 55-67.
- Block, Ned, Owen Flanagan, and Guven Guzeldere, ed. 1997. *The Nature of Consciousness*. Cambridge MA: MIT Press. [NC]
- Butler, Joseph. 1975. Of Personal Identity, in PI, 99-105. Originally published as an appendix to *The Analogy of Religion*, London 1736.
- Conn, Christopher H. 2003. *Locke on Essence and Identity*. Dordrecht: Kluwer.
- Dretske, Fred. 1997. Conscious Experience, in NC, 773-788. Originally published in *Mind* 102 (1993) 263-283.
- Garrett, Don. unpub. Locke on the Transfer of Consciousness. Paper read at the American Philosophical Association, Pacific Division, Albuquerque NM, April 2000.
- Grice, H.P. 1975. Personal Identity, in PI, 73-95. Originally published in *Mind*, 50 (1941).
- Guzeldere, Guven. 1997a. Introduction. The Many Faces of Consciousness: A Field Guide, in NC, 1-68. Originally published in *Journal of Consciousness* 2 (1995) 30-51, 112-143.
- . 1997b. Is Consciousness the Perception of What Passes in One's Own Mind?, in NC, 789-806. Originally published in *Conscious Experience*, ed. by T. Metzinger (1995) 335-357.
- Hume, David. 2000. *A Treatise of Human Nature*, ed. by David Fate Norton and Mary J. Norton. Oxford: Oxford University Press. Originally published London 1739-40.
- Locke, John. 1975. *An Essay concerning Human Understanding*, ed. by Peter H. Nidditch. Oxford: Clarendon Press. Originally published London 1690.
- Mishori, Daniel. 2004. Locke on the Inner Sense and Inner Observation. *Locke Studies* 4: 145-81.
- Nagel, Thomas. 1975. Brain Bisection and the Unity of Consciousness, in PI, 227-45.
- Perry, John, ed. 1975a. *Personal Identity*. Berkeley: University of California Press. [PI]
- . 1975b. The Problem of Personal Identity, in PI, 3-30.
- . 1975c. Personal Identity, Memory, and the Problem of Circularity, in PI, 135-55.
- Quinton, Anthony. 1975. The Soul, in PI, 53-72. Originally published in *Journal of Philosophy*, 59 (1962).
- Reid, Thomas. 2002. *Essays on the Intellectual Powers of Man*, ed. by Derek R. Brookes and Knud Haakonsson. Edinburgh: Edinburgh University Press. Originally published Edinburgh 1785.
- Rosenthal, David. 1997. A Theory of Consciousness, in NC, 729-754. Originally published as ZiF Technical Report (1990).
- Ryle, Gilbert. 1949. *The Concept of Mind*. London.
- Shoemaker, Sydney. 1975. Personal Identity and Memory, in PI, 119-34.
- Thiel, Udo. 1983. *Lockes Theorie der personalen Identitat*. Bonn: Bouvier Verlag.
- . 1991. Cudworth and Seventeenth-Century Theories of Consciousness, in *The Uses of Antiquity*, ed. by Stephen Gaukroger. Dordrecht: Kluwer, 79-99.
- . 2000. The Trinity and Human Personal Identity, in *English Philosophy in the Age of Locke*, ed. by M. A. Stewart. Oxford: Clarendon Press, 217-243.
- Yolton, John W. 2004. *The Two Intellectual Worlds of John Locke: Man, Person, and Spirits in the Essay*. Ithaca NY: Cornell University Press.