

*Locke on Consciousness, Personal Identity
and the Idea of Duration*

By GIDEON YAFFE
University of Southern California

It is as undeniable as it is unhelpful to say that for Locke personal identity consists in the identity of consciousness. It is undeniable because he just comes out and says as much in passages like the following: “[T]he same consciousness being preserv’d. . . the personal Identity is preserv’d.” (II.xxvii.13)¹ It is unhelpful, however, for two reasons: First, it is unclear what consciousness is; what portion of a mind’s mental activity at a time is its “consciousness”? Second, even if we knew what, of all the myriad things going on in my mind now is my “consciousness” and we knew what, of all the myriad things that went on in the mind of a child who, in 1976 was forced to wear an embarrassing sailor suit, is his “consciousness”, it would still be unclear what conditions must be satisfied for the two “consciousnesses” to be the same. The simple memory theory of personal identity—the theory according to which later and earlier person-stages are stages of the same person just in case the later can remember the experience of the earlier—tries to solve both problems at once. The theory equates consciousness with any conscious act of awareness and then insists that two acts of awareness are the same in the relevant sense if they have the same content, if they are awarenesses of the very same thing. What makes my “consciousness” and the sailor-suit-wearing boy’s the same, on this view, is that we are both aware of the same event, and in the same way.

The simple memory theory has the great virtue of intelligibility, but it has many familiar problems, and there are familiar problems with attributing it to Locke. I won’t rehearse either here. My aim, instead, is to make some progress in determining what Locke actually understands “consciousness” to be and to use what I learn to shed some light on his theory of personal identity. Consciousness is, obviously, a form of awareness; this it has in common with sensation, which is also, for Locke, a form of

awareness. And consciousness is, obviously, directed inward. As Locke puts it, “[c]onsciousness is the perception of what passes in a Man’s own mind.” (II.i.19). In this way it differs from sensory perception, which is awareness not of what goes on in the mind, but of what goes on outside it. But can more be said? Are there further, and more important, differences between sensation and consciousness? They differ in their objects, but do they differ in other important ways as well? To answer this question, I will be looking in what might seem, at first, a peculiar place; I will be looking at Locke’s discussion of the idea of duration in II.xiv. My claim is that the argument of that chapter sheds significant light on Locke’s notion of consciousness and what makes it different from sensory perception. Most of this paper will be spent drawing from Locke’s discussion of duration a difference between consciousness and sensory perception. However, at the end, I’ll return to the “same consciousness” theory of personal identity and say a word about how that theory might be understood in light of what we will have learned about the nature of consciousness, as Locke understands it.

The notion of duration that Locke has in mind when, in II.xiv, he undertakes to explain how we acquire the idea, is inherited from Descartes.² Descartes was concerned with the question of whether duration was something over and above the things that possessed duration or whether it was, instead, a quality that emerged from the existence of a thing. He sides with the latter view, making the following remark about duration:

[W]e should regard the duration of a thing simply as a mode under which we conceive the thing in so far as it continues to exist.³

So, for Descartes a thing has the quality of duration just in case it has the general property of continuous existence. We might think of the property this way: if an object’s path through time is continuous, then it has the property of duration. On the view of duration that Descartes offers it seems that particular amounts of duration are to duration as figures are to extension; both are particular ways in which something continuous (extension in time, extension in space) can be terminated. Locke endorses this conception, writing that “we call the Existence, or the Continuation of the Existence of our selves, or anything else. . .the *Duration* of our selves, or any such other thing. . .” (II.xiv.3).⁴

It is worth pausing for a moment to consider a bit further how Locke must be thinking of continuity of existence, of duration, given the way he construes a number of closely-related concepts. Consider, for instance, his infamous gloss on the “*principium individuationis*”:

[T]he *principium Individuationis*. . .is Existence it self, which determines a Being of any sort to a particular time and place incommunicable to two Beings of the same kind. . .Let us suppose an Atom. . .existing in a determined time and place:

‘tis evident, that, considered in any instant of its Existence, it is, in that instant, the same with it self. For being, at that instant, what it is, and nothing else, it is the same, and so must continue, as long as its Existence is continued: for so long it will be the same, and no other. (II.xxvii.3)

Here Locke uses the notion of continuity of existence to explain—or, arguably, to debunk—the idea of that which distinguishes each thing from others, which gives a thing its identity. At a time, a thing is identical to all and only those things of the same kind that occupy the same place as it; over time a thing is identical to all and only those things in which its existence is continued. Thus, for Locke the notion of existence is interdefinable with the notion of occupying a place. Locke takes it to be axiomatic that two things of the same kind cannot occupy the same place at the same time. But, more importantly, Locke uses that axiom to say what it is to exist at a time: it is to occupy a place that others of the same kind cannot occupy. To exist continuously, then, is to continuously occupy places that distinct others of one’s kind cannot occupy.

To understand what, exactly, this means consider Locke’s example of an atom and make two simplifying assumptions: assume that the atom occupies exactly one point at any one time, and imagine that the atom’s universe is one dimensional; it consists of nothing but a number line. Consider the set of places occupied at some time or another by the atom. Call that set P. If the atom exists continuously, P must be what I will call “densely packed”: for any two members of P—two places occupied by the object—there is a member of P that lies between them; that is, there is a member of P that is closer to each of the two members than they are to each other. Thus, for a set of places to be densely packed, there must be some measure of distance between places. However, to be densely packed is not enough for continuous existence, for the atom might leap, say, from 0 to 1 and then return to cover the points in between 0 and 1. In such a case, the atom’s occupation of place would not be continuous, but would involve a leap. To prevent this possibility we must imagine that P is also what I will call “densely ordered”: the path that the thing travels through the members of P over time must be continuous in the mathematical sense. That is, if the atom continuously exists and if at some time in its lifetime the atom occupies point 0 and at another later time occupies point 1, then it must move through all of the points between 0 and 1 and must do so continuously. If the set of places occupied by a thing is densely ordered, it is also densely packed (although, for reasons that I hope become clear, it is nonetheless helpful to have the notion of a densely packed set of places). So, we can say that a thing enjoys continuity of existence only if the set of places that it occupies during its lifetime is densely ordered. This is a necessary condition on continuity of existence; as we will see in a moment, it is not sufficient.

But first, consider the notion of “place”. Place turns out to be an extraordinarily elusive notion for Locke—more elusive than the example of an

atom captures. What kind of property a thing's place is depends on what kind of thing it is and on our purposes in designating it. The place of an atom, an extended object, is physical position or position in space. The place of a chess piece, however, is, for many purposes, determined by the square it occupies on the board; the king may be said to be occupying the same place as before even though it has been moved within the square that it occupied, and even though the board has been moved across the room (II.xiii.7–8). As Locke says,

[T]his Modification of Distance, we call *Place*, being made by Men, for their common use, that by it they might be able to design the particular Position of Things, where they had occasion for such Designation, Men consider and determine of this *Place*, by reference to those adjacent things, which best served to their present Purpose, without considering other things, which to another Purpose would better *determine the Place* of the same thing. (II.xiii.9)

The relativity of a thing's place to our purposes in identifying its place is surely related, for Locke, to the relativity of a thing's kind to our purposes in identifying the thing's kind, a well-known, and well-explored, feature of Locke's account of kind membership.⁵ Locke also suggests, however, that a thing's place need not even be a physical location, not even a relative physical location, for there are notions of distance that are non-physical and with them come corresponding notions of adjacency and of place. For instance, he offers the following example:

[I]f any one should ask, in what Place are the Verses, which report the Story of *Nisus* and *Eurialus*, 'twould be very improper to determine this Place, by saying, they were in such a part of the Earth, or in *Bodley's Library*: But the right Designation of the place, would be by the parts of *Virgil's Works*; and the proper Answer would be, That these Verses were about the middle of the Ninth Book of his *AENEIDS*; And that they have been always constantly in the same Place ever since *Virgil* was printed. (II.xiii.9)

It seems to me virtually impossible to give good sense to Locke's idea here without considering the place of the story as something distinct from position in space in the ordinary sense. It seems, rather, that with the identification of the relevant kind of thing comes a corresponding notion of distance and a corresponding notion of place which is such that two things of the same identified kind cannot at the same time occupy the same place, in that sense of place.

In short, then, for a thing to enjoy the property of duration, or continuous existence, is, at least in part, for the set of places that it occupies to be densely ordered, with the understanding that the notion of place itself, and the notion of distance between places that that concept requires, are determined by the kind to which the thing belongs. To have the idea of a thing's duration, then,

is to represent the set of places that thing occupies to be densely ordered; such a conception, then, involves, in the background, a conception of the kind to which the thing belongs, a kind that brings with it conceptions of place and of distance among places.

However, there's more to the notion of continuous existence, even, than this. To see this, it helps to consider what Locke contrasts continuous existence with, namely successive existence. He writes,

Only as to things whose Existence is in succession, such as are the Actions of finite Beings, v.g. *Motion* and *Thought*, both which consist in a continued train of Succession, concerning their Diversity there can be no question: Because each perishing the moment it begins, they cannot exist in different times, or in different places, as permanent Beings can at different times exist in distant places. (II.xxvii.2)

The idea here is something like this: when I move my finger from, say, point 0 to point 1 on a number line, the tip of my finger successively occupies each of the points on the interval; but would it be right to say that the movement of my finger—that is, the action—successively occupied those points? No, for the movement is not at point 0 and later at point 1/2 and yet later at point 1, as the tip of my finger is. The part of the movement that is going on when my finger is at point 1/2 is gone as soon as my finger is no longer at that point; it doesn't exist again later at point 3/4 and yet again at point 1; rather new movement parts exist at those points. Notice, however, that there is no meaningful sense in which the set of places occupied by the tip of my finger is distinct from the set of places occupied by the movement of my finger. If the movement takes place anywhere, it takes place on the points from 0 to 1. But if the sets of places occupied by the two things is the same, then it's not possible for one and not the other to be densely ordered; whatever features the one set has the other has as well. However, the tip of the finger exists continuously while the movement of the finger exists successively. What follows is that there must be more to continuous existence than is captured by appeal to sets of places that are densely ordered.

What more is there? The answer is contained, it seems to me, in the passage just quoted. What Locke is noticing is that time of existence is an essential property of the parts of successively existing things, and it is these parts that can be said to occupy certain locations; not so in continuously existing things. It is also parts of continuously existing things that occupy locations—it is the tip of the finger, not the whole finger, that occupies, at a certain time, point 0 on the number line. But it is not an essential property of that part that it exist at any particular time. It does not “perish the moment it begins”.

To summarize, for Locke continuous existence can be defined like so:

A thing exists continuously *if and only if* (1) The set of places it occupies is densely ordered, and (2) The parts of the thing that occupy particular places

are capable of existing at different times; they are not excluded from doing so by their essential properties.

Notice that, and this will be of some importance later, if you were to recognize that a thing that exists at a particular time is exactly the same thing as something existing at another time, that would be enough to recognize that what you are encountering is not something that “perpetually perishes”; if the same thing exists at different times, then time of existence can’t be among its essential properties.

Return, now, to Locke’s discussion of duration in II.xiv. Given the contrast between continuous and successive existence, the remark with which Locke opens that chapter is extraordinarily provocative. He writes,

THERE is another sort of Distance, or Length, the *Idea* whereof we get. . .from the fleeting and perpetually perishing parts of Succession. This we call *Duration*. (II.xiv.1)

The reason this remark is provocative is that Locke is telling us that we get the idea of a continuous, uninterrupted existence from the perception of something that does not possess that property; we get it from a succession that has “perpetually perishing parts” and therefore fails the second condition above for continuous existence. Thus, Locke is claiming that our minds pull a rabbit from a hat: it is from something that lacks duration that we derive our idea of it. How does this come to pass?

In answering this question, Locke makes two claims, one positive and one negative. The positive claim is that we derive the idea of duration from reflection on the succession of our ideas. The negative claim is that we cannot derive the idea from the sensation of motion, the sensation, that is, of the succession of places occupied by an object. So, the rabbit cannot be pulled from just any hat; not any old perceived succession is useful for acquiring the idea of duration; only a succession of one’s own ideas will do. Thus, Locke must explain how we get the idea of duration from reflection on the succession of our ideas, and the explanation must not promiscuously allow for the acquisition of the idea from the perception of other successions, such as the succession of places occupied by a thing in motion. Conversely, the explanation for why we cannot get the idea of duration from the sensation of motion must not rule out the possibility that we are able to acquire the idea from reflection on the succession of our ideas.

Let’s look first at what Locke has to say in support of the negative claim, the claim that we cannot derive the idea of duration from the perception of motion. What he has to say on this score is easily misunderstood. It is clear that he is attacking our capacity to get the idea of succession from the sensory perception of motion. Given what the idea of duration is, this is a reasonable strategy. After all, you can’t have the idea of duration without

having an idea that represents the places a thing occupies as densely ordered; but such an idea must, then, represent those places as ordered—that is, as in succession. Thus, if you can't have the idea of succession from the perception of motion, we can be certain that you can't have the idea of duration from that perception. But, still, it is far from clear why he thinks we lack the capacity to acquire the idea of succession from the perception of motion. In support of this claim, he writes,

[B]y reflecting on the appearing of various Ideas, one after another in our Understandings, we get the Notion of Succession; which if any one should think, we did rather get from our Observation of Motion by our Senses, he will, perhaps, be of my Mind, when he considers, that even Motion produces in his Mind an Idea of Succession, no otherwise than as it produces there a continued train of distinguishable Ideas. For a Man looking upon a Body really moving, perceives yet no Motion at all, unless that Motion produces a constant train of successive Ideas. v.g. a Man becalmed at Sea, out of sight of Land, in a fair Day, may look on the Sun, or Sea, or Ship, a whole hour together, and perceive no Motion at all in either; though it be certain, that two, and perhaps all of them, have moved, during that time, a great way: But as soon as he perceives either of them to have changed distance with some other Body, as soon as this Motion produces any new Idea in him, then he perceives, that there has been Motion. (II.xiv.6)

What reason does Locke give in this passage for thinking that we can't get the idea of succession from the perception of motion? There is a way of reading what he says in support of that claim that involves attributing to him a really silly argument, and much in the passage just quoted encourages the attribution. The silly argument runs like this: "Sometimes a thing is moving, but an observer can't tell. Such an observer doesn't have the idea of succession from the perception of the thing even though the thing is, in fact, occupying a succession of distinct places. Therefore, the idea of succession is not derived from the perception of motion." The problem, of course, is that if the observer can't tell that the thing is in motion, then he isn't actually sensing motion. If a hypothetical observer who doesn't sense motion still had a certain concept, then that would be evidence that that concept is not derived from the sensation of motion. But the fact that a hypothetical observer who doesn't sense motion, lacks the idea of succession leaves the sensation of motion as one of the candidate ways in which the concept is acquired; if he'd only noticed that the thing he saw was in motion, we might say, then he would have had the idea of succession. That, however, is the opposite conclusion from the one Locke intends to draw. Now, in the passage just quoted, Locke does add that someone who does sense motion has a succession of ideas. But that point doesn't help, since, for such a person, there are two things in succession and the observer is aware of both: he is aware through sensation of the successive places occupied by the object, and he is aware through reflection of the successive states of his mind. Who's to say that he doesn't

derive his idea of succession from the former of these two successions, the successive places occupied by the object?

But if Locke isn't offering this silly argument in support of the negative claim, what is he saying? To understand the point that he's after, it helps to consider a gruesome example that he offers a few paragraphs later:

Let a Cannon-Bullet pass through a Room, and in its way take with it any Limb, or fleshy Parts of a Man; 'tis as clear as any Demonstration can be, that it must strike successively the two sides of the Room: 'Tis also evident, that it must touch one part of the Flesh first, and another after; and so in Succession: And yet I believe, no Body, who ever felt the pain of such a shot, or heard the blow against the two distant Walls, could perceive any Succession, either in the pain, or sound of so swift a stroke. Such a part of Duration as this, wherein we perceive no Succession, is that which we may call an *Instant*; and is *that which takes up the time of only one Idea* in our Minds, without the Succession of another, wherein therefore we perceive no Succession at all. (II.xiv.10)

Had Locke had access to cameras, he might have made this point by analogy to them. His claim is that the senses have the equivalent of a shutter speed, and just as the film represents all that happens while the shutter is open as simultaneous, and not in succession, so the mind registers all that happens in succession within a certain interval of time as simultaneous. All that happens from the moment the cannonball enters the room to the moment it exits is represented by one static idea, and so is represented not as successive, but as taking place at once. Just as a photograph of the event would show a cannonball streak stretching from one side of the room to the other—thus representing the cannonball in many places at once—so does the idea of the observer. And, in fact—and here's where the analogy with the camera breaks down—the mind does this across the various modes of sensation, across, for instance, sight and touch. The observer in Locke's example represents as simultaneous two sounds (the ball hitting the entry and exit walls) and his feeling of pain; if his eyes were open, what he saw would, presumably, also be represented in the same single, complex, cross-modal idea or "snapshot" of the event.

Let's call the view that Locke is expounding here the "halting sensation view". Under it, the physical sense organs register information more quickly than the information can be represented in idea. Hence, the ideas that are formed as a result of impressions on the sense organs represent not what is going on right now, but rather all that has been "stored up" in the sense organs since the last idea was formed. On the halting sensation view, the sense organs are like an information hopper: they hold information until it can be transferred as a single batch into ideas. On the halting sensation view, then, sensory ideas—the ideas you are having when you are awake and your sense organs are being stimulated—are, in a possibly misleading sense,

memories: they represent events that are an instant in the past. We do not experience these ideas as memories; they are experienced as representations of what is going on right now, but, in fact, they represent a world that has just immediately passed.

The halting sensation view might be contrasted with a different position, a position Locke is rejecting, that I'll call the "smooth sensation view". Under the smooth sensation view, in sensory perception mental representations of an object's position align with changes in the object's position. Or, in other words, on the smooth sensation view, when an object changes position there is an accompanying change in the mind's representation of the object's position. The object is never represented as occupying more than one position at once. We might think of the smooth sensation view as the halting view at the limit, in which the intervals represented by ideas are infinitesimally small. On the halting sensation view, there is, necessarily, a delay between the occurrence of an event and the appearance in the mind of a representation of that event; events can't show up in the mind, as it were, until there is enough recorded in the sensory organs to transfer into idea. If your finger is at point 0 at time 0, an idea cannot represent it as being there until, say, time 3 when the idea representing all of the locations of the finger between times 0 and 3 pops into the head. This is why sensory ideas, on the halting sensation view, are necessarily something like memories. However, under the smooth sensation view, there is no conceptual obstacle to imagining that ideas represent events just as they happen. If your finger is at point 0 at time 0, there is no reason why there couldn't be an idea at time 0 that represents it as occupying the very location that it occupies at that time. Nor, however, is there any obstacle from the smooth sensation view, to imagining that sensory ideas are delayed. Perhaps you don't represent the finger's position at time 0 until time 3, even though you represent it's position at time 0 + epsilon at time 3 + epsilon.

The differences between the smooth and halting views is perhaps better illustrated in the following figure.

In explanation, imagine an object, a point, that sits at point 0 on a number line. At time 0 the point begins to move to the right at exactly 1 unit per second. And imagine that, under the halting sensation view, all that takes place within a three second period is captured in a single "snapshot". The gray points indicate the locations that the object is represented as occupying at each moment in time, under the halting sensation view. The black points indicate the location that the object is represented as occupying at each moment in time, under the smooth sensation view. Under the smooth sensation view, at any given moment the object is represented as occupying only one location, and, we'll assume, it is represented as occupying the location that it occupied three seconds previously; at time 6, for instance, it is represented as occupying point 3, which is exactly the point which it occupied at time 3. On the halting sensation view, the object is represented, for the first three seconds, as sitting at point 0, since it sat at that point for the three seconds

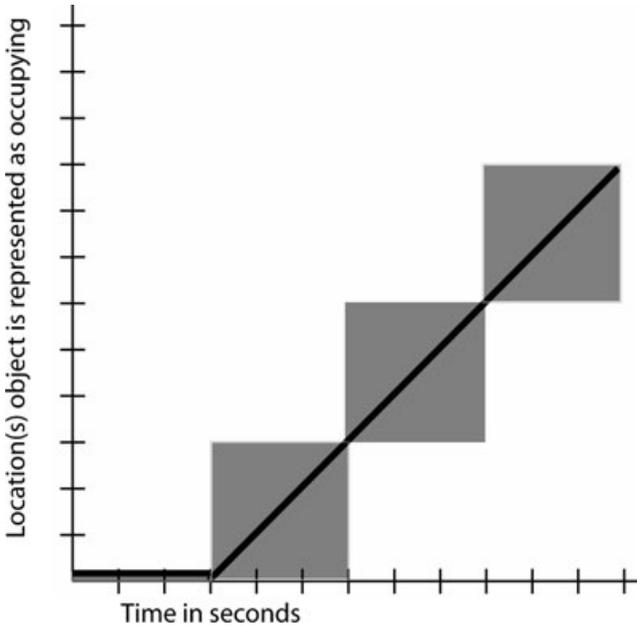


Figure 1. From time -3 to time 0 , the represented object sits at point 0 before starting to move at one unit per second from left to right on a number line. Grey points indicate the locations the object is represented as occupying by the halting idea, black by the smooth.

preceding time 0 . At time 3 , however, a new idea is produced that represents the object as occupying, at once, all the places that it occupied from time 0 to time 3 , namely, all the points between 0 and 3 ; at time 6 , a new idea appears that represents the object as occupying all of the locations it occupied between time 3 and 6 , and so on.

Both the halting sensation view and the smooth sensation view are accounts of the dynamics of sensory representation. They are accounts, that is, of the mechanism of transfer of information from our sensory organs to our ideas. On the version of the smooth sensation view represented in Figure 1, there is a delay between the moment information hits the sensory organs and the moment it is transferred to an idea—a delay of three seconds, as on the halting view—but once information begins to flow from the sensory organs to the idea, it does so smoothly, rather than in a batch, as on the halting sensation view.

The halting sensation view implies that it is actually impossible to have the idea of succession—to represent things as taking place one after another—from a very short observation period that does not span multiple ideas; the nature of the transition from sensory impression to idea excludes that

possibility in just the way that a camera that opened the shutter for 1/24 of a second at a time could not represent events succeeding each other at a shorter interval while the shutter was open as anything but simultaneous. If a photograph of a ball streaking through the air represents the ball as being in a succession of locations, it is because the person viewing the photograph knows, already, that balls don't jump through space and then back again, but, instead, travel in continuous paths. It is not the photograph itself, then, that represents the locations of the ball as successive; the photograph represents them as occupied all at the same time by the ball; if the viewer of the photograph sees a succession, it is because he supplies more than can be found in the photograph itself. Similarly, if sensation is halting, then our sensory ideas necessarily represent all that takes place within (what Locke calls) an instant as non-successive.

So, Locke is not arguing for the negative claim by arguing, simply-mindedly, that we don't have a succession of distinct ideas when we don't sense motion, and do when we do. Rather, he is pointing out that the ideas that we have immediately from sensation, whether or not we sense motion, cannot, by the very nature of the mechanism through which they are formed, represent events as successive. If the halting sensation view is right, and if (although this is a big "if") all and only the ideas that we have were those that are formed immediately in response to impressions on our sense organs, then we would have no ideas of succession. A mind so limited would be limited indeed for it would, if Locke is right, represent all the events that it encounters, in whatever order, as taking place in no order at all.

Of course, for Locke we have many ideas other than those that are formed immediately in response to impressions on the sense organs. Putting aside, for now, ideas of reflection, we also have all of the ideas that have their root in sensory experience, but which are formed through the exercise of some combination of the three important forms of mental operation: abstraction, composition and comparing. (Abstraction is the means by which we create general ideas from the ideas of particulars that experience affords. Composition is the means by which we combine ideas to create complexes that may or may not be found in experience. And comparing is the means by which we create ideas of relations from the ideas of relata.) Mightn't the idea of duration, the idea of continuous existence, be formed by taking ideas supplied by sensation, under the halting sensation view, and manipulating them? If those operations are capable of providing us with ideas of one quality formed from ideas of entirely distinct qualities, then it might be possible to use them to create an idea of succession from ideas that, necessarily, fail to represent succession.

The prospects for creating the idea of duration from the available materials using only abstraction and composition seem dim. To use an analogy, say that I have two photographs each of which represents a distinct event. I can't create for myself a collage representing the one event as preceding the other

just by taping those photos together, or by cutting parts out of them and then taping the resulting pieces together. Since the photos don't provide me with any representation of succession, I'm at a loss to create such a representation through the photographic analogues of composition and abstraction.

Comparing—the mental operator through which we form ideas of relations by setting side-by-side ideas of relata—seems more likely to provide what is needed. If I can get the idea of the relation “to the left of” by comparing the idea of my right hand and the idea of my left, why couldn't I get the ideas of “before” and “after” by comparing my idea of the docked boat with my idea of the boat at sea? Nothing that Locke says seems to rule out this possibility entirely, although it is not the direction that he goes.⁶ Instead of trying to exploit his notion of comparing to explain how we get the idea of succession, Locke instead notices that when sensing something successive we do indeed get succession, as it were, into our heads. It isn't represented there: the ideas immediately formed in response to sensory impressions do not represent it. But it is there in the ideas themselves, for they are in succession. All by itself this does not make it true that we have the idea of succession—ideas have many properties that a particular person might lack any idea of—but it suggests a tool for explaining how we get the idea of succession. Perhaps we get it not by looking out, but by looking in.

This brings us to the positive claim, the claim that we acquire the idea of duration from reflection on the succession of our own ideas. Notice, however, that there is a *prima facie* objection to the positive claim that is encountered before we even see Locke's reasons for it. If our sensory ideas are not capable of representing succession, how could our ideas of reflection make that possible? Say Locke is right that in sensation the mind is limited in the way a camera is: all that takes place within certain short periods of time is necessarily represented as non-successive. Why doesn't it follow that ideas of reflection, too, are so limited? You might think that whether you look outward, as in sensation, or inward, as in reflection, the ideas that are formed are like snapshots of that at which they are aimed, and so cannot represent succession. Or, to put the point another way, if Locke holds not just the halting sensation view, but what we might call “the halting reflection view”, then he would be committed to the claim that reflection is no more capable of providing the idea of succession than sensation is. The way we think about an object's movement as it takes place before our eyes is by having ideas of it. Similarly, the way we think about our own mind's movement as it takes place before our inner eye is by having ideas of it. If the inner eye supplies ideas of what's going on in our minds in the same halting manner as the eyes in our sockets supply ideas of what's going on outside, then there would be no more reason to think reflection capable of providing the idea of succession than sensation is. If the halting reflection view is true, then each idea of reflection would represent, say, three successive sensory ideas as simultaneous, rather than in succession. It would be not just

the sound of the cannonball hitting one wall, the pain, and the sound of it hitting the other that would be represented as simultaneous; we would also represent the ideas of those three events as simultaneous. What does this objection show? It shows that since Locke thinks that reflection is capable of providing the idea of succession, he is committed to thinking that reflection is different from sensation in this respect. Reflection, unlike sensation, is not merely halting. Perhaps reflection, by contrast, is smooth. Perhaps, that is, reflection provides us with ideas of the states of our minds that represent only one state at a time.

But how would smoothness help? If sensation were only smooth, and not halting, would we be able to get the idea of succession from the perception of motion? No. To have an idea of the successive locations occupied by an ordinary physical object, you'd need to have, at once, an idea of each of those locations. You can't have an idea of succession unless you have an idea that represents the various parts of the succession, the things that are in order. It would not be enough to represent each of the locations the object occupies; your idea, rather, has to embody a point of view from which you can see, as it were, more than one location at once. Notice that under the halting sensation view, sensory ideas do this much: the sensory idea of the cannonball represents it as occupying both the location on one side of the room and the location on the other. But representing a single thing as occupying distinct locations is not enough to represent those locations as in succession: the idea must also represent those locations as ordered, as following one another in a sequence. Sensory ideas, on the halting sensation view, fail to represent succession because they fail to have this feature. But notice that if sensation were smooth our sensory ideas would fail to represent succession for the opposite reason: they would fail to do so because they would never represent an object as occupying distinct places; they would never provide us with more than one part of the succession at a time, and so would make it impossible to represent succession at all. Simplifying, we might say that under the halting sensation view, each of our ideas presents a segment of a time line; the trouble is that our ideas don't tell us how the points of the segment are ordered, or that they are ordered at all. On the smooth sensation view, at any given instant a sensory idea presents only a point on a time line; the trouble is that it never provides us with the multiple points needed to make a succession.

What if sensation were both halting and smooth? What if, that is, impressions on your sense organs gave you two ideas: one, the halting idea, that represents all that happened in the last instant, although with no ordering within that instant, and another, the smooth idea, that distinctly represents the location of the object in the previous instant at each of the times during that instant? If this were the case, it seems we could have the idea of succession from the perception of motion. Consider, again, Figure 1. From 3 to 6 seconds, you experience, in succession, the object to be occupying each

of the locations from 0 to 3; this is your smoothly representing idea. In that same time interval, from 3 to 6 seconds, you also have an idea of the “streak” of the object: an idea that represents the object as occupying, at once, all of the locations from 0 to 3. If you could put these two ideas together, you’d have the idea of succession: the second idea tells you what the things are that are to be in succession; it provides you with all of the locations of the object. The first idea, the smooth idea, tells you how these things are to be ordered: they are to be ordered as they are ordered in the smooth experience of them. To use an analogy: imagine that you look at a photograph of the streak of the cannonball, and while you are looking at it a cannonball-shaped spotlight moves from left to right across the streak. The photograph (the halting idea) shows you all the positions the cannonball occupied; the spotlight (the smooth idea) shows you the order in which it occupied them. If reflection is both halting and smooth, then consciousness gives you a dynamic picture of the state of your mind during the previous instant, and, at once, a fixed picture of the state it was in during that instant. By putting the two together you are able to get the idea of the succession of your mind’s states. In your halting idea, reflection represents the bundle of thoughts you had an instant ago; in your smooth idea, reflection represents the order in which you had each of the thoughts in the bundle. This is true, that is, if reflection is both halting and smooth.

It’s clear that Locke takes sensation to be halting, and not smooth. After all, he clearly thinks sensation is halting and also thinks that it fails to give us the idea of succession. Since if it were also smooth it would give us that idea, he must think it is not smooth. But what about reflection? We know that he thinks that reflection is capable of giving us the idea of succession. So he must think that either it is so capable in virtue of the fact that it is both halting and smooth, or else for some other reason. At this point, it would be helpful to have some textual evidence on the point. As far as I know, however, there is none. What we have, instead, is only the following: Locke explicitly takes sensation to be halting and explicitly takes the fact that sensation is halting as part of the explanation for our failure to derive the idea of succession from the sensory perception of motion. While this suggests that he would solve the problem by appeal to a difference in this respect between reflection and sensation—he would explain reflection’s ability to provide us with the idea of succession through appeal to its being both halting and smooth—it does not entail it. It remains possible, that is, that he takes reflection to differ from sensation in some other way that allows it to provide us with the idea of succession. But what way could that be? Notice that Locke cannot simply point to the fact that the object of reflection—namely the succession of our ideas—is successive to explain reflection’s ability to provide us with the idea of succession. After all, motions, too, are successive, Locke thinks, but sensation is not, on those grounds, able to provide us with the idea of succession. It seems more likely that he identifies sensation as halting

precisely because he takes himself to be thereby contrasting sensation with reflection and thereby identifying the features of reflection in virtue of which it provides us with the idea of succession. That is, given that (a) he makes the distinction between halting and smooth awareness, and (b) that distinction can help us to distinguish between awareness that can provide us with the idea of succession and awareness that cannot—an awareness that is both halting and smooth can give us the idea, but not awareness that is just one or the other—and (c) he takes reflection to be capable, and sensation to be incapable, of giving us the idea, it seems very likely that he takes reflection to be both halting and smooth. There is powerful, albeit indirect evidence, then, to suggest that Locke takes reflection to be both halting and smooth and takes that fact about it to explain how we are able to acquire the idea of succession from reflection.

Still, the position raises many questions. Not least among them is this: where it seems that empirical evidence could be mustered to show that sensation is halting—in fact, Locke’s cannonball example is a piece of empirical evidence supportive of the claim—it is hard to see how comparable empirical evidence could be provided in support of the contention that reflection is halting, especially if it is also smooth. One might wonder what grounds there could be for such a claim, given that it is not a priori. Evidence could, in theory, be provided by introspection. However, untrained introspection does not seem to provide it, and training oneself to introspect on the representational power of introspection itself is at least very difficult and perhaps impossible. Further, the mere fact that reflection’s being both halting and smooth would help to explain how it is that we have the idea of succession does not provide support for the claim. After all, the idea of succession might, instead, be innate. Relatedly, where we could measure the length of time between events represented as simultaneous by sensation it is hard to imagine how we could make comparable measurements of the “period” of halting reflection. When an empiricist requires unsupported, and even possibly unsupportable, empirical claims in order to support his contention that a particular idea has its source in experience, there seems to be a problem. Still, defensible or not, if Locke accepts that reflection is both halting and smooth, that can tell us something important about his conception of consciousness.

More, however, can be gleaned, still, from Locke’s discussion of duration. Locke doesn’t think that it follows immediately, and without further explanation, from the fact that reflection provides us with the idea of succession that it furnishes us with the idea of duration, or continuous existence. We need to explain how it is that once we have acquired the idea of the succession of our ideas, we are able to acquire the idea of continuous existence from there. This is the positive claim of the chapter on duration and we can learn more about what Locke takes consciousness to be by considering what he says in defense of it.

Locke states the positive claim as follows:

‘Tis evident to any one who will but observe what Passes in his own Mind, that there is a train of *Ideas*, which constantly succeed one another in his Understanding, as long as he is awake. *Reflection* on these appearances of several *Ideas* one after another in our Minds, is that which furnishes us with the *Idea of Succession*: And the distance between any parts of that Succession, or between the appearance of any two *Ideas* in our Minds, is that we call *Duration*. (II.xiv.3)

The claim is that by reflection on the succession of ideas, we come to have an idea of something existing continuously over the course of the succession of ideas. The statement of this view is immediately followed by a statement of a reason favoring it:

For whilst we are thinking, or whilst we receive successively several *Ideas* in our Minds, we know that we do exist; and so we call the Existence, or the Continuation of the Existence of our selves, or any thing else, Commensurate to the succession of any *Ideas* in our Minds, the *Duration* of our selves, or any such other thing co-existing with our Thinking. (II.xiv.3)

Part of the point here seems to be deeply Cartesian: while you are thinking you know that you exist. This point is supposed to help explain why a succession of ideas gives rise to the idea of your own duration, the idea of your own continuous existence. But how does the Cartesian point help? What I suggest, and I’ll elaborate on the point, is that the Cartesian point helps us to see encounters with ideas—which are “perpetually perishing”, they are things the times of existence of which are essential to them—as encounters, also, with things capable of existing at different times. Given that this is part of what continuous existence is—it is the densely ordered successive existence of things capable of existing at multiple times—the Cartesian point is necessary for making the leap from the idea of succession to that of duration.

To see what I mean, return to Figure 1 above and imagine that the object that is “moving” is not a regular physical object, but your mind; and imagine that the “locations” of the mind, its places, are its ideas. In short, imagine that the y-axis represents distinct ideas, one for each point on the y-axis, and the x-axis, still, represents time. If reflection is both halting and smooth, at any given time, you have two ideas of reflection: one, the smooth idea, that represents your mind as having the idea that it had exactly three seconds ago, and, also, the halting idea, that represents your mind as having had all of the various ideas that it had in the previous instant. But, given the Cartesian point, both of these two ideas also represent you, a thing over and above the ideas that they represent you as having. Interestingly, however, the smooth idea represents you at only one time, while the halting idea

represents you at all of the times in the previous instant. At a typical time—at, say, time $3\frac{1}{2}$ —you have, then, a representation of yourself at time $1\frac{1}{2}$, supplied by the smooth idea, and a representation of yourself at that time, and at times both before and after that time, supplied by the halting idea. But these ideas both represent you in a way which cannot but lead you to identify the mind you are with the one you were and will be; each embodies a *cogito* thought. It is as though you thought, at once, “I see, therefore I am” and “I hear, therefore I am”; there’s no room to wonder if the one who sees and the one who hears are distinct from one another; it’s simply impossible not to identify them. Similarly, to have in your mind at once, in the instant after the cannonball leaves the room, both an idea of the set of ideas that you had during the previous instant and a continuously changing idea of the ideas that went through your mind during that instant, is to have two thoughts and two corresponding *cogito* thoughts; you think of all of what you just thought and of just what you were thinking an instant ago and you identify the two thinkers. But what this means, then, is that reflection has given you the idea of yourself as continuously existing. Your smooth idea is monitoring a mind; and your halting idea tells you it is the same mind that it just perceived and is about to perceive, despite changes in its states. If reflection is both smooth and halting then, given the Cartesian point, it seems possible to acquire the idea of your own continuous existence, your own duration, from reflection on the succession of your ideas. To put the point another way: The halting idea supplies you with an idea of a densely packed set of your mind’s places; your smooth idea supplies you with enough to recognize those places as, in addition, densely ordered; the fact that each is a *cogito* thought, and thus involves a thought not about an idea, but about a mind, and only one mind at that, supplies you with enough to realize that what you are encountering is not a succession of things that perish as soon as they are born, but, instead, a succession of things, of minds, that are capable of existence at more than one time. In short, the halting idea and the smooth idea, given the Cartesian point, supply what is needed to have an idea of your own continuous existence.

I started this paper by saying that I would draw from Locke’s discussion of duration a difference between consciousness, on the one hand, and sensory perception, on the other. What we’ve learned is that where Locke takes sensation to be halting, he takes reflection to be both smooth and halting. We acquire two ideas through reflection and they appear through different routes: one through a halting mechanism and another through a smooth one. What does this result tell us about Locke’s “same consciousness” theory of personal identity?

To answer that question, start with an answer to a different question: Can the explanation just given for how it is that, through reflection, we acquire our idea of duration allow for the acquisition of the idea from a succession of ideas in which there are temporal gaps between steps in the succession?

Imagine someone, for instance, who has a horrible case of narcolepsy: he is never awake for more than an instant, in Locke's technical sense of the term "instant", and while he is asleep he does not dream at all. Such a person has a single idea; then he has no ideas; then he awakes and has another single idea before falling unconscious again; and so on. In between any two ideas in succession there is a time when he is not conscious. Can such a person acquire the idea of duration from reflection on the succession of his ideas? Locke's answer, I believe, is yes. The reason is that from the first personal point of view, there is no gap in the succession. Locke makes this point in the following remark:

When that succession of *Ideas* ceases, our perception of Duration ceases with it; which every one clearly experiments in himself, whilst he sleeps soundly, whether an hour, or a day; a month, or a year; of which Duration of things, whilst he sleeps, or thinks not, he has no perception at all, but it is quite lost to him; and the moment wherein he leaves off to think, till the moment he begins to think again, seems to him to have no distance. . . . But if sleep commonly unites the distant parts of Duration, it is, because during that time we have no Succession of *Ideas* in our Minds. For if a Man, during his Sleep, dreams, and variety of *Ideas* make themselves perceptible in his Mind one after another, he hath then, during such a dreaming, a Sense of *Duration*, and of the length of it. By which it is to me very clear, that Men derive their *Ideas* of Duration, from their *Reflection on the train of the Ideas*, they observe to succeed one another in their own Understandings, without which Observation they can have no Notion of *Duration*, whatever may happen in the World. (II.xiv.4)

Applying Locke's point to our narcoleptic, and referring to the Figure 2, we might describe the situation like this: Imagine that the narcoleptic awakes for three second intervals and goes unconscious for one second intervals. Imagine that over the course of the interval from time 0 to time 3, he passes through all of the ideas, represented on the y-axis, from 0 to 3; at time 3 he falls unconscious for one second. Now consider his situation when he awakes at time 4. He has a halting idea of the ideas that he had over the course of the interval from 0 to 3, and a smooth idea of the state of his mind at time 0. Thus, he is in precisely the same mental condition that he would have been at time 3, had he not gone unconscious at all; he does not represent the feature of his mind that distinguishes it from a normal person's, namely that he was unconscious from time 3 to time 4. If reflection on a succession of ideas, because it is both smooth and halting, can give us the idea of duration, it is because it creates the appearance of an encounter with a continuously existing thing, oneself. So gaps such as those encountered by our narcoleptic can confound the process of acquiring the idea of duration only if they have an impact on the appearance. They don't, claims Locke, and so they can't.⁷

How does the narcoleptic in this example differ from the rest of us? It seems that he differs only in the length of time he is awake and the length

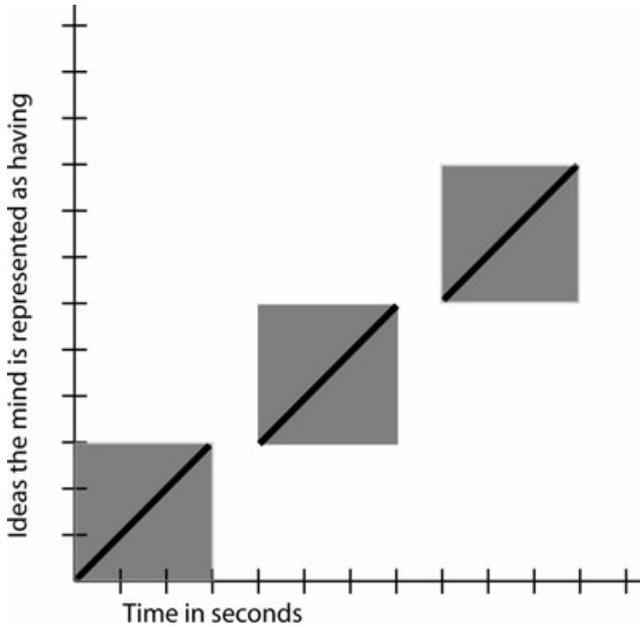


Figure 2. The represented mind has all of the ideas in the interval from 0 to 3 on the y-axis in the interval from time -4 to time -1 ; it has the ideas from 3 to 6 in the interval from time 0 to time 3. From times 3 to 4 and 7 to 8 it is completely unconscious.

of time he is asleep. Locke is insistent that “Men think not always” (cf. II.i.10) and uses the example of dreamless sleep to support the claim. If he is to be taken at his word, then when waking from truly dreamless sleep you should have no sense of time having passed from the moment you slipped into unconsciousness until the moment you awoke. And similarly, you should have the idea of yourself as having existed continuously from the moment of sleep until the moment of waking (although you take that amount of time to be very short). In short, nothing about the fact of dreamless sleep, which Locke takes to be a part of all of our lives, bars the possibility of acquiring the idea of one’s own duration, one’s continuous existence. Now it is, of course, the fact of gaps in conscious experience that make the problem of personal identity interesting: how could our conscious psychological states constitute identity through time when there are times when they simply aren’t present? But if those gaps pose no problem for acquiring the idea of oneself as continuously, gaplessly, existing because they can be bridged by consciousness, then perhaps consciousness can provide the same bridge in constituting identity through time.

Now, at the outset, I noted two challenges for interpreting Locke’s “same consciousness” theory of personal identity. The first problem is that it isn’t

clear what consciousness is: which mental activities are acts of consciousness? The second problem is that it isn't clear what the identity conditions are for acts of consciousness: what makes a later act the same as an earlier in the sense relevant to personal identity? The differences between sensation and consciousness that emerge from consideration of Locke's account of the acquisition of the idea of duration help us to meet both challenges. How it does so with respect to the first is probably clear: Consciousness is both a smooth and halting awareness of one's mental states. This distinguishes consciousness from other forms of awareness, and, in particular, from sensation. Further, since consciousness always involves two ideas, a halting idea of the mental states of the previous instant, and a smooth idea of a particular mental state enjoyed during that instant, it provides one with a conception of oneself as continuously existing, as having duration. This is a distinctive feature of consciousness, which distinguishes it from other mental activities.

To see how what has been said here helps to meet the second challenge, the need to supply identity conditions for acts of consciousness, consider the following famous passage:

[A]s far as...consciousness can be extended backwards to any past Action or Thought, so far reaches the Identity of that *Person*; it is the same *self* now it was then; and 'tis by the same self with this present one that now reflects on it, that that Action was done. (II.xxvii.9)

The notion of consciousness "extending backwards" has been interpreted as a reference to memory. But given what consciousness is, for Locke, it is better to understand that notion like so: First, consider the basic case. What does it mean for my consciousness to extend back to the actions I performed seconds ago, assuming that I did not go unconscious in the interval? For consciousness to extend back to that past action, there must be a succession of ideas, beginning with ideas had at the time of that past action and ending with my current ideas, from which the idea of duration, of my continuous existence, can be derived. We've seen what's required for that to be possible: I must be able to reflect on the succession of ideas and have both a halting and a smooth idea of them, each of which involves a *cogito* thought. When there is a succession of ideas from the past person-stage to the present from which the present can derive the idea of his continuous existence, there is identity. As we've seen, this can be possible even across gaps in conscious experience, as in dreamless sleep. We can also extend the basic case: perhaps person-stages A and C can be thought of as stages of the same person if there's some other person-stage B such that the succession of ideas from A to B allows B to derive the idea of continuous existence while the succession from B to C allows C to do so. And, of course, we can imagine allowing identity whenever such a sequence of person-stages can be constructed, even if such a construction requires many more person-stages than three. Although this

extension of the basic case will provide the tools with which to defend Locke against putative counterexamples, such as that of Reid's brave officer, it is not an *ad hoc* addition to the theory designed merely for this purpose. On the contrary: since at bottom the possibility of deriving the idea of duration from a succession of ideas rests on the possibility of identifying the two thinkers encountered in two *cogito* thoughts, one should expect inheritance of these capacities: if A is in a position to identify himself in this special way with B and B with C then A must be the same person as C, even if he cannot identify himself with C in the same way. In any event, this, I suggest, is what there is to Locke's theory of personal identity.

Let me end with a point that deserves more elaboration than I will provide here. Descartes thought that in the *cogito*, in thinking and thereby perceiving oneself as something over and above one's thoughts, we each found ourselves perceiving a substance, a thing that depended for its existence on nothing other than God. For Locke, however, what matters for identity is not that you are a substantial thing that exists continuously through time, but, instead, that you are in position to have the experience of yourself as such a thing. As it turns out, the experience of one's own substantiality can be present even in changes of substance and so the self that exists continuously cannot be a substance in the Cartesian sense. But, still, Locke's theory of personal identity owes a deep debt to Descartes, for Locke tries to exploit the psychological features of *cogito* thoughts, features to which Descartes himself was very sensitive, without insisting on the metaphysics that Descartes took those psychological features to support. To be a single thing through time, for both Descartes and Locke, is to be encountered in thought as continuously existing; but for Locke and not for Descartes, the encounter is all there is to the thing.⁸

Notes

¹ All references of this form are by book, chapter and section to John Locke, An Essay Concerning Human Understanding, Peter Nidditch, ed., Oxford: Clarendon Press, 1975.

² The same point is made in Michael Ayers, Locke: Epistemology and Ontology, London: Routledge, 1991, v.1, p. 225f.

³ Rene Descartes, Principles of Philosophy, part I, section 55 in CSM: The Philosophical Writings of Descartes, John Cottingham, Robert Stoothoff and Dugald Murdoch, eds., Cambridge: Cambridge University Press, 1985, p. 211.

⁴ Also, Descartes distinguishes duration from time, saying that time is measured duration, or continuous existence divided, in thought, although not in reality, into particular quantities (Principles I.57, CSM I 212). We might then measure a thing's continuous existence, or its manner of possessing duration, by imposing on it divisions; until such divisions are imposed, the object is not rightly said to possess the properties associated with various measures of time, such as the property of being in a certain place for an hour. Locke, too, distinguishes duration from time and thinks of time as measured duration. For instance, the title of II.xiv.17 reads "Time is Duration set out by Measures.", and a few paragraphs later Locke writes,

We must therefore carefully distinguish betwixt Duration it self, and the measures we make use of to judge of its length. Duration in it self is to be considered, as going on in

one constant, equal, uniform Course: but none of the measures of it, which we make use of, can be known to do so. (II.xiv.21)

The proposition that time and duration are distinct, and that duration is prior to any particular measure of it, implies that a thing's duration is not, in itself, divided. A thing has duration only if it exists at all of the points in time in between two points at which it exists.

⁵ For some discussion see, as a start, Peter Geach, "Identity" in *The Review of Metaphysics*, 21: 3–12, 1967; Bruce Langtry, "Locke and the Relativisation of Identity" in *Philosophical Studies*, 27: 401–9, 1975; Vere Chappell, "Locke and Relative Identity" in *History of Philosophy Quarterly*, 6 (1): 69–83, 1989.

⁶ Although it is nothing more than speculation, it is possible that he does not go this way because, although the idea of succession is an idea of relation, the idea of duration is not; perhaps he holds that if our idea of succession was merely an idea of relation, and not an idea of the successive states of a particular thing that it could not be used to form the idea of duration.

⁷ Notice that we might doubt that the narcoleptic's condition is accurately represented in Figure 2. After all, the figure shows the narcoleptic at time 4 having a halting idea that fails to represent the states of his mind over the previous three seconds. Rather, it represents the states of his mind over the first three of the previous four seconds. Why does his halting idea fail to represent his bout of unconsciousness? It is almost as though Locke is imagining that the information sits in the organs of inner sense awaiting a receptive mind; when the narcoleptic wakes the information that was sitting there suddenly has somewhere to go. By contrast, when the narcoleptic is unconscious, the organs of inner sense are picking up nothing, for there is nothing to pick up.

⁸ Versions of this paper were presented at New York University and at the University of California at Irvine. I'm grateful to audiences on both occasions. The paper was especially improved thanks to comments and suggestions made by Vere Chappell, Don Garrett, Mark Johnston, Janet Levin, Ed McCann, Alan Nelson, Lex Newman, John Richardson, Jim Van Cleve, and an anonymous referee for this journal.