**REVIEW ARTICLE**

*Locke on Personal Identity: Consciousness and Concernment.* By Galen Strawson. Princeton UP, 2011.

*The Early Modern Subject.* By Udo Thiel. Oxford UP, 2011.[[1]](#footnote-1)

Antonia LoLordo ([lolordo@virginia.edu](mailto:lolordo@virginia.edu))

September 9, 2015

**I**

In his 1769 *Defence of Mr. Locke’s Opinion Concerning Personal Identity*, Edmund Law explained that

[T]he word person, as is well observed by Mr. Locke . . . is properly a forensic term . . . denoting some such quality or modification in man as denominates him a moral agent, or an accountable creature; renders him the proper subject of laws, and a true object of rewards or punishments.[[2]](#footnote-2)

Law’s reading has received a great deal of attention recently. In his *Locke on Personal Identity*, Galen Strawson says, “Law provided an essentially correct account of Locke’s position” (2). In *The Early Modern Subject*,Udo Thiel says, “Law’s account of Locke’s theory remains one of the best to this day” (205).[[3]](#footnote-3) Indeed, Law’s two related claims – that the notion of a person is fundamentally a moral notion and that persons are modes – are central to both their interpretations. They are also central to the interpretation I put forth in *Locke’s Moral Man* (Oxford, 2012).But Law’s various admirers have taken his two main claims in very different directions.

Thiel notes – as he did in his influential 1998 “Personal Identity” – that Locke is not alone in treating persons as modes. The term ‘person’ was commonly used to refer to “a particular aspect, quality, or function of the individual human being” (26): consider Cicero’s *persona*, a role or character; Pufendorf’s notion of a person as whatever bears rights and duties; and Hobbes’ natural and artificial persons. Cicero, Pufendorf, and Hobbes all also emphasize the moral dimension of persons; the mode claim and the moral claim fit well together, though neither requires the other.

For Thiel, the Lockean person is “the human subject considered with regard to all those thoughts and actions of which it is or can become conscious” (121).[[4]](#footnote-4) This includes both a cognitive dimension and an ethical one. The cognitive dimension involves consciousness, which Thiel describes mainly in negative terms. It is not reflection (114), higher-order perception (116), intuition (118-119), or memory (118-119). Rather, the “relation of consciousness to thought … is immediate in the sense that consciousness belongs to thought itself” (116).

As Thiel notes, Locke does not say all that much about consciousness (110). He says particularly little about what it is for a consciousness to persist through time, although this is obviously crucial for his account of personal identity. Thiel explains that

[W]hen [Locke] speaks of ‘the same continued consciousness’ … this expression should be read not as referring to some unchanging property of the self, nor as suggesting that (acts of) consciousness at different points in time can be numerically identical, but rather in terms of the idea of a connectedness of distinct (acts of) consciousness (123).

This is correct, if a bit vague. Thiel adds that continued consciousness in this sense is not enough for moral responsibility:

[T]o establish a guilty mind it would not be sufficient to point to an underlying chain of direct, overlapping consciousness-links. I must also be able to bring the past misdeed to mind … It is … required that there be an ability to recall the past action and thus to attribute it to one’s own self, so that one can be held responsible for it (124).

Thus the ethical dimension requires memory as well as continued consciousness. It also requires what Thiel calls self-concern, whose function is

[T]o relate the present human subject, as person, to its own future; for the consciousness of thoughts and sensations includes the consciousness of pleasure and pain; and ‘that which is conscious of Pleasure and Pain’ desires ‘that that *self*, that is conscious, should be happy’ … it is through its link with self-concern that consciousness relates to the future person: we are now motivated to act in such a way as to avoid future pain and to attain happiness. This includes the desire that I will not have to ascribe actions to myself which will result in punishment – that is, misery (127).

Although both dimensions are important, Thiel argues that ultimately the cognitive dimension has priority. One reason is that Locke tells us “it is impossible to make personal Identity to consist in any thing but consciousness” (2.27.21). The second is that Locke clearly conceives of selves as including morally neutral experiences like seeing the Thames overflow (128).

The notions of consciousness and concern are crucial to Strawson’s interpretation as well, and his development of them is one of the book’s strengths. On Strawson’s reading, consciousness has both an occurent and a dispositional sense. To be conscious of something in the occurent sense is to experience it immediately, from the inside, in the way in which we experience our own feelings but nobody else’s. To be conscious of something in the dispositional sense is to be *able* to experience it in this way (33). The dispositional sense is the one that’s usually relevant in discussions of personal identity. For Strawson, as for Thiel, consciousness does not require higher-order thought or an explicit representation of the self or its acts (42). Rather, talk of consciousness is simply talk of “the experiential qualitative character that experience has for those who have it as they have it” (45). In this, Strawson points out, Locke is following Aristotle and Descartes, among others.

The objects of consciousness are varied: we are conscious of our bodies, our immaterial souls (if we have them), and our actions and experiences, both past and present (30). Consciousness thus includes autobiographical memory, although it goes far beyond it. To be conscious of something, Strawson argues, is to be (or to be able to be) aware of it, but also to be concerned in it – to have an interest in it. So any conscious being who can feel pleasure or pain thereby has concern for its body, its immaterial soul (if it has one), and its past, present, and future experiences. Concernment is what links consciousness with responsibility and hence what makes it relevant to personal identity. We are creatures who are “capable of a Law” (2.27.26): our actions can be assessed in moral terms, and we will be rewarded or punished for the past actions we are conscious of on the Day of Judgment. Because we are concerned for our future selves, this motivates action.

Thiel thinks of concern as fairly narrow – it’s limited to happiness and misery – and as distinct from consciousness. Strawson understands concern more broadly and as intertwined with consciousness. On his view, you are concerned for whatever you are conscious of: a body, an immaterial soul (if you have one), and various past and present actions and experiences. Thus neither of Thiel’s two reasons for giving consciousness priority to concern applies in Strawson’s case. This is a nice result.

**II**

The first question for any mode interpretation of Locke on personal identity is this: what are persons modes *of*? Locke disclaims the traditional Cartesian and late-scholastic senses of the term ‘mode’, apologizing for using it “in a somewhat different sense from its ordinary signification” (2.12.4). Still, his modes do not subsist by themselves but are “Dependences on, or Affections of, Substances” (2.12.4).[[5]](#footnote-5) So what substance do persons depend on?

Law’s answer is clear. Persons are modes of ‘men’, that is, human beings – thinking substances, together with their living animal bodies, if the substance that thinks in us *isn’t* the living animal body. After reading Thiel, I became convinced that Law’s answer is correct, although in earlier work I tended to think of persons simply as modes of thinking substances.

However, Thiel’s considered view is *not* that persons are modes of human beings:

[A]lthough [the] concepts [of person, spiritual substance, and man] are closely related and may be applied to the same individual being, they denote different aspects, respectively, under which we may consider the human subject; and they need to be distinguished from one another for an account of identity, for each of these concepts carries with it different identity criteria (107).

Again, “Locke regards soul or spirit, man, and person(ality) as different abstract ideas under which we may consider the human subject” (109). Thus both persons and human beings are modes of individual substances. In other words, we can consider an individual substance as either a person or a human being. But apart from the way in which we conceive it, this individual substance belongs to no kind and has no essential properties or persistence conditions. Discussing “the question … of how much a thing may change without losing its identity – the question about its essential constituents”, Thiel says that “[s]ince Locke rejects substantial forms and denies that we have knowledge of the real essences of substances, he argues that the answer to this question must be determined by ‘nominal essences’ or abstract ideas” (104). Thus Locke’s account of identity in general is subjectivist (102).

This part of Thiel’s argument is very quick. I believe his view is that the joint between material and immaterial substances is carved by nature, and all other joints are up to us. But he also endorses Kaufman 2008, saying, “Locke’s thinking on the ontological status of organisms is rather muddled” (138 n57). I found Thiel’s brief treatment of identity in general unsatisfying. On most issues he considers all the relevant secondary literature. Here he doesn’t consider alternate interpretations at all. And he comes dangerously close to saying that persons are modes *because* their persistence conditions derive from their nominal essences – an argument that Locke never gives and that is incompatible with his use of organisms as the paradigm examples of substances.[[6]](#footnote-6) However, little in Thiel’s account of personal identity depends on whether he is right about Locke’s treatment of identity in general.

What about Strawson? Understanding what he thinks persons are modes of requires understanding what is both the most novel and the most troubling feature of his interpretation: the distinction between persons and subjects of experience. And to understand that distinction and why Strawson draws it, we’ll have to look at all the main features of his interpretation.

**III**

For Strawson, subjects of experience are things that possess the capacities set out in 2.27.9 and 2.27.26: they are intelligent agents with reason and reflection who are capable of a law and can consider themselves as themselves, the same thinking things in different times and places (131). Subjects of experience are “at any given time concretely identical with, i.e. wholly substantially realized by”, their components – a human material body and, if we have immaterial souls, an immaterial soul – but which body and which soul are involved can change (14). Persons are moral entities, units of accountability (22). A person consists of a subject of experience and a set of actions and experiences (14). More precisely, a person consists of a subject of experience, as it exists at a moment, together with some of its history.

Strawson’s Locke takes for granted that subjects of experience persist through time. His project is to explain why subjects are responsible for certain past actions and which of those actions they are responsible for:

[T]he question of what … a subject of experience’s Personhood or Personal identity consists in, at that time … is simply a question about what that subject of experience is (morally and legally) responsible for, at that time. In this respect, Locke’s account of Personal identity begins where neo-Lockean accounts of personal identity end (23).[[7]](#footnote-7)

The explanation is simple: subjects of experience are responsible for whatever past actions they are now conscious of and hence concerned in. On Strawson’s reading, talking about personal identity is just a way of talking about responsibility. So only the past actions a subject is now conscious of can be attributed to her as a person.

Strawson thus in effect combines two lines of interpretation that have been popular recently but that I had hitherto thought of as incompatible.[[8]](#footnote-8) One line of interpretation – the one I favor – reads Locke as investigating the nature and persistence conditions of persons, and concluding that persons are ways of conceiving thinking substances (together with their bodies, if the body is not the substance that thinks). The other line of interpretation – put forth by Atherton 1983 and, perhaps, Weinberg 2012 – denies that Locke is investigating the nature or persistence conditions of persons at all. Instead, it reads Locke as investigating the idea of identity, in the course of which investigation he simply assumes that persons (or consciousnesses, on Weinberg’s version) persist through time. Strawson’s distinction between subjects of experience and persons makes these two lines of interpretation compatible.

Distinguishing persons from subjects of experience as Strawson does solves a number of problems. It means that Butler’s famous objection that “consciousness of personal identity presupposes, and therefore cannot constitute, personal identity” simply misses the point:

Personal identity through time, understood as human subject-of-experience identity through time, is indeed presupposed. It’s taken as given by Locke … In saying that Consciousness constitutes Personal identity, he’s merely offering a criterion for picking out those parts of such a subject’s past that that subject is still responsible for (96).

Indeed, rebutting the “widely held” charge that “Locke’s account of personal identity … is circular and inconsistent, and blatantly so” is one of the main motivations of Strawson’s book (1).

A second alleged advantage of Strawson’s interpretation is that it can deal with the notoriously tricky fatal error passage (2.27.13), while the ‘radical theory’ that consciousness alone makes identity – which he uses as his foil throughout the book – cannot (121).[[9]](#footnote-9) On the radical theory, if consciousness of an action is transferred to you, then it’s your action and you can be justly punished or rewarded for it. Thus, there’s no reason for God’s justice to prevent transfers of consciousness. And thus, if Locke really held the radical theory, he would be contradicting himself when he claimed that “the goodness of God” will not allow creatures to “by a fatal error of theirs transfer from one to another, that consciousness, which draws Reward or Punishment with it” (2.27.13).

The difficulties posed by the fatal error passage are well known, and Strawson is right that we’re better off if we can avoid attributing self-contradiction to Locke. But Strawson proposes two alternate explanations of the fatal error, and they both have problems as well. The first explanation is that Locke is “making a slip while his meaning remains quite clear” (116). When he says that God won’t allow “one intellectual substance [to] have represented to it, as done by itself, what it never did, and was perhaps done by some other agent”, Locke is switching the referent of ‘agent’ from a person to an intellectual substance. So according to Strawson’s first explanation, Locke is confused about one of the central distinctions of 2.27 – the distinction between a person and a thinking substance. Attributing this confusion to Locke might be better than attributing an outright contradiction to him, but it’s still not great.

On Strawson’s second explanation, there is no slip. Rather:

As things actually are, all **I**-transfers [i.e. transfers from one immaterial substance to another] would in fact be **P**-transfers [i.e. transfers from one person to another], so God has a reason to prevent them … Nevertheless **I**-transfers that are not **P**-transfers are possible as far as we know. There is no logical or conceptual reason why we should think that Personal identity requires sameness of immaterial substance (116-117).

Transfers from one person to another would be unjust, so God prevents them. And since “[a]s things actually are” (116), transfers from one immaterial substance to another would be transfers from one person to another, God prevents transfers from one immaterial substance to another too.[[10]](#footnote-10) Thus the radical theory must be false. Consciousness alone does not make personal identity. It’s merely what determines which of a subject’s actions can be attributed to the person she now is.

If you grant that “as things actually are”, any transfers from one soul to another would also be transfers from one person to another, then Strawson’s second explanation works. But – as I’ll explain in section V below – I don’t think we should grant this.

A third advantage of the way Strawson distinguishes persons from subjects of experience is that it makes misappropriation impossible. Because subjects can only be conscious of their own past actions and experiences,[[11]](#footnote-11) they cannot wrongly acquire responsibility for an action that was never performed, or that was performed by some other subject. And because what makes a present subject responsible for a past action is that she is now conscious of it, there’s no sense in which she *should* be conscious of any past action she is not actually conscious of.

Indeed, Strawson thinks that subjects of experience typically are conscious of only a few past actions:

[T]here are likely to be many, many actions that we *considered as human beings* have performed which we won’t be Conscious of anymore, and which therefore won’t be actions on our part *considered as Persons* – so that we won’t be responsible for them (27).

[S]uppose the initial pool of candidates for being part of the Person that I now am includes all the actions and experiences … that have ever been part of the total spatio-temporal career of the human being known as GS. Only a few of them will pass the test of being part of the Person that I now am, for only a few of them will pass the accessibility-to-present-Consciousness test (37).

The actions and experiences of which [a subject, say,] John is Conscious, and which therefore form part of his Personal identity, are a (small) subset of the complete set of the actions and experiences of John the human subject of experience (121-122).

It might be unnecessary to quote three separate passages: Strawson’s prose is very clear. However, I find the claim that subjects of experience are only conscious of a small number of their past actions so strange that his commitment to it needs underlining.

It strikes me as very un-Lockean.[[12]](#footnote-12) A comparison to Thiel’s Locke – who is a practitioner of what Strawson’s father would have called descriptive metaphysics, at least in the case of personal identity – is useful. Thiel explains that

The importance [Locke] attaches to issues such as agency, self-concern, and moral and legal responsibility require an approach that is consistent with ‘how we normally speak and think’ (126 n. 18).

This is exactly right. Locke thinks we already have a great many true moral beliefs – many of them obtained on the basis of Revelation – and wants to articulate a conception of personal identity that helps ground those beliefs. Thus his account of personal identity is supposed to fit our ordinary practices fairly well.[[13]](#footnote-13) Moreover, there’d be no point in using thought experiments involving responsibility to test our intuitions about personal identity if you thought we were deeply confused about what we’re responsible for. And we ordinarily think of ourselves as responsible for the vast majority of actions performed by the subjects we are.[[14]](#footnote-14) Thus it seems to me that if Locke’s account turned out to imply that subjects are only responsible for a few of their actions, it would be a failure on its own terms.

Strawson might reply that I’m focusing on the wrong thing. What subjects are responsible for doesn’t matter. (It’s probably a category mistake to ask what subjects are responsible for.) What matters is what persons are responsible for, and Lockean persons *are* responsible for the vast majority of their actions.

But I am not convinced. One way to get at what’s troubling me here is to consider what God will do about actions that have been lost to consciousness on Judgment Day. Strawson explains that

The Day of Judgment plays an important part in Locke’s discussion of personal identity, but those who – like myself – don’t believe in God or the Day of Judgment shouldn’t think that their place in Locke’s theory makes it less interesting. All questions about the Day of Judgment can be converted into wholly earthly questions about one’s overall moral status or being as one stands here now, firmly on the ground, with no prospect of going anywhere else. The idea of the Day of Judgment is, no doubt, a fantasy, but the fundamental idea behind Locke’s discussion of personal identity doesn’t depend on it in any way … the respect in which his conception of personal identity … is intuitively natural is independent of the story of the Day of Judgment (xii-xiii).

I share Strawson’s views on the Day of Judgment and on the truth of Locke’s religious beliefs more generally. But I worry that they have led him to treat Locke’s religious beliefs as more marginal to his account of personal identity than they actually are. (I worry about this in my own work too.) And although I agree Locke’s theory remains interesting if you reject God and the Last Judgment, I am unconvinced that all questions about the Day of Judgment can be converted into earthly questions.

Here’s an example of an action that’s lost to consciousness on Judgment Day. (It also suggests an example of a question that cannot be converted into earthly terms.) Say that a subject steals a horse, and on the Day of Judgment is no longer conscious of the theft. That subject cannot be punished for the theft, since she’s not conscious of it and hence not the person who stole the horse. But given what persons are, they can’t be transferred between subjects. So, on Strawson’s view, there’s no way for God to resurrect the person who stole the horse. And so, on Strawson’s view, there’s no one who can be held responsible for the theft. This hardly seems consistent with the demands of divine rectification.

**IV**

I’m tempted to describe the claim that subjects are only conscious of a small number of their past actions as a fatal flaw in Strawson’s account. But this isn’t fair, since it could easily be excised. If you distinguish persons from subjects, you open up the possibility of past actions that can be attributed to a present *subject* but not a present *person*. But you could still say that in the actual world, subjects are conscious of the vast majority of their past actions. This wouldn’t obviate the distinction: even if subjects are conscious of all the actions they performed in the past, it would still be their consciousness of those actions – not merely the fact that they performed them – which made them responsible for them.

A comparison with appropriation interpretations like my own might help. I think it’s crucial for such interpretations to distinguish between what a person appropriates and what’s available to her for appropriation. Perhaps, for instance, only the actions you remember are available for appropriation. But appropriation interpretations can still say that as a matter of fact, persons appropriate more or less everything available to them. This doesn’t make them collapse into a memory interpretation: the reason a particular past action is *my* past action is that I’ve appropriated it, not that it’s available for appropriation.

You might think it’s an empirical question what proportion of her past actions and experiences a given subject is now conscious of. But it also depends on what’s involved in consciousness. Strawson says very little about what the ability to experience something immediately, from the inside, amounts to when he first introduces the distinction between occurent and dispositional senses.[[15]](#footnote-15) But later he says that consciousness, understood dispositionally,

Certainly reaches beyond what one can now bring back to mind completely unaided, and perhaps also beyond what one can remember when suitably prompted or shocked. For it reaches back – in Locke’s view – to all of those past actions that one is Conscious of in one’s conscience or “heart” … [This] may not be fully known or accessible to one now, but that doesn’t matter, for there is nevertheless a fact of the matter about what one is Conscious of (52).

Let’s think about the Day of Judgment again. On Judgment Day, Locke tells us (quoting 1 Corinthians), the secrets of all hearts shall be laid open and “The Sentence shall be justified by the consciousness all Persons shall have, that they *themselves* … are the *same*, that committed those Actions, and deserve that Punishment for them” (2.27.26). Strawson explains that

The only actions or experiences that are up for punishment or reward are the ones … one is then Conscious of … *Vast quantities* of one’s life’s actions and experiences are unrecorded. (If one’s whole life flashes before one’s eyes, few events will snag on one.) Those that remain are all and only those one still feels involved in, in such a way as still to feel that they’re something that *one* did or experienced, where “*one*” means the person one feels or experiences oneself to be … and where this feeling or experience is not a matter of one’s choice, but a matter of what shows in one’s heart when its secrets are laid open. This tiny subset of the actions of the human being that one is constitutes the Person … one now is (142-143).

One way to understand what Strawson thinks is going on is that God restores autobiographical memory, thereby making a subject’s whole life flash before her eyes. When God does this, the subject need not feel involved in all the actions and experiences she remembers. In other words, she need not be concerned for all of them.[[16]](#footnote-16) And she will only be held responsible for those actions she *is* concerned for.

What determines which actions a subject is concerned for? Here Strawson brings in another ingredient, the notion of moral character:

Locke’s overall picture is also deeply and plausibly informed – it seems to me – by the idea that one’s overall forensic condition, one’s fundamental moral standing, at any time, either now or on the Day of Judgment, lies in one’s overall moral character or moral being at that time (143) … the set of actions and experiences (of the human being that one is) of which one is still Conscious in the accountability-entailing way is at least partly a function of one’s present overall moral character or moral being considered (insofar as it can be) *independently* of the idea of the set of actions and experiences (143-44).

Strawson is careful to say just that this is “an important part” of “a generally speaking Lockean picture” (144). Presumably he recognizes that there’s no talk of moral character in *Essay* 2.27.

One of Strawson’s reasons for emphasizing the notion of moral character is interesting. We should, he argues, emphasize the notion of moral character because it helps us make sense of repentance. If I have genuinely and sincerely repented for some past action, I do not have the same character as when I performed it. Thus, I am not the same person, and thus, I am not responsible for that action.

I hadn’t thought about how repentance fits into Locke’s account of personal identity before reading Strawson. (Thiel (139) also credits Strawson for bringing the issue of repentance to his attention). The link between repentance and moral responsibility is interesting and important. However, I’m not convinced that the link is via personal identity. Locke thinks that we won’t be punished for the sins we’ve repented, and he clearly thinks repentance is important: “faith and repentance … are the indispensable conditions of the new covenant, to be performed by all those who would obtain eternal life (*The Reasonableness of Christianity* §172). But he doesn’t mention repentance at all in *Essay* 2.27. One obvious explanation is that he doesn’t think repentance makes s a different person but that it eliminates punishment for some other reason.[[17]](#footnote-17)

**V**

Now that we’ve seen some implications of Strawson’s distinction between subjects and persons, let’s go back to what he thinks persons are modes *of.* Strawson doesn’t use the term ‘mode’ at all. But he explains that Locke’s term ‘person’ has:

[T]he odd property of being naturally used to denote a thing or object of some sort, and not merely a property of a thing (you and I are Persons, and we’re not, we suppose, just properties or sets of properties), even while at the same time operating in effect as a term for something that we do take to be a property or aspect of that most paradigmatic of things or objects, a … human being … In fact many terms have this kind of doubleness; consider the words ‘seamstress’’ and ‘carpenter’ (13).[[18]](#footnote-18)

Or, I add, constable, father, and dictator, which Locke classifies as relations (2.25.4, 2.28.3) but which denote things and not just relational properties. Thus,

A Person is indeed an object of a certain sort, and must exemplify a certain sort of temporal continuity, if it is to continue to exist … a Person is also an object-considered-specifically-in-respect-of-a-certain-property. It’s a subject-of-experience-considered-in-the-Ciceronian-Hobbesian-Pufendorfian way, or, in Locke’s version of this idea, a subject of experience considered – at any particular time at which it is considered – specifically in respect of the reach of its field of Consciousness at that time (93).

This emphasis on the “Ciceronian-Hobbesian-Pufendorfian” way of thinking is something we saw already, in Thiel, and indeed Strawson credits Thiel 1998 here.

Reading this, part of me wanted a little more precision. But ultimately, I think, Strawson is wise to leave it where he does. He would probably have been even wiser to just talk about how the term ‘person’ functions, and not bring up what persons are at all. For once we start talking about what persons are, it’s very easy to start reifying them – and then start worrying about some kind of overpopulation problem. (Aren’t the person and the thinking substance two different things in the same place at the same time, which Locke can’t allow?) Strawson addresses this worry by insisting, quite rightly, that the mode is not a thing in addition to the substance it modifies:

There isn’t any mysterious duplication or ghosting here, giving rise to Persons that are somehow additional to subjects of experience. John the Person standing before us now *is* John the human subject of experience standing before us now. John the Person is simply John the human subject of experience considered specifically with respect to his overall moral or forensic status (129).

Thus, overpopulation worries are misplaced.

However, although there’s no reason to worry about the relationship between persons and subjects, I think there *is* reason to worry about the relation between subjects and their substantial realization – bodies (if materialism is true) or body-soul composites (if materialism is false). Let’s assume materialism for the time being, for simplicity. Strawson never mentions the distinction between organisms and masses, but it’s natural to ask whether the bodies in question are organisms or masses of matter. [[19]](#footnote-19) In other words, are subjects *realized* by living organisms? Or *are* they living organisms?[[20]](#footnote-20)

I spent quite a lot of time trying to figure this out – far more than should have been necessary. At one point Strawson remarks that the book may “be judged to be more difficult than it is, because of the respects in which it departs from patterns of interpretation . . . that have become entrenched” (xiii). I did find the book difficult – surprisingly so, given that it’s a short, non-technical book with a clear prose style. But I don’t think Strawson’s diagnosis of the difficulty is correct. He admits that the book “lacks a standard expository structure” (3) – it’s a record of his thoughts as he read 2.27 over and over – and I think this is the source of its difficulty. The structure has advantages as well as disadvantages: the book is fun to read and has some narrative drive. But at times it’s more work than it should be to figure out Strawson’s position.

Here’s a reason to think that subjects of experience just are living organisms. It’s the only way to make sense of Strawson’s remarks about materialism:

Suppose … that materialism is true, and that one’s whole psychological being … is wholly located in one’s brain. Suppose further that all the individual material particles composing one’s brain have over the years been replaced many times. Is this something to worry about …? Plainly not. This is how things actually are with us (so Locke suspects …), and it doesn’t put our continuing existence in question in any way at all. The … subject of experience that one is, survives the process of complete substantial turnover with its diachronic identity untouched. So too, the person that one is … (97).

Or again:

[T]he materialists who ‘place thinking in a system of fleeting animal spirits’ are already committed to denying that the persistence of a single continuing substance is necessary for personal responsibility. They’re already committed to the view that persons can survive change in their thinking substance (given that they hold that persons’ thinking substance is wholly material (103; cf. 125).

Sameness of substantial realizer requires sameness of material particles only if the realizer is a mass of matter. Materialists are only committed to *actual* survival through change in thinking substance if by ‘thinking substance’ they mean ‘mass of matter’.

Here’s a second reason to think that subjects are living organisms (still assuming materialism). Strawson says that Locke assumes the existence of subjects who can survive change in their substantial realization. If he means that Locke assumes the existence of subjects who can survive change in the organism that realizes them, this is crazy. None of Locke’s contemporaries would have found this assumption plausible or even familiar. Hence it would be very odd for Locke to make the assumption without or even mentioning it, let alone arguing for it.

Here, on the other hand, is a reason to think that subjects are *realized* by living organisms. If subjects just *are* living organisms, it’s hard to see how persons could be transferred from one organism to another. Locke’s thought experiments would have us imagining something impossible.

It turns out that this isn’t much of a reason. For Strawson thinks that Locke’s thought experiments *are* supposed to be impossible. Consider what he says about the Nestor case:

[A] large part of its force derives precisely from the fact that such a case is in fact impossible. It’s a *per impossibile* point designed to make the force of the notion of Consciousness clear: if, *per impossibile,* one really did somehow become Conscious of one of Nestor’s actions, one would indeed be the same person as Nestor, as far as that particular action was concerned; *that’s* how strong the notion of Consciousness is (128).

Strawson doesn’t actually say why the Nestor case is impossible or what sense of possibility he has in mind. But – now giving up the assumption of materialism, since the Nestor case involves soul switching – if subjects simply *are* body-soul composites, then soul switching is metaphysically impossible, and obviously so.

This is implausible. If Locke thought that soul switching is metaphysically impossible, he would not have said the following:

Whether if the same thinking substance … be changed, it can be the same person … cannot be resolved, but by those, who know what kind of substances they are, that do think; and whether the consciousness of past actions can be transferred from one thinking substance to another (2.27.13).

Moreover, as William Uzgalis pointed out in his review of Strawson’s book, Locke clearly does think that body switching is possible. For Locke thinks that as far as we know, we will be resurrected with different bodies – not simply different masses of matter, but different living organisms. And since he also thinks that materialism is epistemically possible, he must think that body switching is epistemically possible. This rules out any interpretation that commits him to the belief that body switching is metaphysically impossible.

**VI**

Strawson uses the radical theory that consciousness alone makes identity as his foil throughout the book. He argues that the radical theory is philosophically incoherent. If consciousness alone makes identity, then any collection of actions and experiences whatsoever can be the actions and experiences of a single person. But this is false:

A person … can’t just be an action-and-experience pot into which one can put absolutely any action and experience … a person … is, necessarily, something that has a certain personality or moral-characteral coherence. This is something Locke takes for granted, in assuming a continuing human subject of experience as the entity for whom the question of its Personal identity … arises … Otherwise we no longer really have any idea what we’re talking about. Or rather: we’re certainly not talking about persons … No account of Locke’s views can be correct if it attributes to him a theory of what persons are … that threatens to conflict with this basic assumption about the nature of persons – the assumption that they have, necessarily, a certain moral-characteral coherence (132-133).

Here’s an example of a being whose character is not coherent enough for her to count as a person:

[S]uppose the thousand actions or experiences transferred to [a subject are] … quite fantastically disparate, coming from a thousand different cultures, two sexes, and several different sexual orientations. There can be no characteral coherence in a subject of experience so constituted. It’s natural to think that there can’t really be a person there at all, a single entity that we can think of as a person (129).

Strawson doesn’t give any textual evidence that Locke accepts the need for characteral coherence. Relying on the intuition that such cases are impossible, he claims that Locke would not even have entertained the idea of ‘stewpot persons’ (130).

I do think it’s sometimes enough to say that a philosopher cannot hold a theory just because that theory is philosophically incoherent. But if you’re going to say this, there better be some consensus that the theory really is incoherent. It can’t just seem incoherent relative to your particular intuitions. And I don’t share Strawson’s sense that the being just described could not be a person. I’ll grant that it’s a bit hard to imagine a subject of experience who’s conscious of all these disparate actions and experiences. But if there were such a subject, it seems obvious that she’d be a person like myself, only with a more interesting history.

This is an instance of a more general worry I have. I assume that Strawson was attracted to his topic because he saw something true and important in Locke’s account of personal identity. And it’s not a vice for historians of philosophy to have philosophical views, or for those views to influence their interpretations. We can’t really avoid having views, and we can’t really avoid reading other philosophers through them. Nevertheless, in Strawson’s discussion of stewpot persons – as in his remarks about the Day of Judgment and his suggestion about how few past actions subjects are now responsible for – I worry that Strawson’s Locke is too close to Strawson.

**VII**

There’s a spectrum of methodology in the history of philosophy today. Strawson stands at one end, and Thiel near the other. Thiel has read and synthesized an enormous amount of material, – both little-known primary texts and secondary literature – and for that reason alone his book is an invaluable resource. The result is a long, richly contextualized, clearly laid out, and for the most part carefully argued book, which puts together a fascinating picture of late-17th-century debates. In contrast, Strawson’s book refers to little secondary literature and almost no primary literature beyond the *Essay.*

I have spent far more time on Strawson’s book than on Thiel’s. This is not because I thought it was the better book, but because of the kind of book that each one is. Strawson’s book is essentially one long argument, and the conclusion it argues for is novel and interesting. I’ve argued that Strawson’s interpretation can’t be right. But this isn’t a reason not to read his entertaining, original, and smart book.

In contrast, Thiel’s book as a whole has no one line of argument, no over-arching interpretive position. His discussion of Locke is unified by the two claims from Law I began with, two claims whose truth and importance I entirely agree with him about. There are some minor points I disagree with, but to mention them in this context would be nit picking. I learned a great deal from Thiel’s book, and I wish I’d been able to read it before writing my own. Anyone interested in Locke or in personal identity should read both Strawson and Thiel right away.

**References**

Margaret Atherton (1983). “Locke’s Theory of Personal Identity”. *Midwest Studies in Philosophy* 8, 273-93.

Dan Kaufman (2008). “The Resurrection of the Same Body and the Ontological Status of Organisms: What Locke Should Have (and Could Have) told Stillingfleet”. In ed. Hoffman, Owen, and Yaffe, *Contemporary Perspectives on Early Modern Philosophy: Essays in Honor of Vere Chappell.* Broadview.

John Locke (1801). *The Works of John Locke*, *in Ten Volumes.* London: J. Johnson, etc. Facsimile edition Elibron, 2005.

Antonia LoLordo (2013). “Three Problems in Locke’s Ontology of Substance and Mode Ontology”. In *Contemporary Perspectives on Early Modern Philosophy*, ed. Martin Lenz and Anik Waldow. Springer.

\_\_\_\_\_\_\_\_\_ (2012). *Locke’s Moral Man.* Oxford, 2012.

Udo Thiel (1998). “Personal Identity. In *The Cambridge History of 17th Century Philosophy.*  Edited by Michael Ayers and Daniel Garber. Cambridge, 1998.

William Uzgalis (2012). Review of Galen Strawson, *Locke on Personal Identity: Consciousness and Concernment. Notre Dame Philosophical Reviews* 1202.02.07. http://ndpr.nd.edu/news/28700-locke-on-personal-identity-consciousness-and-concernment/

Shelley Weinberg (2012). “The Metaphysical Fact of Consciousness in Locke’s Theory of Personal Identity”. *Journal of the History of Philosophy*, 50: 387-415.

1. Thiel’s discussion stretches from Descartes to Hume. I’ll confine myself to the chapters on Locke and his critics, which occupy roughly a third of the book. [↑](#footnote-ref-1)
2. Locke, *Works* 3.179–180. Law’s *Defense* is reprinted in its entirety as Appendix 2 of Strawson’s book. Strawson’s Appendix 1 reprints *Essay* 2.27, with Strawson’s ‘translation’ into contemporary English – intended for students or anyone else who finds 17th century English difficult – on facing pages. [↑](#footnote-ref-2)
3. However, Thiel qualifies his admiration somewhat, noting that “While Law seems to capture the notion of a Lockean person correctly, accounting for it in terms of a mixed mode, he overemphasizes the forensic aspect of that notion” (209). [↑](#footnote-ref-3)
4. We’ll see that ‘subject’ is a crucial ontological category for Strawson. But despite the title of Thiel’s book, the category ‘subject’ doesn’t doany real philosophical work for him; it’s just shorthand. [↑](#footnote-ref-4)
5. The relevant notion of dependence is ontological, but it’s not the inherence relation traditionally held to obtain between substance and mode. If modes depended on substances by inhering in them, qualities would be modes – but they are not. And many of Locke’s examples of modes aren’t properties, so they can’t inhere in anything: consider, for instance, a triangle (2.12.4), theft (2.12.5), the number two (2.13.1), an inch (2.13.4), a tune (2.18.3), or a rainbow (2.18.4). See LoLordo 2013 for more on dependence. [↑](#footnote-ref-5)
6. Locke articulates persistence conditions for organisms. But since organisms are paradigmatic substances for him, their real essences are unknown. Hence the persistence conditions he provides must be derived from their nominal essences instead. [↑](#footnote-ref-6)
7. Strawson capitalizes terms like ‘Personhood’, ‘Personal identity’, and ‘Consciousness’ to emphasize that they’re being used in a technical sense. I won’t follow him in this. [↑](#footnote-ref-7)
8. He does not mention any of these alternate interpretations. [↑](#footnote-ref-8)
9. Oddly, especially given how comprehensive the book typically is, Thiel does not try to explain the fatal error passage. [↑](#footnote-ref-9)
10. I’m not entirely sure what Strawson means by “as things actually are”. He doesn’t mean “in the actual world”, since the point of the fatal error passage is that in the actual world, there *aren’t* any transfers from one immaterial substance to another. Perhaps he means something like “according to the actual laws of nature”. [↑](#footnote-ref-10)
11. I am not entirely sure whether this is supposed to be an empirical claim, or whether it’s built into Strawson’s notion of consciousness. [↑](#footnote-ref-11)
12. I also find it philosophically odd. Strawson refers to “*the fact that* you may no longer rightly be held responsible for everything you did” (92, italics mine). Given the ‘everything’ and the ‘may’, this is obviously true. I can no longer be rightly held responsible for everything I did because I’ve done some things I couldn’t rightly have been held responsible for even at the moment I did them. (The things I did as a toddler, for instance.) But Strawson means something much more radical: that I (as a person) am no longer responsible for many of the actions that I (as a subject) performed in the past and that I (as a person) was responsible for at the moment when I performed them. This seems to me to imply a revision of our ordinary moral practices that’s drastic enough to almost destroy them. Strawson might not be too worried about this, however. At one point he says that “Locke’s theory of personal identity may be completely right … *given the assumption that the notion of moral responsibility is a coherent one at all*” (92, italics mine). [↑](#footnote-ref-12)
13. I don’t think it has to fit *perfectly.* Locke might well think that once we see what personal identity consists in, we will make some minor adjustments to the way in which we assign praise and blame. But I do not think that he allows the possibility that those practices are fundamentally wrong. [↑](#footnote-ref-13)
14. Only the vast majority, since we make an exception for the actions of toddlers, people in certain altered states of consciousness, and the like. [↑](#footnote-ref-14)
15. He also glosses it in terms of accessibility: “To be Conscious of *x* at a given time is for *x* to be accessible to one’s experience at that time, not necessarily for one to be actually experiencing *x* at that time” (34). But without explaining what accessibility involves, this isn’t very helpful. [↑](#footnote-ref-15)
16. Is God restoring your consciousness of them? You could say that consciousness, in the technical sense, requires concernment, so a person with autobiographical memory of an action but no concernment in it isn’t conscious of it. Or you could say that subjects are responsible for just those past actions they are conscious of *and concerned in*. As Strawson points out, it doesn’t matter which way you go here. I’ve tended to speak as if consciousness implies concernment. [↑](#footnote-ref-16)
17. Locke describes repentance as “not only a sorrow for sins past, but (what is a natural consequence of such sorrow, if it be real) a turning from them into a new and contrary life” (*The Reasonableness of Christianity* §170). But to think that “new life” implies “new person” would be over-reading. [↑](#footnote-ref-17)
18. Cf. 78: The term person “is sometimes better taken as a term for a fundamental property or aspect of a kind of thing. One might say that for Locke, ‘person’ functions more like ‘baker or ‘rogue’ than ‘human being. For ‘baker’ and ‘rogue’ function smoothly as names for things in spite of the fact that they highlight properties of things”. [↑](#footnote-ref-18)
19. Failure to draw this distinction also causes trouble in Strawson’s discussion of the Resurrection. He says that

    Locke’s account must provide for something whose reality is in his time publicly presupposed on all sides, i.e. our personal responsibility over time up to and including the Day of Judgment, when we’re resurrected with bodies that can’t plausibly be thought to be particle-for-particle identical with our bodies at any time of our lives (101).

    But while Locke’s contemporaries found it difficult to maintain that the resurrected body is particle-by-particle identical with the pre-mortem body, many of them also found it irrelevant: to be resurrected with the same body is not to be resurrected with the same *particles.* It’s difficult, if not impossible, to make sense of the contemporaries debates about the resurrection of the body without distinguishing masses from organisms. See Thiel, 85ff for an account of those debates. [↑](#footnote-ref-19)
20. You might think the answer is, both. Presumably organisms are realized by masses, so if subjects are realized by organisms, they’re indirectly realized by masses. However, it’ll be clear in a minute that this doesn’t help. [↑](#footnote-ref-20)