

## THERE IS NO PROBLEM OF THE SELF

Eric T. Olson, Churchill College, Cambridge CB3 0DS, UK  
Email: [eto20@cam.ac.uk](mailto:eto20@cam.ac.uk)

---

**Abstract:** Because there is no agreed use of the term ‘self’, or characteristic features or even paradigm cases of selves, there is no idea of ‘the self’ to figure in philosophical problems. The term leads to troubles otherwise avoidable; and because legitimate discussions under the heading of ‘self’ are really about other things, it is gratuitous. I propose that we stop speaking of selves.

---

### I

People often speak as if there were a serious philosophical problem about *selves*. Is there a self? Is the self knowable? How does the self relate to the body? These and other questions are thought to make up something called the *problem of the self*.

I doubt seriously that there is any such problem. Not because the self is unproblematic, or because there are unproblematically no such things as selves. My trouble is that a problem must be a problem *about* something: even if there are no selves, there must at least be some problematic idea or concept of a self, if there is to be a problem of the self. As far as I can see there is no such idea. What is a self? For every answer to this question, there is another answer not only incompatible with it, but wholly unrelated. There is virtually no agreement about the characteristic features of selves: depending on whom you believe, selves may be concrete or abstract, material or immaterial, permanent or ephemeral, naturally occurring or human constructions, essentially subjective or publicly observable, the same or not the same things as people. There are not even any agreed paradigm cases of selves, things we could point to or describe and say, ‘A self is one of *those*.’ But no concept could be so problematic that no one could agree about *anything* to do with it. For lack of a subject matter, then, there is no problem of the self.

So I shall argue, anyway. I am not just quibbling about a word. Real philosophy is at stake. Many philosophers assume that there is something properly called the problem or problems of the self, and write as if everyone knew what they meant by ‘self’. This often leads to obscurity and muddle. If I am right, the muddle arises because those philosophers believe in a concept that doesn’t exist.

I claim also that if we look at those books and articles with titles like ‘The Self’ or ‘Problems of the Self’, we will find that they are most typically not about ‘selves’ at all, but about other things — different things in different cases. In fact the matters discussed under the heading of ‘self’ are *so* various that no one can seriously say that they are all about some one thing, the self. Because legitimate inquiries that go under the heading of ‘self’ are really about something else, and can be (and typically are) put in other terms, we can easily do without the word ‘self’.

For these reasons, I believe that philosophers would do well to avoid the word ‘self’ in their theorizing.<sup>1</sup>

## II

What would the problem of the self be about, if there were such a problem? What is this ‘self’ whose existence and properties appear to be the subject of so much debate? What distinguishes problems about the self from problems about other topics — causation or intrinsic value, say? If there is any genuine problem of the self, there must be *some* nontrivial answer, however vague or incomplete, to these questions.

There are several ways of trying to meet this demand. The most satisfactory would be to give a definition of the term or an analysis of the concept. A definition of ‘self’ would say, ‘ $x$  is  $y$ ’s self if and only if . . .’, with the dots filled in by some synonymous or at least logically equivalent phrase. I say that the phrase to be explained is ‘ $x$  is  $y$ ’s self’ rather than ‘ $x$  is a self’ (less formally, ‘one’s self’ rather than simply ‘a self’) because it is part of the meaning of the word ‘self’, in its typical philosophical uses, that a self be *someone’s* self. (That much, anyway, seems clear.) Or if there could be ‘unowned’ selves, it is part of the idea of a self that it at least be the sort of thing that *could* be the self of a particular person or thing. If someone could explain ‘ $x$  is a self’ but not ‘ $x$  is  $y$ ’s self’, or if someone could distinguish selves from non-selves but had no idea what made a given self the self of a particular person, I think most ‘self’-users would doubt that he knew what the word meant. The word ‘self’ is like the word ‘body’, as in the phrase ‘human body’: anyone who knows what the words ‘human body’ mean must have at least some idea of what it is for something to be the body, or at least a body, of a particular person (van Inwagen, 1980). We can easily define ‘ $x$  is a self’ in terms of ‘ $x$  is  $y$ ’s self’:  $x$  is a self just in case there is (or could be) some  $y$  such that  $x$  is  $y$ ’s self. But there is no obvious way of deriving an account of the meaning of ‘ $x$  is  $y$ ’s self’ from a definition of ‘ $x$  is a self’. Thus, even if we succeeded in defining ‘ $x$  is a self’ we should still face the task of accounting for ‘ $x$  is  $y$ ’s self’.

If we use a word that we can’t define, or that we can’t define in a way that others who use it will accept, we must be on our guard — the more so if the word is a piece of philosophical jargon that has no place in ordinary language. (Part of the trouble with ‘self’ is that it is jargon masquerading as ordinary language.) It will be a real question whether those who use the word are in any sense talking about the same thing. Anyone who uses a word whose definition is disputed is obliged to say what she means by it. If we cannot do so, we have to wonder whether we understand what we are saying.

But a word isn’t necessarily in serious trouble just because it has no agreed definition. There are other ways of saying what a self is than by explicitly defining it. We could mention some characteristic features of selves — features that may not necessarily be shared by all and only selves, but which selves at least typically have and non-selves typically lack. A longish list of such features would give us a fair idea of what a self is.

Failing both a definition and a list of characteristic features, we could at least begin to answer the question, What is a self? by referring to some paradigm cases of selves — particular things that everyone agrees are selves, or that would be selves if there were any selves — and some paradigm cases of non-selves. Of course, this strategy faces well-known difficulties: even if you can figure out which particular things I am referring to, you might not be able to guess what it is about those things that makes them all selves, and so have no idea what other things count as selves or why. (If I point to a shoe, a lampshade, a book, and a carrot, and say, ‘*Melves* are things like

those,' obviously that won't suffice to tell you what I mean by 'melf'; nor will it help much if I go on to tell you that stones and forks are not melves.) Still, I should have given you at least *some* information about what a self is.

Let me illustrate these rather abstract points with a concrete example. There is no agreed account of the meaning of the term 'mind' or 'mental', of what makes a phenomenon mental rather than non-mental. Intentionality ('aboutness'), subjectivity ('what it is like'), immediate accessibility to consciousness, and many other features have been proposed as 'marks of the mental'; but none of these accounts is widely accepted. This is a problem. In a sense, we don't know what the mind is. At the same time, we know rather a lot about the mind. Nearly everyone agrees that mental phenomena, if there are any, often *have* intentionality and subjectivity, and are often accessible to consciousness, even if these features don't suffice to define the mental. We agree on a wide range of typical and characteristic features of mental phenomena: everyone agrees, for instance, that the desire for rain is satisfied by rain and frustrated by persistent dry weather, and that being in the presence of rain usually causes one to believe that it's raining. And we agree on paradigm cases: no one doubts that beliefs, wants, memories, intentions, sensations, emotions, and dreams, if there are such things, are mental phenomena, and that earthquakes and temperatures are not. Even those who deny that there are any mental phenomena can reasonably claim to understand what it is whose existence they are denying. Thus, although there is some dispute about just what the subject matter of psychology or the philosophy of mind may be, there seems to be enough agreement that many of our questions can be said to be questions *about the mind*.

On the other hand, if a large minority of philosophers and psychologists thought that beliefs, though they exist, were not mental phenomena at all, but took something completely different — photosynthesis, say — as paradigmatic of the mental, it would be clear that there was serious confusion afoot. Matters would be even worse if there were respected participants in the debate who thought that many non-mental but few mental phenomena had properties like intentionality or subjectivity. Although there would certainly be a problem about the *word* 'mind' in that case, it would be a good deal less clear that there was a philosophical problem *of the mind*. For what, beyond the word, could it be a problem about? The case of the self is rather like that.

### III

Let us consider some attempts to say what a self might be. It doesn't matter whether we take these proposals to be definitions of 'self' or only as giving essential or paradigmatic or at any rate salient features of selves. Take them as proposed answers to the ordinary question, What is a self?

*Account 1.* One's self is that unchanging, simple substance to which one's impressions and ideas have reference.

This is apparently what Hume sought in vain within himself (1978, p. 251). We may wonder what Hume meant by ideas 'having reference to' one's self. But the main thing to be said about this account is that most present-day philosophers see no reason to believe that any unchanging, simple substance has anything to do with anyone's ideas or impressions. If this is what selves are, there are simply no selves, and the

problem of the self is of little more than historical interest. For an account of 'self' that reflects today's concerns we must look elsewhere.

Here is an updated version of something like Hume's account:

*Account 2.* One's self is the inner subject of one's conscious experiences.

(Campbell, 1957, p. 74; Harré, 1987, p. 99.) Variants include 'One's self is the bearer of one's personal identity over time' (Berofsky, 1995, p. 234); '... that which views the world through one's eyes' (Nagel, 1986, p. 55); '... the tautological subject of one's actions' (Rée, 1974, p. 188n.); 'the cause of everything one does' (Minsky, 1985, p. 232); etc.

Now I take it that *I* am the subject of my conscious experiences, the bearer of my identity over time, and the cause of everything I do. Otherwise they wouldn't be *my* experiences, *my* identity, or things that *I* do. On this account, then, I am my self, and you are yours. Everyone is identical with his self. If there should turn out to be no selves, there would be no such beings as you and I. And although some philosophers accept this inference, others reject it: they think that the existence of selves is problematic in a way that the existence of midwives and plumbers, or of you and me, is not (e.g. Harré, 1987, p. 103). In fact those who argue that 'there is no self' rarely go on to conclude from this that they themselves do not exist. *Their* problem of the self must be about something else.

Although the above account seems to imply that selves are people, it doesn't say that the two terms are synonymous. But many philosophers use the terms interchangeably (e.g. Mellor, 1991; Bermúdez, 1997). So we might say:

*Account 3.* One's self is just that person, himself.

Selves are just people, human beings. To say that the self is so-and-so is to say that people, you and I, are so-and-so. This makes it rather awkward to say that one *has* a self. What could it be for me to *have* a person, or to *have* myself? The truth, rather, is that we *are* selves. But we could presumably explain away talk of having a self as a hangover from the eighteenth-century use of the word 'self'.

If this account were widely accepted, I shouldn't be writing this essay. To be sure, the word 'person' is problematic enough that anyone who relies on it in philosophical discussion had better say what she means by it. There is a good deal of dispute about its definition, and about whether it can be defined at all. Compared with 'self', though, 'person' is a model of clarity and accord. Everyone agrees that people (persons) characteristically have certain mental capacities such as rationality and self-consciousness, and certain moral attributes such as accountability for their actions (or at any rate that they are capable of acquiring such features). There is a fair consensus about what things count as people: no one doubts that you and I and Boris Yeltsin are people, and that houses and bronze statues of people aren't. Although there are disputed cases (foetuses, infants, adults suffering from severe senile dementia), their number is small compared with the number of items we can confidently classify as people or non-people. In this respect 'person' is no worse off than most other nouns. We know how many people there are (a few hard cases aside, once more) and how to count them. The word 'person' is well enough understood for there to be philosophical problems about people.

But many philosophers explicitly deny that ‘self’ and ‘person’ are interchangeable. For example, some say that dogs and dolphins are, or have, selves, even though they aren’t people (Lowe, 1996, p. 49n.). More seriously, it is often said that ‘person’ has to do with a publicly observable, social being whereas a self is essentially something inner or private (e.g. Harré, 1984, p. 26; McCall, 1990, pp. 12–15; see also Campbell, 1957, p. 93; Hamlyn, 1984, p. 188). Others say that people but not selves can be described, identified, and counted (Abelson, 1977, p. 87). This suggests that you and I are not selves, but rather have them — assuming, at least, that *we* are people, or that we are publicly observable or countable or identifiable. What, then, are selves? We might say:

*Account 4.* One’s self is that indescribable and unidentifiable private, inner being within one.

(‘By “self” I mean the personal unity I take myself to be, my singular inner being, so to speak’ [Harré, 1984, p. 26].)

Now this is completely unhelpful. We might as well say that the self is ‘the elusive “I” that shows an alarming tendency to disappear when we try to introspect it’, a definition (tongue in cheek, I assume) found in a popular philosophical dictionary (Blackburn, 1994, p. 344). You can’t explain what a self is by saying that it is something inner and ineffable, any more than you can do so by saying that it is a certain kind of ‘I’. Of course, if the self really is something ineffable, this is no criticism. You can’t describe the indescribable. The challenge for those who think they understand this account is to persuade the rest of us to take them seriously.

‘One’s self is what one identifies oneself with, what a person cares most about, the loss of which amounts, for him, to self-destruction, either partial or total’ (Abelson, 1977, p. 91; see also James, 1905, p. 291; Berofsky, 1995). Thus we might write,

*Account 5.* One’s self is what one values above all else.

Of course, there may be a number of things that someone values above all else, and whose destruction would be as bad for her as her own destruction. So it would not be uncommon for someone to have several different selves at once, on this account (‘The piano is my second self,’ said Chopin), as well as having different selves at different times; and different people would often share the very same self. The phrase ‘Mary’s self’ wouldn’t necessarily have unique reference. It would be more like ‘Mary’s brother’. In fact, Mary’s self might *be* her brother. There would be people whose selves were houses, political causes, children, or pets. Many of those who use the word ‘self’ would reject this consequence as absurd.

Let us continue our catalogue:

*Account 6.* One’s self is the unconscious mechanism responsible for the unity of one’s consciousness.

(Brooks 1994, pp. 36, 51.) On this account, there is nothing subjective or ineffable or immaterial about the self; nor are selves people. My self may be literally an organ, a part of my brain, a matter for physiologists to investigate.

Others suggest that the self is not a material part of one but rather an attribute, typically an aspect of one’s personality or character or behaviour:

*Account 7.* One's self is a psychological or behavioural attribute of one.

But which psychological or behavioural attribute? No one would count my fear of close spaces as a serious candidate for being my self. One view is that one's self is the way one sees oneself, or a certain set of one's beliefs about oneself: one's self is roughly one's self-image (Rogers, 1951, pp. 498; Harré, 1984, p. 26; Marx and Hillix, 1973, p. 605). I suppose Dennett's account of the self as a 'center of narrative gravity' might be something like this, though it is far from clear just what he means by that phrase (1991, pp. 416, 427). On this view a person's self is typically unique, and one may have different selves at different times, perhaps even more than one self at once. Others say that one's self is by definition unique: it is (for example) what is expressed by one's uniquely characteristic actions (Kenny, 1988, p. 33). In that case, for all I know I may not have a self at all, as I cannot rule out the possibility that somewhere in the universe there is someone exactly like me in the relevant respects. For that matter, it is consistent with everything we know that no one has a self. Even those who consider it an open question whether there are any selves are unlikely to accept that there could fail to be selves for *that* reason. By contrast, others take the self to be 'man as he really is, not as he appears to himself' (Jung, 1968, p. 186), or that collection of features shared by all and only human beings in all times and places (Solomon, 1988, p. 4). One's self is something like universal human nature. The very idea of the self, on this account, ensures that we all have the same self, and that there could not be more than one.

Naturally none of these variants will be acceptable to those who think that *we* are selves, or that selves are things that think or experience. No one seriously supposes that *he* is a psychological attribute. No psychological attribute could think about Vienna, or sleep badly, or drink coffee.

Let us consider one more popular account:

*Account 8.* One's self is an aggregate of or a construction out of one's sense-experiences.

(Ayer, 1946, p. 125; see also Broad, 1925, p. 282; Marx and Hillix, 1973, p. 605.) Let us set aside the serious question of just what an 'aggregate' or 'construction' might be. We might put the view more loosely by saying that one's self is one's mental life, the sum-total of all of one's thoughts and experiences, or perhaps some selected portion thereof. This, I suppose, is the 'bundle theory of the self' often attributed to Hume. Are you and I selves on this account? That depends on whether we are literally made out of thoughts and experiences — whether an aggregate of thoughts and experiences is the sort of thing that can do and be all of the things that you and I can do and be. (Can a bundle of thoughts ride a bicycle? Could a bundle of thoughts have consisted of completely different thoughts?) This is a familiar problem about personal identity.

#### IV

What, then, of the problem of the self? What is it a problem about? There is clearly nothing that satisfies all of these accounts. There aren't even two or three similar kinds of things that each satisfy most of them. It should be equally clear that there is no one thing — no single idea — that all of these accounts could reasonably be seen as trying to capture. There is no one sort of thing that some believe is a construction out

of sense-impressions and others take to be a mental attribute, a simple substance, an organ, a human being, or in some cases even a house or a hamster. It should also be clear that there are no agreed characteristic attributes of selves, or even any generally accepted cases. (We can't even pick out a self in a purely relational way, for example as 'Bertrand Russell's self', without controversy, for on some accounts of 'self' there are no selves to pick out, while on others Russell might have had any number of different selves.) I conclude that those who use the word 'self', if they are saying anything coherent at all, must be talking about completely different things. Thus, there is no such idea as the idea of the self, and therefore nothing for the 'problem of the self' to be a problem about.

There are several replies that someone might make to this argument. First, our catalogue of answers to the question, What is a self? is incomplete. We may have neglected one that would solve the problem. Of course, the mere fact that there are other accounts different from those mentioned is no help to those who believe in the problem of the self. Simply extending the list will only make matters worse. What we need is not just an account of the self that would command wider assent than any of these, but one that would synthesize them and show them all to reflect a part of some larger, common idea. But we can say no more about this possibility until someone produces a candidate for that role.

Second, one might argue that this lack of agreement about selves is part of the very problem I claim not to exist. Doesn't this just show that the problem of the self is more serious than we thought? That it is not only about the nature and existence of selves and their relation to more familiar entities such as people and human beings, but about the very meaning of the term? I can imagine someone comparing the problem of the self with the problem of *race*. There is no accepted account of what the term 'race' or any of its determinates such as 'black', 'white', 'coloured', 'slavic', 'oriental', etc. mean. For this and other reasons, there is no end to the rubbish that has been said and written about race. But aren't those confused writings and sayings *about race*? If there were no problem of the self for the reasons I have given, wouldn't it follow, absurdly, that there was no race problem?

The analogy is strained. Even if no one knows what race or individual races are, there is wide agreement about a great deal of 'race lore'. Whatever 'black' or 'white' might mean, everyone agrees that 'black' people *typically* (though not always) have dark, curly hair and dark skin compared with 'white' people. Everyone agrees that race, if there is such a thing, is an inherited trait, even if there is dispute about how the race of one's parents determines one's own race. Most of us are disposed to class most people into 'black', 'white', etc. in the same way, however arbitrary or unfair those classifications may be. Everyone accepts certain paradigm cases: Nelson Mandela is clearly 'black', Margaret Thatcher is 'white', Mao Tse-Tung is 'oriental', and so on. So even the problematic term 'race' fares far better than 'self' in terms of characteristic features, consistency of application, and paradigm cases.

At any rate, the comparison with 'race' can hardly be cheering for theorists of the self. Of course there is a *social* 'race problem'. But that problem is not so much about race itself, whatever that may be, as about people's attitudes. It arises not because different people belong to different races, but because people think they do, or rather because they treat others differently on the basis of their outward appearance or that of their ancestors. If everyone stopped doing that, the problem would go away. In this

sense there is indeed a ‘problem of the self’. The problem is that people use the word ‘self’ as if everyone knew what it meant when in fact there is no agreement about what it means, and that this leads them into needless troubles. *That* ‘problem of the self’ would be solved if people simply stopped using the word and others like it. But that is not the problem of the self that my opponents believe in. There is a social problem about the *word* ‘self’. There is no philosophical problem about selves.

A third reply would be to concede that there is no single concept of the self, but to insist that there are, nonetheless, many different concepts properly so called. There are different kinds of selves. Since the term is ambiguous, it may be a mistake to speak of *the* self without specifying which ‘self’ we mean. But we do have ‘the Humean self’, ‘the inner-subject self’, ‘the personal self’, ‘the ineffable self’, ‘the evaluative self’, and so on, each with its own set of problems. Don’t these ideas give us a set of problems of the self?

Only in the most trivial sense. Once we concede that the various uses of the term ‘self’ have nothing more in common than the word, we can see that it could be a pun to say that they are all nevertheless problems of the self. We might as well say that the ‘problem of property’ includes both debates about the legal institution of ownership and the metaphysical problem of universals. On the other hand, if the various accounts of the self do have some interesting content in common, we ought to be able to say what it is, and thus give an account of the meaning of the word, however vague or incomplete, that everyone can accept.

## V

I said I was going to argue not only that there is no problem of the self, properly so called, but also that careless use of the term ‘self’ creates trouble we could otherwise avoid. I will mention a few examples of that trouble.

Some philosophers seem to be aware that the term ‘self’ is wildly ambiguous. Owen Flanagan, for example, writes,

The word ‘self’ has many meanings — personality, character, an individual’s central character traits, the way(s) one carries oneself in the world, the way one represents oneself to oneself and to others, the dynamic integrated system of thoughts, emotions, lived events, and so on, that make up who one is from the God’s eye point of view. All these senses are useful (1996, p. vii).

We should expect someone of this view to take care always to explain what he meant when he used the word ‘self’. But Flanagan goes on to say that Augustine’s *Confessions* is ‘the story of a single self’, and to ask whether every individual has ‘one and only one self’, without giving any indication of whether he means central character traits, the way one represents oneself to others, or any of the other items on his list, leaving the reader to guess what claim is intended (pp. 95 f.).<sup>2</sup>

This can lead to more serious trouble. Consider this quotation:

[2] Elsewhere he writes, ‘One useful way of conceiving of the self is as a kind of structured life . . .’ (p. 67), and then says on the next page that the various twists and turns in one’s life ‘have to be part of the life of a single self’. Unless something can both be and have a life, the point would have been more clearly put in different words.



If A's brain is put into B's body, would A's Self move into B's body? Clearly, if the bodies were different (A might be a man, and B a woman) then the Self could hardly be the same — for our notion of Self is surely bound up with our potentialities and our behaviour. At most there could be but a kind of inner core of Self (which might be memories?) remaining after a radical change of body. Suppose, though, that A and B are identical twins and equally fit. Would they swap selves with a brain transplant?' (Gregory, 1981, p. 491)

Well, *would* the result of putting your brain into the head of your identical twin and his brain into yours be that you got what was previously his self and he got the self that was once yours? There are many questions we might ask about this imaginary case. We can ask what would happen to *you*: which, if any, of the two resulting people, if they would be people at all, would be you. We can ask whether the being got by putting your brain into your twin's cranium would have memories, personality, self-image, and other psychological features that were more like yours or more like those of your twin, or which person — you or your twin or someone else — he would think he was. And so on. But what could it mean to ask whether you and he would exchange *selves*? Suppose someone insisted that you would indeed exchange selves. Would you agree or disagree? What sort of argument would support or undermine this claim? What consequences, if any, would it have for the way the other questions are answered? Until the author tells us what he means by 'self', we can have no idea. Yet he apparently sees no need to do so.

The same trouble arises when people ask whether someone with a split personality has more than one self, or whether we create two different selves by cutting the cerebral commissures, or whether one can become a different self through religious conversion, psychic trauma, education, or some other experience. Anyone familiar with philosophical writing that purports to be about the self will recognize claims like these:

The self exists (does not exist).

The self is identical (is not identical) with the body.

The self is not an object.

A human being can have (cannot have) more than one self at once.

Human beings have selves but lobsters don't.

The first-person pronoun purports to pick out a self *qua* self.

Without further elucidation, I don't understand any of these claims well enough to know even whether I agree or disagree. And for anyone who thinks he understands them perfectly well, there will be another who feels equally confident, but who understands them in a completely different way. The reason is that unless the word 'self' as it figures in these sentences is explicitly given some special meaning, they simply don't say anything at all.

Let us consider one more example. Galen Strawson writes,

I will call my view the Pearl view, because it suggests that many mental selves exist, one at a time and one after another, like pearls on a string, in the case of something like a human being (1997, p. 424).

Each self, apparently, exists for as long as one's attention is focused on some one thing, typically two or three seconds, then perishes. But what is it that a human being gets a new one of every few seconds? That is not so clear, despite Strawson's attempts to explain what he means by 'self'. The idea seems to be something like this:

One's self is that distinct, mental thing within one that is a subject of experience and a single thing within any hiatus-free period of experience.

(The bit about one's self being 'within' one is a guess; Strawson doesn't say what makes a self the self of a particular person.) I take it that something is a 'subject of experience' just in case it *has* experiences — just in case it sees, feels, hears, and so on. But *I* see, feel and hear. The question, then, is whether I am 'distinct', 'mental', or 'a single thing' (it doesn't matter for present purposes what those terms of art mean). If I am, it appears to follow from Strawson's view that I am my self, and that Strawson is Strawson's self. Otherwise, if I have a self at all, it must be a subject of experience other than me — something *else* that sees, feels, and hears everything that I see, feel, and hear. I hope we may assume that there aren't *two* sentient beings living within my skin, I and my self. And Strawson seems inclined to think that there are such things as selves. Thus, although he never says so explicitly, his view seems to be that each person is ordinarily identical with his self.

If this is right, then Strawson is telling us that he himself exists (or existed) for only two or three seconds. Or rather, he, Galen Strawson, does not exist at all — for what could make it the case that one rather than any other of those billion or so two-second beings that are the successive 'selves' of a certain tall, blond-haired human being was Strawson? The article attributed to him was in fact written by a vast committee of authors, none of whom contributed so much as an entire sentence. And the same goes for you and me, unless your attention span is considerably longer: you are not the being who read the previous sentence a moment ago. You didn't exist then. The slightest lapse of attention is literally fatal: it destroys you and replaces you with someone else — someone rather like you, but numerically different.

That, surely, is incredible. Anyone who reaches the conclusion that none of us exist, or at least none for longer than a few seconds, must have gone wrong somewhere, just as certainly as someone who concludes that motion is impossible. At any rate, we can accept Strawson's conclusion (or try to accept it) only if it is clear that it follows inexorably from premises whose truth is even more obvious than that you and I exist, and that we existed ten seconds ago. Strawson apparently thinks that it follows inexorably from the fact that our thoughts don't flow in a single, unified stream, but are disjointed and gappy. The premise is probably true. But I can't begin to see how this entails Strawson's conclusion. Even if there *are* things that last only as long as one's attention remains focussed on something — 'sets of neuron-constituting atoms in a certain state of activation' (p. 425), if you like — why suppose that you and I are such things? Why couldn't we be human beings? Why couldn't human beings be the subjects of disjointed thoughts and experiences? What is it about lapses of attention that makes it impossible for *anything* to survive them?

My point is not to criticize Strawson's view as implausible or unsupported, but rather to illustrate the perils of relying on terms like 'self' in doing philosophy. I doubt that Strawson would have reached this absurd conclusion had he put his questions in other terms. If he had simply asked, What sort of thing am I? (Am I a mental thing, a substance, unified synchronically and diachronically, etc.? Am I the sort of thing I ordinarily believe myself to be?) instead of asking about the nature of 'the mental self', I doubt that he would have answered, 'I am a thing that lasts only as long as its attention remains fixed and then perishes and is replaced by something else.' The claim that a human being acquires a new 'mental self' every few seconds may

sound surprising. The claim that *you and I* did not exist five seconds ago turns the world upside down. I suspect that Strawson didn't mean to endorse this view at all, and that he was lured into saying something that implies it by the seductions of the word 'self'. (If he did mean it, why not make it plain? As if the nonexistence of people were a corollary too trivial to mention!) But if that is not what he meant, I have no idea what 'the Pearl view' is meant to be a theory about.

## VI

If there is nothing properly called the problem of the self, what of those books and articles that appear to be about the self? Must we commit them to the flames? Not at all. Despite their titles, they are typically not about 'the self' at all. They are about issues like these:

*Personal identity.* What does it take for you and me to persist through time? What determines how many people there are at any given time? (Is the number of human people always the same as the number of human animals, for example?) What sort of things are you and I? Are we immaterial substances? Mere 'bundles' of thoughts? Living organisms? Material objects different from but 'constituted by' organisms?

*Semantics.* What are the semantic properties of first-person pronouns such as 'I'? How and to what do they refer, if they refer at all? What distinguishes first-person beliefs such as my belief that I have brown eyes from third-person beliefs such as the belief that Eric Olson has brown eyes? Does this difference involve irreducibly subjective facts?

*Philosophy of mind.* What is it for one's mental contents to be unified, and to what degree is this ordinarily the case? What, if anything, causes this unity? To what extent are we aware of what goes on in our minds? What is self-consciousness, and how does it relate to consciousness in general?

*Moral psychology.* What is it that one cares about most, that one identifies with in the sense of regarding its flourishing as a large part of one's own well-being? What makes a project, belief, value, pattern of behaviour, or personality trait autonomous or authentic, fully one's own, and not merely the result of our upbringing or peer pressure? How do these issues connect with moral responsibility?

*Cognitive psychology.* What is involved in forms of reflexive conduct such as knowing one's own mind, in the sense of having settled, consistent and realistic intentions? Or knowing one's own capacities and propensities, in the sense of being able to judge realistically what one can do and is likely to do? Or recognizing one's standing as one agent among others and seeing one's desires in relation to those of others? Or being in command of oneself, in the sense of being able to match one's conduct to one's intentions? How do these reflexive abilities relate to one another? How are they acquired? What happens when they are absent? How does one's mental picture of oneself relate to the way one really is? How does one acquire that distinctiveness that makes one different from others?

*Epistemology.* What are the varieties of first-person knowledge? How do we get knowledge of our own psychological states? Is it something like sense perception? What kinds of first-person knowledge are immune to error through misidentification? What is the nature and extent of proprioception and other forms of first-person knowledge of one's physical properties?

And so on.

You might think that these, and others like them, are the very ‘problems of the self’ whose existence I have denied. I grant that they are not completely unrelated. When you pick up a book whose title includes the free-standing noun ‘self’, you have at least some idea of what it will be about. (It won’t be about cookery or geology.) But we shouldn’t make too much of this. These problems have even less in common than the various accounts of what a self is discussed earlier. If they did have some idea in common, then once again we ought to be able to use that idea to explain what everyone means by ‘self’. Yet no one has been able to do this.

Moreover, all of these problems can be put without using the word ‘self’, as the way I put them shows. *All* of the intelligible content of what are called problems of the self can be captured in this way. If we had answers to all of these questions, there would be no further ‘problem of the self’ remaining to be solved. There is nothing left over that can be expressed only in terms of words like ‘self’. One of the unfortunate consequences of using the word ‘self’ in doing philosophy is that it encourages us to look for entities that we have no other reason to believe in. Once we have accounted for people, their mental features, their relation to those human animals we call their bodies, and so on, we think we need to say something about ‘the self’ as well. There is no good reason to think so.

Or so I claim. Of course, merely putting a number of so-called problems of the self in other terms doesn’t show that the term ‘self’ is superfluous. I may have overlooked legitimate problems or questions that can be put only in terms of the free-standing noun ‘self’ or some equivalent term. In that case I should have to retract my claim that there is no legitimate problem of the self — though the problem so revealed is unlikely to be the problem commonly thought to bear that name.

But if the word ‘self’ really has no agreed meaning, and leads us into troubles we could otherwise avoid, and if we can easily get on with our legitimate philosophical inquiries without it, there can be no reason, other than tradition, to continue to speak of the self.<sup>3</sup>

### References

- Abelson, Raziel (1977), *Persons: A study in philosophical psychology* (London: Macmillan).  
 Alston, William P. (1977), ‘Self-intervention and the structure of motivation’, in *The Self: Psychological and philosophical perspectives*, ed. Theodore Mischel (Oxford: Blackwell).  
 Ayer, A.J. (1946), *Language, Truth and Logic*, 2nd. ed. (New York: Dover).  
 Bermúdez, José (1997). ‘Reduction and the self’, *Journal of Consciousness Studies*, 4 (5–6), pp. 459–66.  
 Berofsky, Bernard (1995), *Liberation From Self: A theory of personal autonomy* (Cambridge: Cambridge University Press).  
 Blackburn, Simon (1994), *Dictionary of Philosophy* (Oxford: Oxford University Press).  
 Bradley, F.H. (1893), *Appearance and Reality* (London: Allen and Unwin).  
 Broad, C.D. (1925), *The Mind and its Place in Nature* (London: Routledge and Kegan Paul).  
 Brooks, D.H.M. (1994), *The Unity of the Mind* (London: St. Martins).  
 Campbell, C.A. (1957), *On Selfhood and Godhood* (London: Allen and Unwin).  
 Dennett, Daniel C. (1991), *Consciousness Explained* (Boston, MA: Little, Brown).

[3] For helpful comments on earlier versions of this paper I thank Shaun Gallagher, Hugh Mellor and two referees for this journal.

- Flanagan, Owen (1996), *Self Expressions: Mind, morals, and the meaning of life* (New York: Oxford University Press).
- Flew, Antony (1949), 'Selves', *Mind*, **63**, pp. 355–8.
- Flew, Antony (1993), 'People themselves, and/or their selves?', *Philosophy*, **68**, pp. 546–8.
- Gregory, Richard (1981), *Mind in Science* (Harmondsworth: Penguin).
- Hamlyn, D.W. (1984), *Metaphysics* (Cambridge: Cambridge University Press).
- Harré, Rom (1984), *Personal Being: A theory for personal psychology* (Cambridge, MA: Harvard University Press).
- Harré, Rom (1987), 'Persons and selves', in *Persons and Personality: A contemporary inquiry*, ed. A. Peacocke and G. Gillet (Oxford: Blackwell).
- Hume, David (1978), *A Treatise on Human Nature* (Oxford: Clarendon Press; original work published in 1739).
- James, William (1905), *The Principles of Psychology*, Vol. 1. (London: Macmillan; original work published in 1890).
- Jung, C.G. (1968), *The Archetypes and the Collective Unconscious*, 2nd. ed. (London: Routledge).
- Kenny, Anthony (1988). *The Self* (Milwaukee, WI: Marquette University Press).
- Lowe, E.J. (1996), *Subjects of Experience* (Cambridge: Cambridge University Press).
- McCall, Catherine (1990), *Concepts of Person: An analysis of concepts of person, self and human being* (Aldershot, UK: Avebury).
- Marx, Melvin and Hillix, W. (1973), *Systems Theories in Psychology* (New York: McGraw Hill).
- Mellor, D.H. (1991), *Matters of Metaphysics* (Cambridge: Cambridge University Press).
- Minsky, Marvin (1985), *The Society of Mind* (New York: Simon and Schuster).
- Nagel, Thomas (1986), *The View from Nowhere* (New York: Oxford University Press).
- Rée, Jonathan (1974), *Descartes* (London: Allen Lane).
- Rogers, Carl R. (1951), *Client-centered Therapy* (Boston, MA: Houghton).
- Solomon, Robert (1988), *Continental Philosophy Since 1750: The rise and fall of the self* (Oxford: Oxford University Press).
- Strawson, Galen (1997), 'The self', *Journal of Consciousness Studies* 4 (5–6), pp. 405–28.
- Toulmin, Stephen E. (1977), 'Self-knowledge and knowledge of the "self"', in *The Self: Psychological and philosophical perspectives*, ed. Theodore Mischel (Oxford: Blackwell).
- van Inwagen, Peter (1980), 'Philosophers and the words "human body"', in *Time and Cause*, ed. van Inwagen (Dordrecht: Reidel).

Paper received October 1997