

‘THE SELF’

Galen Strawson, * Jesus College, Oxford OX1 3DW, UK
Email: galen.strawson@philosophy.oxford.ac.uk

- I know that I exist; the question is, what is this ‘I’ that I know? (Descartes 1641)
- The soul, so far as we can conceive it, is nothing but a system or train of different perceptions. (Hume 1739)
- *What* was I before I came to self-consciousness? . . . I did not exist at all, for I was not an I. The I exists only insofar as it is conscious of itself. . . . *The self posits itself*, and by virtue of this mere self-assertion it exists. (Fichte 1794–5)
- The ‘Self’ . . . , when carefully examined, is found to consist mainly of . . . peculiar motions in the head or between the head and throat. (James 1890)
- The ego continuously constitutes itself as existing. (Husserl 1929)
- Any fixed categorization of the Self is a big goof. (Ginsberg 1963)
- The self which is reflexively referred to is synthesized in that very act of reflexive self-reference. (Nozick 1981)
- The self . . . is a mythical entity. . . . It is a philosophical muddle to allow the space which differentiates ‘my self’ from ‘myself’ to generate the illusion of a mysterious entity distinct from . . . the human being. (Kenny 1988)
- A self . . . is . . . an abstraction . . . , [a] Center of Narrative Gravity. (Dennett 1991)
- My body is an object all right, but my self jolly well is not! (Farrell 1996)¹

I: Introduction

The substantival phrase ‘the self’ is very unnatural in most speech contexts in most languages, and some conclude from this that it’s an illusion to think that there is such a thing as the self, an illusion that arises from nothing more than an improper use of language. This, however, is implausible. People are not that stupid. The problem of the self doesn’t arise from an unnatural use of language which arises from nowhere. On the contrary: use of a phrase like ‘the self’ arises from a prior and independent sense that there is such a thing as the self. The phrase may be unusual in ordinary speech; it may have no obvious direct translation in many languages. Nevertheless all languages have words which lend themselves naturally to playing the role that ‘the self’ plays in English, however murky that role may be. The phrase certainly means something to most people. It has a natural use in religious, philosophical, and psychological contexts, which are very natural contexts of discussion for human beings. I think there is a real philosophical problem about the existence and nature of the self, not just a relatively uninteresting problem about why we think there’s a problem. It is too quick to say that a ‘grammatical error . . . is the essence of the theory of the self’, or that ‘“the self” is a piece of philosopher’s nonsense consisting in a misunderstanding of the reflexive pronoun’ (Kenny, 1988, p. 4).

* I am grateful to Derek Parfit, Shaun Gallagher, Jonathan Shear, Keith Sutherland, and P.F. Strawson for their comments on a draft of this paper.

¹ Sources of quotations: Descartes (1985, vol. 2, p. 18); Hume (1978, p. 657); Fichte (1982, pp. 97–8); James (1950, vol. 1, p. 301); Husserl (1973, p. 66); Ginsberg (1963); Nozick (1981, p. 91); Kenny (1988, pp. 3–4); Dennett (1991, pp. 426–7); Farrell (1996, p. 519).

The first task is to get the problem into focus. I will recommend one approach, first in outline, then in slightly more detail. (I will model the problem of the self, rather than attempting to model the self.) I think the problem requires a straightforwardly metaphysical approach; but I also think that metaphysics must wait on phenomenology, in a sense I will explain. Most recent discussion of the problem by analytic philosophers has started from work in philosophical logic (in the large sense of the term).² This work may have a contribution to make, but a more phenomenological starting point is needed.

I will use the expression ‘the self’ freely — I am already doing so — but I don’t want to exclude in advance the view that there is no such thing as the self, and the expression will often function as a loose name for what one might equally well call ‘the self-phenomenon’, i.e. all those undoubtedly real phenomena that lead us to think and talk in terms of something called the self, whether or not there is such a thing.

II: The Problem of the Self

Many people believe in the self, conceived of as a distinct thing, although they are not clear what it is. Why do they believe in it? Because they have a distinct *sense* of, or experience as of, the self, and they take it that it is not delusory. This sense of the self is the source in experience of the philosophical problem of the self. So the first thing to do is to track the problem to this source in order to get a better idea of what it is. The first question to ask is the *phenomenological question*:

What is the nature of the sense of the self?

But this, in the first instance, is best taken as a question explicitly about human beings: as the *local phenomenological question*

(1) What is the nature of the human sense of the self?

Whatever the answer to (1) is, it raises the *general phenomenological question*

(2) Are there other possibilities, when it comes to a sense of the self? (Can we describe the minimal case of genuine possession of a sense of the self?)

The answers to (1) and (2) raise the *conditions question*

(3) What are the grounds or preconditions of possession of a sense of the self?

and this question raises a battery of subsidiary questions. But progress is being made, at least potentially. For, if one can produce satisfactory answers to (1), (2) and (3), one will be in a good position to raise and answer the *factual question*, the fundamental and straightforwardly metaphysical question

(4) Is there (could there be) such a thing as the self?

I think one has to answer (1) and (2), and probably (3), in order to answer (4) properly.

III: The Local Question; Cognitive Phenomenology

I will now go through the plan in more detail, and sketch how I think some of the answers should go. The first question is the local phenomenological question: What is the nature of the ordinary human sense of the self? This raises a prior question: Can

² See, for example, the essays collected in Cassam (1994).

one generalize about the human sense of the self? I think the answer is Yes: the aspects of the sense of the self that are of principal concern, when it comes to the philosophical problem of the self, are very basic. They are situated below any level of plausible cultural variation.³ They are conceptual rather than affective: it is the *cognitive phenomenology* of the sense of the self that is fundamentally in question, i.e. the conceptual structure of the sense of the self, the structure of the sense of the self considered (as far as possible) independently of any emotional aspects that it may have. The cognitive phenomenology of the self is bound up with the affective phenomenology of the self in complicated ways, but emotional or affective aspects of the sense of the self will be of concern (e.g. in section VIII) only in so far as emotions shape or weight conceptions.

What, then, is the ordinary, human sense of the self, in so far as we can generalize about it? I propose that it is (at least) the sense that people have of themselves as being, specifically, a mental presence; a mental someone; a single mental thing that is a conscious subject of experience, that has a certain character or personality, and that is in some sense distinct from all its particular experiences, thoughts, and so on, and indeed from all other things. It is crucial that it is thought of as a distinctively mental phenomenon, and I will usually speak of the ‘mental self’ from now on (the qualifier ‘mental’ may be understood wherever omitted).

Is the sense of the mental self, as so far described, really something ordinary? I believe so. It comes to every normal human being, in some form, in childhood.⁴ The early realization of the fact that one’s thoughts are unobservable by others, the experience of the profound sense in which one is alone in one’s head — these are among the very deepest facts about the character of human life, and found the sense of the mental self. It is perhaps most often vivid when one is alone and thinking, but it can be equally vivid in a room full of people. It connects with a feeling that nearly everyone has had intensely at some time — the feeling that one’s body is just a vehicle or vessel for the mental thing that is what one really or most essentially is. I believe that the primary or fundamental way in which we conceive of ourselves is as a distinct mental thing — sex addicts, athletes, and supermodels included. Analytic philosophers may find it hard to see — or remember — this, given their training, and they risk losing sight of the point in derision.

This is not to deny that we also naturally conceive of ourselves as mental-and-non-mental things, human beings considered as a whole. We do. Nor is it to claim that the sense of the mental self always incorporates some sort of belief in an immaterial soul, or in life after bodily death. It doesn’t. Philosophical materialists who believe, as I do, that we are wholly physical beings, and that the theory of evolution by natural selection is true, and that animal consciousness of the sort with which we are familiar evolved by purely physical natural processes on a planet where no such consciousness previously existed, have this sense of the mental self as strongly as anyone else.

In more detail: I propose that the mental self is ordinarily conceived or experienced as:

³ Work in evolutionary psychology suggests that doubts about the possibility of generalization that derive from considerations of cultural difference can be easily dealt with. See e.g. Barkow *et al.* (1992).

⁴ It certainly does not require the special kind of experience recorded by Nagel (1986, pp. 54–7) or Richard Hughes (1929, ch. 6), for this is by no means universal.

- (1) a *thing*, in some robust sense
- (2) a *mental* thing, in some sense
- (3,4) a *single* thing that is single both *synchronically* considered and *diachronically* considered
- (5) *ontically distinct* from all other things
- (6) a *subject of experience*, a conscious feeler and thinker
- (7) an *agent*
- (8) a thing that has a certain character or *personality*

This is an intentionally strong proposal, and it may be thought to be too strong in various ways. Most of (1)–(8) can be contested, and the list may well contain redundancy, but it provides a framework for discussion. There are various entailment relations between the eight elements that need to be exposed; (1) – (6) are closely linked. (1) also raises the general question ‘What is a thing?’ — a question that will be important when the fundamental factual question (‘Is there such a thing as the self?’) is considered.

I don’t think the list omits anything essential to a genuine sense of the mental self, even if it includes some things that are not essential. I will assume that this is true for the purposes of this paper: a primitive framework can show the structure of a problem even if it is not complete. It can be the best way to proceed even if the problem resists regimentation in terms of necessary and sufficient conditions. If an omission were identified, it could simply be added in to the existing framework.

(2) is the only one of the eight properties that is not attributed as naturally to the embodied human being as to the putative mental self, and it may be suggested that the sense of the mental self is just a delusory projection from the experience of embodiment. Perhaps the so-called self is just the human being incompletely grasped and illegitimately spiritualized. This is a popular view, but I am not yet in a position to assess it.⁵ Some argue from the fact that use of the word ‘I’ to refer to the supposed mental self does not ordinarily stand out as distinct from use of the word ‘I’ to refer to the human being considered as a whole to the conclusion that we have no good reason to distinguish them. To this it may be replied that appeal to facts about public language use is often irrelevant when considering facts about meaning and reference, and is spectacularly inappropriate in the case of the problem of the self.⁶

IV: Phenomenology and Metaphysics

Equipped with an answer to the local question, one can go on to raise the general question: ‘Are there other possibilities, so far as a sense of the mental self (or SMS) is concerned?’ Given the assumption that the list of eight properties doesn’t omit anything essential to a genuine sense of the self, this amounts to the question whether one can dispense with any of (1)–(8) while still having something that qualifies as a genuine SMS. It enquires, among other things, after the minimal case of a SMS. The answer is partly a matter of terminological decision, but for the most part not.

⁵ For older versions of the view, see e.g. James (1950, ch 10). See also Bermúdez *et al.* (1995).

⁶ This point is developed in Strawson (forthcoming).

How might the answer go? I don't yet know, but if I had to commit myself it would be as follows: (4) and (8) are not necessary to a sense of the mental self, even in the human case (see sections VIII and IX). (6) is secure, but a serious doubt can be raised about (7). (2) and (5) need careful qualification if they are to survive. (1) and (3) can be challenged but effectively defended.

Objection: 'Surely the phenomenological investigation loses something crucial at this point? It is no longer rooted in the human case, so it is no longer independent of specifically philosophical theories about what selves actually are or can be: such theories are bound to be part of what governs our judgements about whether some thinned down SMS can count as a genuine SMS, once we go beyond the human case.'

I believe that a detailed attempt to answer the general phenomenological question will show that this is not so: our basic judgements about whether anything less than (1)–(8) can count as a genuine SMS can remain comfortably independent, in any respect that matters, of metaphysical philosophical theorizing about the nature of the self. In fact I think they can be sufficiently supported by reference to unusual human cases.

So much for the claim that phenomenology is substantially independent of metaphysics. What about the other way round? Here I think there is a fundamental dependence: metaphysical investigation of the nature of the self is subordinate to phenomenological investigation of the sense of the self. There is a strong phenomenological constraint on any acceptable answer to the metaphysical question which can be expressed by saying that the factual question 'Is there such a thing as the mental self?' is equivalent to the question 'Is any (genuine) sense of the self an accurate representation of anything that exists?'⁷

This equivalence claim can be split in two:

(E1) If there is such a thing as the self, then some SMS is an accurate representation of something that exists,

(E2) If some SMS is an accurate representation of something that exists, then there is such a thing as the self.

(E1) and (E2) may seem trivial, but both may be challenged. The first as follows:

(C1) There is really no very good reason to think that if the self exists, then there is some SMS that is an accurate (if partial) representation of its nature. Perhaps the mental self, as it is in itself, is ineffable, quite unlike any experience of it.

(C1) is Kantian in spirit. The second rejection is a response made when some particular SMS has been presented:

(C2) This SMS you have outlined is indeed an accurate representation of something that exists, but the thing of which it is an accurate representation does not qualify for the title 'the mental self' because it does not have feature F (e.g. it is not an immaterial, \pm immortal, \pm whatever, substance).

The force of (E1) and (E2) consists precisely in the fact that they reject proposals like (C1) and (C2). In this way they impose a substantial constraint on metaphysical theorizing about the self. According to (E1), nothing can count as a mental self unless it possesses all the properties attributed to the self by some genuine SMS, whatever other properties it may possess. It rules out metaphysical claims about the self that

⁷ I take it that a representation R of a thing X is accurate if (and only if) X really has the properties R represents it as having. R need not be complete to be accurate.

fail to respect limits on the concept of the self revealed by the phenomenological investigation. It states a necessary condition on qualifying for the title of self. (E2), by contrast, states that nothing can fail to count as a mental self if it possesses all the properties that feature in some SMS, whatever other properties it may possess or lack. It states a sufficient condition on qualifying for the title of self — it lays it down that there is no further test to pass.

To make the equivalence claim, then, is to say that one must have well-developed answers to phenomenological questions about the experience of the self before one can begin to answer metaphysical questions about the self. The equivalence claim excludes two forms of metaphysical excess — extravagance and miserliness. Extravagance is blocked by showing that we cannot answer the question ‘Is there such a thing as the self?’ by saying ‘Yes there is (or may be), but we have (or may have) no understanding of its ultimate nature’. Miserliness is blocked by showing that we cannot answer by saying ‘Well, there is *something* of which the sense of the self is an accurate representation, but it does not follow that there is any such thing as the self.’

If the answers to the phenomenological questions go well, we should be left with a pretty good idea of what we are asking when we ask the factual, metaphysical question ‘Is there such a thing as the self?’ Any metaphysical speculations that are not properly subordinate to phenomenology can be cheerfully ‘commit[ted]. . . to the flames’ (Hume, 1975, p. 165).⁸

V: Materialism

In sections VI–IX I will give examples of more detailed work within this scheme. Before that I must give a brief account of the sense in which I am a materialist.

Materialists believe that every thing and event in the universe is a wholly physical phenomenon. If they are even remotely realistic in their materialism they admit that conscious experience is part of reality. It follows that they must grant that conscious experience is a wholly physical phenomenon. They must grant that it is wholly physical specifically in its mental, experiential properties. (They must grant that the qualitative character of the taste of bread, considered just as such and independently of anything else that exists, is as much a physical phenomenon as the phenomenon of an electric current flowing in a wire.)

It follows that materialists express themselves very badly when they talk about the mental and the physical as if they were opposed categories. For on their own view, this is exactly like saying that cows and animals are opposed categories — for all mental phenomena, including conscious-experience phenomena *considered specifically as such*, just are physical phenomena, according to them; just as all cows are animals.

So what are materialists doing when they talk as if the mental and the physical were different things? What they presumably mean to do is to distinguish, within the realm of the physical, which is the only realm there is, according to them, between the mental and the non-mental, and, more specifically, between the experiential and the non-experiential; to distinguish, that is, between (A) mental (or experiential) aspects of the physical, and (B) non-mental (or non-experiential) aspects of the physical.⁹

⁸ I should say that I’m rejecting, and not claiming to refute, more unbridled approaches to the metaphysics of the self.

⁹ I need to make the distinction between mental and experiential phenomena, because although all experiential phenomena are mental phenomena, not all mental phenomena are experiential phenom-

This is the difference that is really in question when it comes to the ‘mind–body’ problem, and materialists who persist in talking in terms of the difference between the mental and the physical perpetuate the terms of the dualism they reject in a way that is inconsistent with their own view.¹⁰

Let me rephrase this. When I say that the mental and the experiential are wholly physical, I mean something completely different from what some materialists have apparently meant by saying things like ‘experience is really just neurons firing’. I don’t mean that all that is really going on, in the case of conscious experience, is something that can be discerned and described by current physics, or by any non-revolutionary extension of current physics. Such a view amounts to some kind of radical eliminativism, and is certainly false. My claim is quite different. It is that the experiential considered specifically as such — the portion of reality we have to do with when we consider experiences specifically and solely in respect of the experiential character they have for those who have them as they have them — that ‘just is’ physical. No one who disagrees with this claim is a serious and realistic materialist.¹¹

A further comment is needed. As remarked, thoroughgoing materialists hold that all mental phenomena, including all experiential phenomena, are entirely physical phenomena. But triviality threatens when things are put this way. For now even absolute idealism (in one version, the view that only experiential phenomena exist) can claim to be a materialist position.

The trivializing possibility can be excluded by ruling that anything deserving the name ‘materialism’ must hold that there are non-mental and non-experiential phenomena as well as mental or experiential phenomena. But one can plausibly go further, and take materialism to incorporate what one might call ‘the principle of the necessary involvement of the mental with the non-mental’. Most realistic materialists take it that the existence of each particular mental or experiential phenomenon involves the existence of some particular non-mental, non-experiential phenomenon. More strongly expressed: each particular mental or experiential phenomenon has, essentially, in addition to its mental or experiential character or mode of being, a non-mental character or mode of being. One might call this ‘mental-and-non-mental’ materialism. When I talk of materialism in what follows, I will take it to involve this view.

According to materialism, then, every thing or event has non-mental, non-experiential being, whether or not it also has mental or experiential being. More needs to be said (given that we have knowledge of central aspects of the fundamental reality of the mental just in having experience in the way we do, we need to ask whether it is possible to give some basic positive characterization of the non-mental, perhaps in terms of properties like time, length, position, mass, electric charge, spin, ‘colour’ and ‘flavour’ in the quantum theory sense). But this is enough to make it clear that the present question about whether the self exists in the human case is not a question about whether we might possibly be ‘Cartesian egos’ or immaterial substances. It is

ena: according to ordinary usage, beliefs, likes and dislikes, and so on are mental phenomena, though they have no experiential character.

¹⁰ There is tremendous resistance to abandoning the old mental/physical terminology in favour of the mental/non-mental, experiential/non-experiential terminology, even though the alternative is very clear and is exactly what is required. Cf. Searle (1992, p. 54); also A. Campbell (1994).

¹¹ Hurlburt *et al.* discuss a superficially ‘zombie’-like subject who has ‘no reportable inner experience’ (1994, pp. 391–2), but it becomes clear he does have experience in the current sense.

the question whether the mental self exists given that we are ordinarily embodied, entirely physical living human beings.

VI: Singularity

I have sketched how I think answers to the phenomenological questions should go, described the constraint that phenomenology places on metaphysics, and characterized the sense in which I am a materialist. I will now give samples of more detailed work on the phenomenological questions.

The proposal for consideration is that the mental self is conceived or experienced as (1) a *thing*, (2) a *mental* thing, a *single* thing that is single both (3) *synchronically* considered and (4) *diachronically* considered, (5) a thing that is *ontically distinct* from all other things, (6) a *subject of experience* and (7) an *agent* that has (8) a certain *personality*. In this section I will discuss (3) and (4) in the framework of the *local* phenomenological question, after very brief comments on (1) and (2). In sections VII–IX I will discuss of (4) and (8) in the framework of the *general* phenomenological question. In section IX I will say something about (5).

Thinghood and mentality

What about the claim (1) that the self is conceived of as a thing? In a way, this is the least clear of the eight claims, but the general idea is this: the self isn’t thought of as merely a state or property of something else, or as an event, or process, or series of events. So, in a sense, there is nothing else for it to seem to be, other than a thing. It’s not thought of as being a thing in the way that a stone or a cat is — it’s not thought of as a sort of ethereal concrete object. But it is thought of as a thing of some kind. In particular, it is thought of as something that has the causal character of a thing; something that can undergo things and do things. Bishop Berkeley’s characterization of the self as a ‘thinking . . . principle’ seems as good as any (1975, p. 185). A principle, in this old use, manages to sound like a thing of some sort without sounding anything like a table or a chair.

The second claim, (2), that the self is thought of as something mental, is also unclear. Very briefly, the idea is something like this: when the self is thought of as a thing, its claim to thinghood is taken to be sufficiently grounded in its mental nature alone. It may also have a non-mental nature, as materialists suppose, but its counting as a thing is not thought to depend on its counting as a thing considered in its non-mental nature: the self is the *mental* self. (It’s true and important that many people naturally think of themselves as possessing both mental and non-mental properties, but this doesn’t affect the truth of (2).)

Singularity

Clearly, to think of the self as a thing is already to think of it as single in some way — as *a* thing. But in what way? I have three main claims in mind.

First: in so far as the mental self is thought of as single, it is not thought of as having singularity only in the sense in which a group of things can be said to be a single group. Rather it is thought of as single in the way in which a single marble (e.g.) is single when compared with a single pile of marbles. Developing the Lockean point just made about the fundamental causal component in our idea of a thing, one might say that the mental self is conceived of as something that has the kind of strong unity

of internal causal connectedness that a single marble has, as compared with the much weaker unity of internal causal connectedness found in a pile of marbles.¹²

Second: the mental self's property of singleness is thought of as sufficiently and essentially grounded in its mental nature alone. This closely parallels the idea that the self's claim to thinghood is thought of as sufficiently grounded in its mental nature alone, and the same moves are appropriate. We may suppose that the mental self has non-mental being (the brain-as-revealed-to-physics, say) as well as mental being, and it may be *believed* to have non-mental being. The fact remains that it is thought of as having singleness in a way that is independent of its having singleness when considered in its non-mental nature.

One may express this by saying that its *principle of unity* is taken to be mental. What does 'principle of unity' mean? Well, it is arguable that everything that is conceived of as a single thing or object — electron, atom, neuron, sofa, nation-state — is conceived of as a single thing relative to some principle of unity according to which it *counts* as a single thing. An atom counts as a single thing relative to one principle of unity, and it counts as many things relative to other principles of unity — those which discern subatomic particles. Many associate this point with the view that there are no ultimate facts of the matter about which phenomena are things or objects and which are not; they hold that all principles of objectual unity, as one might call them, are ultimately subjective in character. But this is a further claim. In itself, the claim that everything that is taken to be a single object is so taken relative to some principle of objectual unity is compatible with the view that there are objective principles of objectual unity given which there are right answers to questions about which things are genuinely single objects.

Let me try to put the point about the self in another way: we may suppose that the mental self (the self-phenomenon) has non-mental being as well as mental being, and it may even be widely believed that this is so (few give the matter much thought). The fact remains that it is thought of as having singleness in its mental being in a way that is independent of any singleness that it may have in its non-mental being. In this sense it is taken to be single just as something mental.¹³ I will illustrate this idea after introducing the third main point about singleness.

This is that the mental self is standardly thought to be single in the two ways just characterized both when it is considered (3) synchronically, or as a thing existing at a given time, and when it is considered (4) diachronically, i.e. as a thing that persists through time.

In what follows, I will stretch the meaning of 'synchronic' slightly, and take it to apply to any consideration of the mental self (or self-phenomenon) that is a consideration of it during an experientially unitary or unbroken or hiatus-free period of thought or experience. The notion of a hiatus-free period of thought or experience is important for my purposes, and needs further description (see section IX). For the

¹² Cf. Campbell (1995). A marble, of course, is made of atoms, and is a collection of things from the point of view of an atom. An atom is a collection of things from the point of view of an electron, and perhaps the series continues. This is the point of the comparative formula 'single in the way in which a marble (e.g.) is single when compared with a pile of marbles'.

¹³ Compare 'X is taken to be single just qua something physical (i.e. non-mental)'. The thought that this expresses is not problematic for ordinary thought, and the thought expressed by 'X is taken to be single just qua something mental' is no more problematic.

moment let me simply assert that in the normal course of events truly hiatus-free periods of thought or experience are invariably brief in human beings: a few seconds at the most, a fraction of a second at the least. Our eyes are constantly engaged in saccadic jumps, and reflection reveals the respect in which our minds function in an analogous — if more perceptible — way. (Research by Pöppel and others provides ‘clear evidence that . . . the experienced Now is not a point, but is extended, . . . that the [human] conscious Now is — language and culture independent — of the duration of approximately 3 seconds’, and although this proves nothing about the existence of hiatuses, or about the nature of the self, it is undeniably suggestive.)¹⁴

‘Diachronic’ complements ‘synchronic’ and applies to consideration of the mental self (or self-phenomenon) during any period of conscious thought or experience that includes a break or hiatus. Such periods may range from a fraction of a second to a lifetime.

Now reconsider the second claim — that the mental self is taken to be single just as something mental. This has a synchronic and and diachronic aspect. I will begin with the former. Suppose that someone fully convinces you (perhaps by hypnosis) that your current mental life with all its familiar characteristics, which incorporates your current sense of the single mental self, depends on the activity of three spatially separated brains in three different bodies. Will this immediately annihilate your natural sense of your mental singleness? Surely not. Your thought is likely to be ‘Wow, I have got three brains—I, the single thing or person that I am’ (Kant (1996, A353–4) makes a related point). Your sense of the mental self is overwhelmingly likely to continue unchanged. It doesn’t depend on your believing that you have a single brain or body. Suppose that you find out that there are three separate brains in your single body, collaborating to produce your experience. Again this will not override the experience of mental singleness.

It may be objected that in the case imagined you still have experience as of inhabiting a single body. This is true, given that you are an ordinary human being. But one can equally well imagine a three-bodied creature that naturally experiences itself as three-bodied, and as receiving information (perhaps via different sense modalities) from all three bodies, while still having a strong sense of the single mental self, and thinking of itself as ‘I’. Here the experience of three-bodiedness is likely to make the sense of the singleness of the mental self particularly vivid. It is true that ordinary human experience of oneself as mentally single is deeply shaped by experience of having a single body, but it hardly follows that any possible experience of oneself as mentally single depends essentially on such experience.¹⁵

That is the sense of synchronic singleness I have in mind. Now for the diachronic case. Suppose one experiences one’s mental life as something that has strong diachronic singleness or unity (some do more than others). And suppose that one is then convinced — that it depends for its existence on the successive existence of a series of numerically distinct brains or neuronal entities. Will this annihilate one’s sense of the mental self as a single thing persisting through time? It would be extraordinary if it did: for, by hypothesis, everything else is the same, experientially, as it was before

¹⁴ Ruhnu (1995, p. 168); Pöppel (1978). Citing this research in his essay *The Dimension of the Present Moment*, the Czech immunologist and poet Miroslav Holub writes that ‘in this sense our ego lasts three seconds’ (1990, p. 6).

¹⁵ This is the kind of issue that arises when one asks (3), the ‘conditions’ question.

one made this discovery. This suggests that confrontation with the fact of one's non-mental multiplicity will have no more force to undermine one's sense of the singleness of the mental self in the diachronic case than in the synchronic case.

There is a famous footnote in Kant's discussion of the Third Paralogism (Kant, 1996, A363–4):

An elastic ball which strikes another similar ball in a straight line communicates to the latter its entire motion, and therefore its entire state (if we take account only of positions in space). If, in analogy with such bodies, we postulate substances such that the one communicates representations to the other together with consciousness of them, we can conceive a whole series of substances of which the first transmits its state to the second, the second its own state with that of the preceding substance to the third, and [so on]. The last substance would then be conscious of all the states of the previously changed substances as being its own states, because they would have been transferred to it together with consciousness of them.

Kant's aim is to argue that no experience of the diachronic singleness of the mental self can possibly establish that the mental self or 'I' is in fact a diachronically single substance. My different, compatible claim is that even if one came to believe that the existence of the mental self did *not* involve the existence of a diachronically single substance, there is no reason to suppose that this would undermine one's experience of the mental self as so single.

To summarize: even if one takes it for granted that the mental self (or self-phenomenon) has a non-mental nature or being, one's experience of the mental self as single is independent of any belief that it is single — either synchronically or diachronically — in its non-mental nature or being. This, then, illustrates the respect in which the singularity of the mental self is conceived of as being essentially grounded in its mental nature alone.

It's also true — to diverge from merely phenomenological concerns — that thoughts that occur in a single body or brain (or substance of some other sort) may fail to seem anything like the series of thoughts of a single self or thinker, both when considered 'from the inside' (i.e. from the point of view of the thinker of any given one of the thoughts in question) and when considered from the outside (i.e. by someone who is not the thinker of any of the thoughts, but who has access to the contents of the thoughts, as in a novel). Consider the diachronic case first: imagine that a series of self-conscious thoughts or 'I-thoughts' occurs in the same brain, one at a time, while none of them ever involves any awareness of any thought earlier (or indeed later) than itself, and while no two of them ever stand in any of the relations (of content, temperamental coherence, etc.) in which temporally close pairs of thoughts so often stand when they are the thoughts of a being that we naturally think of as a single thinker.

In this case, it may be said that we lack any mentally grounded reason for saying that there is a single thinker. Some may want to say that there is nevertheless a single thinker, simply because a single brain is the locus of all the thoughts. But why should the fact of non-mental diachronic singleness decisively overrule the natural judgement that there is no plausible candidate for a diachronically single mental self in this case? The fact of non-mental multiplicity in the three-bodies case had no power to defeat the natural judgement of mental singleness. Why should the fact of non-mental

singleness in this case defeat the natural judgement of mental multiplicity (lack of mental singularity)?¹⁶

Now consider the synchronic case: imagine that a single brain is the site of experiential phenomena that are just like the experiential phenomena taking place simultaneously in the brains of three different people (the first thinking exclusively about Vienna, the second exclusively about menhirs, the third exclusively about DNA). Here it is natural to judge that there are three subjects of experience. If one counts the whole brain non-mentally considered as the non-mental being of each of the three apparently distinct thought-thinking selves, then one has multiplicity of selves in spite of non-mental singleness.

The judgement that there are three subjects of experience may seem natural in this case, but it can be cogently challenged. It is very difficult to draw firm conclusions about the number of subjects of experience associated with a single brain from facts about the contents of the experiences associated with that brain. As far as the synchronic case is concerned: it may be a fact about human beings that they can only genuinely entertain one conscious thought at a time, but it does not seem to be an *a priori* truth about conscious thinking in general. As far as the diachronic case is concerned: it is not clear that there is any lower bound on the connectedness of the successive thoughts and experiences of a single subject of experience, any point at which we can confidently say, ‘These experiences are too unconnected and disordered to count as the experiences of a single subject of experience.’¹⁷

Multiplicity?

So far I have taken it for granted that human beings standardly have some sense of the singleness of the mental self. But some may claim to experience the mental self as fragmentary or multiple, and most of us have had experiences that give us — so we feel — some understanding of what they mean.

It seems, however, that the experience of multiplicity can at most affect (4), the sense of the mental self as diachronically single (recall that a sense of the mental self as diachronically single may well be concerned with short periods of time; when I want to consider longer periods of time — weeks, months, years, lifetimes — I will talk about ‘long-term’ continuity). It cannot affect (3), the sense of the mental self as synchronically single (single during any one ‘hiatus-free’ period of thought or experience). Why not? Because any candidate for being an experience of the mental self as synchronically multiple at the present moment will have to be an episode of explicitly self-conscious thought, and there is a crucial (trivial) respect in which no such episode could be experience of the mental self as synchronically multiple. Explicitly self-conscious thought need not always involve some explicit sense of the mental self as something present and involved, even when it has the form ‘*I f*’, or ‘*I am F*’ (‘I forgot the key’, ‘I’m late for my exam’). But whenever it does — and it must if there is to be anything that is a candidate for being an *experience* of the mental self as synchronically multiple at the present moment — there is a fundamental respect in which the mental self must be experienced as single, for the space of that thought at least.

¹⁶ The phenomena of dissociative identity disorder may also support the idea that non-mental singleness is compatible with a multiplicity of mental selves, but the present example is much more extreme.

¹⁷ See Snowdon (forthcoming); also Van Inwagen (1990, section 16, pp. 196–202).

This may seem obvious, but it can be disputed. It may be said that even experience of the mental self synchronically considered can seem to be experience of something shattered and multiple ('My name is legion', *Mark 5.9*). There seem to be forms of human experience that invite such a description. One may be under stress and subject to rapidly changing moods. One may feel oneself pulled in different directions by opposed desires. Human thought-processes can become extraordinarily rapid and tumultuous. But what exactly is being claimed, when it is said that the self may be experienced as synchronically multiple? There seem to be two main possibilities: either the experience is that there are many selves present, or it is (just) that the self is complex in a certain radical way. But in the second case, the experience of radical complexity that is claimed to justify the description 'synchronically multiple' clearly depends on a prior sense of the mental self as synchronically single: in this case 'multiple' is a characterization that is applied to something that must have already presented as single in order for the characterization to be applied at all. What about the first case, in which the experience is that there are many selves present? Well, we may ask who has the experience that there are many selves present. To face the question is to realize that any explicitly self-conscious experience has to present as experience from one single mental point of view. (The word 'mental' is not redundant here, for the three-bodied person that has sensory experience of being three-bodied may have three sensory points of view while still having only one mental 'point of view'.) If so, the experience that there are many selves present is necessarily experience from some single point of view. Even if a single brain is the site of many experiences that there are many selves present, each such experience is necessarily experience from a single point of view. This is the trivial aspect of the claim that experience of the mental self as synchronically multiple is not really possible.¹⁸

It may be added that when one's mind races and tumbles, it is natural to experience oneself as a largely helpless spectator of the pandemonium. To this extent, experience of chaotic dispartateness of contents reinforces a sense of singleness rather than diminishing it. Nor can one experience conflict of desire unless one experiences both desires as one's own.

VII: Personality

So much for a consideration of (3) and (4) — synchronic and diachronic singleness — in the framework of the local phenomenological question (What is the human sense of the self?) I will now consider (4) and (8) — diachronic singularity and personality — in the framework of the general phenomenological question (What senses of the self are possible?) I will begin with personality, and, like William James, I will sometimes talk 'in the first person, leaving my description to be accepted by those to whose introspection it may commend itself as true, and confessing my inability to meet the demands of others, if others there be' (1950, vol. 1, p. 299).

¹⁸ I take it that this conclusion is compatible with the possibility of Husserlian '*splitting of the I*' in transcendental-phenomenological reflection (Husserl, 1973, p. 35), and also with a thought-experiment of Parfit's in which he imagines being able to 'divide his mind' in order to do two separate calculations in two separate streams of consciousness, and then reunite it. He considers his attitude to the process after several divisions and reunions: 'in each of my two streams of consciousness I would believe that I was now, in my other stream, having thoughts and sensations of which, in this stream, I was unaware' (Parfit, 1984, pp. 246–8).

It seems plain that (8) is not a necessary component of any possible sense of the mental self — that experience of the self does not necessarily involve experience of it as something that has a personality. Most people have at some time, and however temporarily, experienced themselves as a kind of bare locus of consciousness — not just as detached, but as void of personality, stripped of particularity of character, a mere (cognitive) point of view. Some have experienced it for long periods of time. It may be the result of exhaustion or solitude, abstract thought or a hot bath. It is also a common feature of severe depression, in which one may experience 'depersonalization'. This is a very accurate term, in my experience and in that of others I have talked to.

Sustained experience of depersonalization is classified as psychotic relative to the normal human condition, but it is of course experientially real, and one can imagine human beings getting stuck in this condition; some do. Equally, one can imagine aliens for whom it is the normal condition. Such an alien may still have a clear sense of the self as a specifically mental thing. It may still have an unimpaired sense of itself as a locus of consciousness, just as we ordinarily do—not only when we suffer depersonalization, but also in everyday life.¹⁹

A very strong form of what may be lost in depersonalization is recorded by Gerard Manley Hopkins, who talks of considering

my self-being, my consciousness and feeling of myself, that taste of myself, of *I* and *me* above and in all things, which is more distinctive than the taste of ale or alum, more distinctive than the smell of walnutleaf or camphor, and is incommunicable by any means to another man Nothing else in nature comes near this unspeakable stress of pitch, distinctiveness, and selving, this selfbeing of my own.²⁰

My enquiries suggest that while some people feel they know exactly what Hopkins means, most find this deeply bewildering: for them, their personality is something that is unnoticed, and in effect undetectable, in the present moment. It's what they look through, or where they look from; not something they look at; a global and invisible condition of their life, like air, not an object of experience. Dramatic differences like these back up the view that we need a phenomenology of the sense of the self before we try to answer the factual question about whether or not there is such a thing.

VIII: The Self In Time; Effects of Character

So much, briefly, for (8). Must any sense of the mental self involve experience of the self as (4) something that has long-term diachronic continuity as a single thing? I think not. The sense of the single mental self may be vivid and complete, at any given time, even if it has to do only with the present, brief, hiatus-free stretch of consciousness, at any given time. Nor do I think that this is just some alien or logical possibility, though it is also that. It lies within the range of human experience. One can be fully

¹⁹ A friend who recently experienced depersonalization found that the thought 'I don't exist' kept occurring to him. It seemed to him that this exactly expressed his experience of himself, although he was aware of the force of Descartes' 'I think, therefore I am', and knew, of course, that there had to be a locus of consciousness where the thought 'I don't exist' occurred. (The case of Meursault is also worth considering, in Camus' book *The Outsider*. So too is his remarkable description of his mother in *The First Man*. See Camus, 1982; 1995.)

²⁰ Hopkins (1959, p. 123); quoted in Glover (1988, p. 59).

aware of the fact that one has long-term continuity as a *living human being* without *ipso facto* having any significant sense of the *mental self* or *subject of experience* as something that has long-term continuity. One can have a vivid sense of oneself as a mental self, and a strong natural tendency to think that that is what one most fundamentally is, while having little or no interest in or commitment to the idea that the I who is now thinking has any past or future.

Human beings differ deeply in a number of ways that affect their experience of the mental self as diachronically continuous. Some people have an excellent 'personal' memory (i.e. memory of their own past life) and an unusual capacity for vivid recollection. Others have a very poor personal memory. And it may not be simply poor. It may also be highly quiescent, and almost never intrude spontaneously into their current thought. These deep differences of memory are matched by equal differences in the force with which people imagine, anticipate, or form intentions about the future.

These differences interact with others. Some people live deeply in narrative mode: they experience their lives in terms of something that has shape and story, narrative trajectory. Some of them are self-narrators in a stronger sense: they regularly rehearse and revise their interpretations of their lives. Some people, again, are great planners, and knit up their lives with long-term projects.

Others are quite different. They have no early ambition, no later sense of vocation, no interest in climbing a career ladder, no tendency to see their life in narrative terms or as constituting a story or a development. Some merely go from one thing to another. They live life in a picaresque or episodic fashion. Some people make few plans and are little concerned with the future. Some live intensely in the present, some are simply aimless.

Many things can encourage or obstruct a sense of the mental self as something that has long-term diachronic continuity. Some people are very consistent in personality or character, whether or not they know it. And this form of steadiness may in some cases strongly underwrite experience of the mental self's continuity. Others are consistent only in their inconsistency, and may for that reason feel themselves to be continually puzzling, and piecemeal. Some go through life as if stunned.

Neither inconsistency nor poor memory is necessary for the episodic experience of life. John Updike writes 'I have the persistent sensation, in my life and art, that I am just beginning' (1989, p. 239). These are the words of a man who has an extremely powerful personal memory and a highly consistent character. I have the same persistent sensation, and learn from Updike that it is nothing essentially to do with my extremely poor personal memory. I believe that it is an accurate description of how things are for many people, when it comes to that sense of oneself as a mental self that is — whether or not it is acknowledged — central to most people's self-conception.

I'm somewhere down the episodic end of the spectrum. I have no sense of my life as a narrative with form, or indeed as a narrative without form. I have little interest in my own past and little concern for the future. My poor personal memory rarely impinges on my present consciousness. Even when I am interested in my past, I'm not interested in it specifically in so far as it is mine. I'm perfectly well aware that it is mine, in so far as I am a human being considered as a whole, but I do not really think of it as mine at all, in so far as 'mine' picks out me as I am now. For me as I am

now, the interest (emotional or otherwise) of my personal memories lies in their experiential content considered independently of the fact that what is remembered happened *to me* — i.e. to the me that is now remembering.²¹ They’re certainly distinctive in their ‘from-the-inside’ character, but this in itself doesn’t mark them as mine in any emotionally significant sense. The one striking exception to this, in my case, used to be — but no longer is — memory of recent embarrassment.

I make plans for the future. To that extent I think of myself perfectly adequately as something that has long-term continuity. But I experience this way of thinking of myself as utterly remote and theoretical, given the most central or fundamental way in which I think of myself, which is as a mental self or someone. Using ‘Me*’ to express this fundamental way in which I think of myself — or to denote me thinking of myself in this way, looking out on things from this perspective — I can accurately express my experience by saying that I do not think of Me* as being something in the future. It is also accurate to shift the ‘not’, and say, more strongly, that what I think of as being in the future is not Me*.

As I write these words, the thought that I have to give a lecture before a large audience in two months’ time causes me some worry, which has familiar physiological manifestations. I feel the anxiety naturally and directly as pertaining to me even though I have no sense that it will be Me* that will be giving the lecture. Indeed it seems completely false to say that it will be Me*. And this is how it feels, not something I believe for theoretical reasons. So why do I feel any anxiety now? I believe that susceptibility to this sort of anticipatory anxiety is innate and ‘hard-wired’, a manifestation of the instinct for self-preservation: my practical concern for my future, which I believe to be within the normal human range, is biologically grounded and autonomous in such a way that it persists as something immediately felt even though it is not supported by any emotionally backed sense on the part of Me* now that Me* will be there in the future. (Not even half an hour away—and certainly not tomorrow.) In so far as I have any sense of Me* (rather than the living human being that I am) as something with a history and future, it seems that this sense is a wispy, short-range product of, and in no way a ground of, my innate predisposition to physiological impulses that develop into experience of anxiety or regret. It dislimns when scrutinized, and it is more accurate to say that it does not exist.

Now for an exception. You might expect me to say that when I think of my death at some unspecified future time, I think that it is not Me* who is going to die, or at least that I do not think that it is Me*. But I do think that it is Me* that is going to die, and I feel fear of death. It’s only when I consider future events *in life* that I do not think it’s Me*. This seems odd, given that my death necessarily comes after any future events in my life, and ought therefore to seem to have even less to do with Me* than any future events in life. But it can be explained. This feature of my attitude to death is principally grounded in susceptibility to the following line of thought: When eternity — eternal nonexistence — is in question, the gap between Me* and death that is created by the fact that I still have an indefinite amount of life to live approximates to nothing (like any finite number compared with infinity). So death — nonexistence for ever — presents itself as having direct relevance for Me* now even

²¹ Here I am strikingly different from J. Campbell, who argues that ‘fission’ (in which one person is imagined to split into two separate people) ‘would mean loss of the right to one’s autobiographical memories, my memories of what I have seen and done’ in some way that mattered (1994, p. 189).

if Me* has no clear future in life — not even tomorrow. On the vast scale of things that one naturally thinks in terms of when thinking of death, death is no significant distance away from Me*, and looms as something that will happen to Me*. This is not to say that I feel or fear that I am going to die now. The thought of eternity doesn't override common sense. But it has an emotional force that makes it seem plain that death faces Me*. If this is Heideggerian authenticity, then Heideggerian authenticity is compatible with lack of any belief in the persisting self.

Note that this line of thought will have equal force for someone who *does* think of their Me* as having a future in life: for if eternity of nonexistence is what you fear, a few years is not a protection. This idea was vivid for me every night as a young child combining an atheist upbringing with great difficulty in going to sleep.

One indirect lesson of this case is important. It is that one's sense of one's temporal nature may vary considerably depending on what one is thinking about. But the general conclusion I draw is that a sense of the self need not necessarily involve (4) a sense of it as something that has long-term continuity.²²

IX: The Self In Time; The 'Stream' of Consciousness

How does the moment-to-moment experience of consciousness relate to the sense of the self? Does it underwrite (4)? I will now consider this question.

I think William James's famous metaphor of the stream of consciousness is inept.²³ Human thought has very little natural phenomenological continuity or experiential flow — if mine is anything to go by. 'Our thought is fluctuating, uncertain, fleeting', as Hume said (1947, p. 194). It keeps slipping from mere consciousness into self-consciousness and out again (one can sit through a whole film without emerging into I-thinking self-consciousness). It is always shooting off, fuzzing, shorting out, spurting and stalling. William James described it as 'like a bird's life, . . . an alternation of flights and perchings' (1950, vol. 1 p. 243), but even this recognition that thought is not a matter of even flow retains a strong notion of continuity, in so far as a bird traces a spatio-temporally continuous path. It fails to take adequate account of the fact that trains of thought are constantly broken by detours — byblows — fissures — white noise. This is especially so when one is just sitting and thinking. Things are different if one's attention is engaged by some ordered and continuous process in the world, like a fast and exciting game, or music, or a talk. In this case thought or experience may be felt to inherit much of the ordered continuity of the phenomenon which occupies it. But it may still seize up, fly off, or flash with perfectly extraneous matter from time to time, and reflection reveals gaps and fadings, disappearances and recommencements even when there is stable succession of content.²⁴ It is arguable that the case of solitary speculative thought — in which the mind is left to its own

²² Narrative personalities may feel there is something chilling and empty in the Episodic life. They may fear it, and judge that it shows lack of wisdom, conduces to lack of moral responsibility, and is 'deficient and empty' (Plutarch, 1939, p. 217). This, however, is ignorance: even in its extreme form this life is no less intense or full, no less emotional and moral.

²³ James (1984, p. 145). Husserl is also heavily committed to the image of the stream, the '*flowing cogito*', the 'flowing conscious life in which the . . . ego lives' (1973, pp. 66, 31). For an excellent discussion of Buddhist uses of the metaphor of the stream see Collins (1982, ch. 8.4).

²⁴ This is just a phenomenological report; compare Dennett's discussion (1991, pp. 189, 237–42) of the 'pandemonium' in the mind-brain as different words, ideas, thoughts, impulses vie for emergence into consciousness.

resources and devices — merely reveals in a relatively dramatic way something that is true to a greater or lesser extent of all thought. There is an important respect in which James Joyce’s use of punctuation in his ‘stream of consciousness’ novel *Ulysses* makes his depiction of the character of the process of consciousness more accurate in the case of the heavily punctuated Stephen Daedalus than in the case of the unpunctuated Molly Bloom. Dorothy Richardson, acknowledged as the inventor of the ‘stream of consciousness’ novel in English, remarked on the ‘perfect imbecility’ of the phrase to describe what she did.²⁵

My claim is not just that there can be radical disjunction at the level of subject matter. Switches of subject matter could be absolute, and still be seamless in the sense that they involved no sensed temporal gap or felt interruption of consciousness. It seems to me, however, that such experience of temporal seamlessness is relatively rare.²⁶ When I am alone and thinking I find that my fundamental experience of consciousness is one of *repeated returns into consciousness from a state of complete, if momentary, unconsciousness*. The (invariably brief) periods of true experiential continuity are usually radically disjunct from one another in this way even when they are not radically disjunct in respect of content. (It is in fact often the same thought — or nearly the same thought — that one returns to after a momentary absence.) The situation is best described, it seems to me, by saying that consciousness is continually *restarting*. There isn’t a basic substrate (as it were) of continuous consciousness interrupted by various lapses and doglegs. Rather, conscious thought has the character of a (nearly continuous) series of radically disjunct irruptions into consciousness from a basic substrate of non-consciousness. It keeps banging out of nothingness; it is a series of comings to. It’s true that belief in the reality of flow may itself contribute to an experience of flow. But I think that the appearance of flow is undercut by even a modest amount of reflection.²⁷

²⁵ This is Richardson’s Miriam Henderson in church:

Certainly it was wrong to listen to sermons . . . stultifying . . . unless they were intellectual . . . lectures like Mr Brough’s . . . that was as bad, because they were not sermons . . . Either kind was bad and ought not to be allowed . . . a homily . . . sermons . . . homilies . . . a quiet homily might be something rather nice . . . and have not *Charity* — sounding brass and tinkling cymbal . . . *Caritas* . . . I have *none* I am sure . . . (Richardson, 1979, p 73).

Compare Molly Bloom in bed:

I want to do the place up someway the dust grows in it I think while Im asleep then we can have music and cigarettes I can accompany him first I must clean the keys of the piano with milk whatll I wear a white rose or those fairy cakes in Liptons at 7¹/₂d a lb or the other ones with the cherries in them and the pinky sugar 11d a couple of lbs of those a nice plant for the middle of the table Id get that cheaper in wait wheres this I saw them not long ago I love flowers . . . (Joyce, 1986, p. 642).

And Stephen Daedalus walking on the beach:

Who watches me here? Who ever anywhere will read these written words? Signs on a white field. Somewhere to someone in your flutiest voice. The good bishop of Cloyne took the veil of the temple out of his shovel hat: veil of space with coloured emblems hatched on its field. Hold hard. Coloured on a flat: yes, that’s right (Joyce, 1986, p. 40).

²⁶ Molly Bloom might seem to be an example of seamlessness across radical change of content, but Shaun Gallagher argues that ‘such radical disjunctions of content actually do disrupt the flow structure — content and form are not independent of one another’ (private correspondence).

²⁷ This experience seems to be in affinity with the Buddhist theory of the way in which consciousness is an interruption of ongoing, unconscious *bhavaṅga* mind, although the Buddhist theory has many special further features. See Collins (1982, pp. 238–47).

‘But perhaps the experience of disjunction is an artefact of introspection. Perhaps unexamined consciousness has true flow, and the facts get distorted by the act of trying to observe what they are.’

This seems highly implausible. Awareness of radical disjunction sometimes surfaces spontaneously and unlooked for. We can become aware that this is what has been happening, we do not see it only when we look. This is my experience, and the claim seems strongly supported by work described by Dennett (1991, e.g. ch. 11). Even if the appearance of disjunction were partly an artefact of intentional introspection, this would be a striking fact about how consciousness appears to itself, something one needed to take account of when considering the underpinnings of the sense of the self. There’s a sense in which this issue is undecidable, for in order to settle it one would need to be able to observe something while it was unobserved. Nevertheless, the view that there is radical disjunction might receive independent support from experimental psychology, and also, more indirectly, from current work on the non-mental neural correlates of consciousness.

I have been arguing — if that’s the word — that the sense of the mental self as something that has long-term continuity lacks a certain sort of direct phenomenological warrant in the moment-to-moment nature of our thought processes. It is not supported at the level of detail by any phenomenon of steady flow. If there is any support for belief in the long-term continuity of the self in the nature of moment-to-moment consciousness, it is derived indirectly from other sources — the massive constancies and developmental coherencies of *content* that often link up experiences through time, and by courtesy of short-term memory, across all the jumps and breaks of flow. One (the human being, the mental-and-nonmental whole) walks from A to B, looking around, thinking of this and that. One works in a room for an hour. Examined in detail, the processes of one’s thought are bitty, scatty, and saccadic in the way described; consciousness is ‘in a perpetual flux’, and different thoughts and experiences ‘succeed each other with an inconceivable rapidity’ (Hume, 1978, p. 252). And yet one is experientially in touch with a great pool of constancies and steady processes of change in one’s environment including, notably, one’s body (of which one is almost constantly aware, however thoughtlessly, both by external sense and by proprioception). If one does not reflect very hard, these constancies and steadinesses of development in the *contents* of one’s consciousness may seem like fundamental characteristics of the *operation* of one’s consciousness, although they are not. This in turn may support the sense of the *mental self* as something uninterrupted and continuous throughout the waking day.

I am not claiming that belief in the flow of consciousness is *necessary* to a sense of the self as something that has long-term continuity. One could think and feel that consciousness was gappy and chaotic and still believe in a mental self that had long-term continuity. This is probably the most common position among those who believe in the self, and the present, weak suggestion is only that belief in the flow of consciousness may be one interesting and suspect source of support for a sense of long-term continuity.

There is more to say, but not here. My central claim remains unchanged: one can have a full sense of the single mental self at any given time without thinking of the self as something that has long-term continuity. According to Reed ‘our sense of self is intimately related to the subjective awareness of the continuity of life. Any break

in personal time [or ‘time-gap experience’] is alarming, because it suggests some disintegration of psychic synthesis’ (Reed, 1987, p. 777). I believe that this is not generally true.

X: The Conditions Question

I have given examples of how one might set about answering phenomenological questions (1) and (2) in preparation for (4), the factual question ‘Does the self exist?’ I have no space to consider (3), the conditions question ‘What are the grounds or preconditions of possession of a sense of the mental self?’, but I think it is best approached by asking the more familiar question ‘What are the grounds or necessary conditions of self-consciousness?’, which has been widely discussed—e.g. by Kant, Fichte, Wundt, James and their followers, and, more recently, by P.F. Strawson (1966, pp. 97–112), Evans (1982, ch. 7), and others (see e.g. the contributors to Bermúdez *et al.*, 1995, and Cassam, 1997). I believe that all discussions in the analytic tradition overestimate the strength of the conditions that can be established as necessary for self-consciousness, but this is a question for another time, and I will now conclude with a wild sketch of how I think the factual question is to be answered.

XI: The Factual Question

Suppose — for the sake of argument — that the answer to the general phenomenological question is as follows: any genuine sense of the self must involve a conception of the self as [(1) + (2) + (3) + (5) + (6)] — as a single, mental thing that is distinct from all other things and a subject of experience — but need not involve a conception of it as (7) an agent, or as having (8) character or personality or (4) longer-term diachronic continuity. If we couple this answer with the equivalence claim (p. 5 above) we get the result that if there is such a thing as a mental self, it must at least fulfil conditions (1), (2), (3), (5) and (6) — one might call these the ‘core conditions’. It must be a distinct, mental thing that is correctly said to be a subject of experience and a single thing within any hiatus-free period of experience; whatever else it may be.²⁸

Is there such a thing? If there is, is it right to call it a self? I can’t legislate on how anyone should use the words ‘self’ and ‘thing’ (cf. note 8 above). It seems to me that the best answer is Yes, but many will think my Yes is close to No, because I don’t think a mental self exists in any sense that will satisfy most of those who want there to be a self. I believe the Buddhists have the truth when they deny the existence of a *persisting* mental self, in the human case, and nearly all of those who want there to be a self want there to be a persisting self.

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.²⁹ According to the Pearl view, each is a distinct existence, an individual physical thing or object, though they may exist for considerably different lengths of time. The Pearl view is not the view that mental selves are

²⁸ Obviously the view that mental selves can have personality and can be agents and have longer-term continuity is not excluded by this proposal. Very few would agree with me that agenthood is dispensable with.

²⁹ It is unlike the ‘bundle’ theory of the self, described but not endorsed by Hume, according to which the self, in so far as it exists at all, is a diachronically extended — perhaps non-continuous — thing constituted of a series of experiences (Hume, 1978, pp. 251–3, 259–63, 633–6, 657–8).

necessarily of relatively short duration—there may be beings whose conscious experience is uninterrupted for hours at a time, or even for the whole of their existence (if I believed in God, this is how I'd expect God to be). But we are not like this: the basic form of our consciousness is that of a gappy series of eruptions of consciousness from a substrate of apparent non-consciousness.

I don't suppose the Pearl view will be much liked. It sounds linguistically odd and counterintuitive. It offends against the everyday use of expressions like 'myself' to refer to enduring human beings, and nearly all theoretical speculation about the self incorporates a deep presumption that if one is arguing for the existence of the mental self one is arguing for something that exists for a substantial period of time. The Pearl view sounds even more implausible as an account of the subject of experience.³⁰

Sometimes we need to speak oddly to see clearly. I think it is important to defend the Pearl view, giving its linguistic counterintuitiveness a chance to diminish through familiarity so that one can judge it on its merits rather than on linguistic gut feeling. Perhaps the most that can be said for it is that it is the best we can do if we commit ourselves in advance to answering Yes to the question 'Is there any straightforward and metaphysically robust sense in which it is legitimate to talk of the mental self as a thing, something that really exists, like a chair or a cat, rather than merely as a Humean or Dennettian fiction?' In my view, that means that there is a lot to be said for it.

The proposal, in any case, is that the mental self — *a* mental self — exists at any given moment of consciousness or during any uninterrupted or hiatus-free period of consciousness.³¹ But it exists only for some short period of time. But it is none the less real, as *real* as any rabbit or Z-particle. And it is as much a *thing* or *object* as any G-type star or grain of salt. And it is as much a *physical* thing as any blood vessel or jackhammer or cow.

I can think of three overlapping tasks one has to undertake in order to develop the proposal. One has to say more about what it is to be a materialist, address the question 'What is a thing (or object)?', and explain further what is meant by 'ontic distinctness'. I will make one comment about each.

(i) In saying that a self is an 'ontically distinct' thing, I mean — at least — that it is not the same thing as anything else ordinarily or naturally identified as a thing. But I don't mean that it is an 'independent or separately existing entity' (Parfit, 1995, p. 18) relative to all other things naturally identified as things — such as atoms, neurons, and brains. Parfit takes a Cartesian immaterial ego to be a paradigm instance of such a separately existing entity, but I take it that a mental self's existence from t_1 to t_2 (I'll suppose this to be a two-second interval) is part of the existence from t_1 to t_2 of a set of neuron-and-neurotransmitter-(etc)-constituting atoms or fundamental particles in a certain state of activation.³²

³⁰ Dennett's account of the self as an 'abstraction', a 'Center of Narrative Gravity' (1991, pp. 426–7) may be the best one can do if one is determined to conceive the self as something that has long-term continuity.

³¹ The notion of uninterruptedness remains vague. Note that many will think that the period of consciousness must be one of explicit self-consciousness (cf. the opening quotation from Nozick), or must at least occur in a being capable of such self-consciousness. But I am not sure that this is the best thing to say.

³² Compare Van Inwagen's account (1990, pp. 94–5) of how an atom may be 'caught up in the life of an organism' while existing both before and after it. One may equally well say that each member of the set of fundamental particles is 'caught up in' the life of a mental self.

Note that this is not any sort of reductionist remark, for the phrase ‘a set of . . . particles in a certain state of activation’, as used by a consistent and realistic materialist, does not refer only or even especially to non-mental phenomena that can be adequately described by current physics or something like it. It refers just as it says, to a set of neuron-and-neurotransmitter-(etc)-constituting particles in a certain state of activation; and this existence and activity, as all genuine and realistic materialists agree, is as much revealed by and constituted by experiential phenomena as by any non-experiential phenomena discernible by physics.

The plausibility of the claim that a mental self is a *thing*, given the way it is characterized in the penultimate paragraph, depends on the success of arguments sketched in (iii) below. But it is at least clear that ontic distinctness is not separate existence. Nor, it seems, is it what Parfit has in mind when he himself distinguishes distinctness from separate existence.

Consider a human being X. I will call the portion of physical reality that consists of X the ‘X-reality’. This is a rough notion — as a physical being X is enmeshed in wide-reaching physical interactions, and is not neatly separable out as a single portion of reality — but it is serviceable none the less. Parfit offers two examples of things that stand in the relation of distinctness without separate existence: a statue and the lump of bronze of which it is made, and a nation and ‘a group of people, on some territory, living together in certain ways’.³³ By contrast, I propose that there is an analogy between the following two relations: (1) the relation between one of X’s little fingers and X, where X is considered statically at a particular moment in time; (2) the relation between a mental self that exists in the X-reality and the X-reality, where the X-reality is considered dynamically as something essentially persisting in time. In other words, I propose that there is some sort of part-whole relation to be discerned, although there is more to be said in description of the whole of which the self is a part. It seems to me that selves are as real, and as much things, as little fingers (actually it is arguable that they have a better claim to count as things than fingers do).

(ii) Genuine, realistic materialism requires acknowledgement that the phenomena of conscious experience are, considered specifically as such, wholly physical, as physical as the phenomena of extension and electricity as studied by physics (section V). This in turn requires the acknowledgement that current physics, considered as a general account of the nature of the physical, is like *Hamlet* without the prince; or at least like *Othello* without Desdemona. No one who doubts this is a serious materialist, as far as I can see. Anyone who has had a standard modern (Western) education is likely to experience a feeling of deep bewilderment — category-blasting amazement — when entering into serious materialism, and considering the question ‘What is the nature of the physical?’ in the context of the thought that the mental (and in particular the experiential) is physical; followed, perhaps, by a deep, pragmatic agnosticism.³⁴

³³ The statue just consists in the lump of bronze, and is therefore not a separately existing entity, but it is not the same as a lump of bronze: for example, we can melt down the statue and so destroy it without destroying the lump of bronze. The existence of the nation ‘just consists in the existence of a group of people, on some territory, living together in certain ways’: it is not a separately existing entity. But it is also ‘not the same as that group of people, or that territory’.

³⁴ Cf. Chomsky (1995, pp. 1–10); Russell (1954, ch. 37).

(iii) The discussion of materialism has many mansions, and provides a setting for considering the question ‘What is a thing or object?’ It is a long question, but the answer suggests that there is no less reason to call the self a thing than there is to call a cat or a rock a thing. It is arguable that disagreement with this last claim is diagnostic of failure to understand what genuine, realistic materialism involves.

‘Come off it. Even if we grant that there is a phenomenon that is reasonably picked out by the phrase “mental self”, why should we accept that the right thing to say about some two-second-long mental-self phenomenon is (a) that it is a *thing* or *object* like a rock or a tiger? Why can’t we insist that the right thing to say is simply (b) that an enduring (‘physical’) object — Louis — has a certain *property*, or (c) that a two-second mental-self phenomenon is just a matter of a certain *process* occurring in an object — so that it is not itself a distinct object existing for two seconds?’

I think that a proper understanding of materialism strips (b) and (c) of any appearance of superiority to (a). As for (c): any claim to the effect that a mental self is best thought of as a process rather than an object can be countered by saying that there is no sense in which a mental self is a process in which a rock is not also and equally a process. So if a rock is a paradigm case of a thing in spite of being equally well thought of as a process, we have no good reason to say that a self is not a thing.³⁵

‘But if there is a process, there must be something — an object or substance — in which it goes on. If something happens, there must be something to which it happens, something which is not just the happening itself.’ This expresses our ordinary understanding of things, but physicists are increasingly content with the view that physical reality is itself a kind of pure process — even if it remains hard to know exactly what this idea amounts to. The view that there is some ultimate stuff to which things happen has increasingly ceded to the idea that the existence of anything worthy of the name ‘ultimate stuff’ consists in the existence of fields of energy — consists, in other words, in the existence of a kind of pure process which is not usefully thought of as something which is happening to a thing distinct from it.

As for (b): the object/property distinction is, as Russell says of the standard distinction between mental and physical, ‘superficial and unreal’ (1954, p. 402). Chronic philosophical difficulties with the question of how to express the relation between substance and property provide strong negative support for this view. However ineluctable it is for us, it seems that the distinction must be as superficial as we must take the distinction between the wavelike nature and particlelike nature of fundamental particles to be. Obviously a great deal more needs to be said. But Kant seems to have got it exactly right in a single sentence: ‘in their relation to substance, [properties] are not in fact subordinated to it, but are the manner of existence of the substance itself’ (1996, A414/B441).

XII: Conclusion

So much for the sketch of my answer to the factual question. I think it expresses a difficult truth, but it is exiguous and probably looks very implausible. It is not designed to persuade, however; it simply marks a possible path. One can think it monstrously implausible without rejecting the approach to the problem of the self proposed in this paper: one can agree about the importance of answering (1) and (2), the two phenomenological questions and (3), the conditions question, even if one wants to give a very different answer to (4), the factual question.

³⁵ In saying this, I don’t mean to show any partiality to the ‘four-dimensionalist’ conception of objects.

References

- Barkow, J.H., Cosmides, L. and Tooby, J. (1992), *The Adapted Mind: Evolutionary Psychology and the Generation of Culture* (New York: Oxford University Press).
- Berkeley, G. (1975), *Philosophical Works*, edited by M. R. Ayers (London: Dent).
- Bermúdez, J. L., Marcel, A. and Eilan, N. (ed. 1995), *The Body and the Self* (Cambridge, MA: MIT Press).
- Campbell, A. (1994), 'Cartesian dualism and the concept of medical placebos', *Journal of Consciousness Studies*, **1** (2), pp. 230–3.
- Campbell, J. (1994), *Past, Space, and Self* (Cambridge, MA: MIT Press).
- Campbell, J. (1995), 'The body image and self-consciousness', in Bermúdez *et al.* (1995).
- Camus, A. (1982), *The Outsider*, translated by Joseph Laredo (London: Hamish Hamilton).
- Camus, A. (1995), *The First Man*, translated by David Haggood (London: Hamish Hamilton).
- Cassam, A-Q.A. (ed. 1994), *Self-knowledge* (Oxford: Oxford University Press).
- Cassam, A-Q.A. (1997), *Self and World* (Oxford: Clarendon Press).
- Chomsky, N. (1995), 'Language and nature', *Mind*, **104**, pp.1–61.
- Collins, S. (1982), *Selfless Persons* (Cambridge: Cambridge University Press).
- Crabbe, J. (ed. forthcoming) *From Soul to Self* (London: Routledge).
- Dennett, D. (1991), *Consciousness Explained* (Boston, MA: Little, Brown).
- Descartes, R. (1985), *The Philosophical Writings of Descartes*, Volumes 1 and 2, translated by J. Cottingham *et al.* (Cambridge: Cambridge University Press).
- Evans, G. (1982), *The Varieties of Reference* (Oxford: Oxford University Press).
- Farrell, B. (1996), Review of Bermúdez *et al.* (1995), *Journal of Consciousness Studies*, **3** (5–6), pp. 517–19.
- Fichte, J.G. (1982), *The Science of Knowledge*, trans. Peter Heath and John Lachs (Cambridge: CUP).
- Ginsberg, A. (1963), 'Statement to the *Burning Bush*', *Burning Bush II*.
- Glover, J. (1988), *I: The Philosophy and Psychology of Personal Identity* (Harmondsworth: Penguin).
- Holub, M. (1990), *The Dimension of the Present Moment* (London: Faber).
- Hopkins, G.M. (1959), *Sermons and Devotional Writings*, ed. C. J. Devlin (London: Oxford University Press).
- Hughes, R. (1929), *A High Wind in Jamaica* (London: Chatto & Windus).
- Hume, D. (1947), *Dialogues Concerning Natural Religion*, ed. N. Kemp Smith (Edinburgh: Nelson).
- Hume, D. (1975), *Enquiries Concerning Human Understanding*, ed. L.A. Selby-Bigge (Oxford: OUP).
- Hume, D. (1978), *A Treatise of Human Nature*, ed. L.A. Selby-Bigge and P. H. Nidditch (Oxford: OUP).
- Hurlburt, R., Happ, F. and Frith, U. (1994), 'Sampling the form of inner experience in three adults with Asperger syndrome', *Psychological Medicine*, **24**, pp. 385–95.
- Husserl, E. (1973), *Cartesian Meditations*, translated by D. Cairns (The Hague: Nijhoff).
- James, W. (1950), *The Principles of Psychology* (New York: Dover).
- James, W. (1984), *Psychology: Briefer Course* (Cambridge, MA: Harvard University Press).
- Joyce, J. (1986), *Ulysses* (Harmondsworth: Penguin).
- Kant, I. (1996), *Critique of Pure Reason*, translated by W. S. Pluhar (Indianapolis: Hackett).
- Kenny, A. (1988), *The Self* (Marquette: Marquette University Press).
- Lowe, E. J. (1996), *Subjects of Experience* (Cambridge: Cambridge University Press).
- Nagel, T. (1986), *The View From Nowhere* (New York: Oxford University Press).
- Nozick, R. (1981), *Philosophical Explanations* (Oxford: Clarendon Press).
- Parfit, D. (1984), *Reasons and Persons* (Oxford: Clarendon Press).
- Parfit, D. (1995), 'The unimportance of identity', in *Identity*, ed. H. Harris (Oxford: Clarendon Press).
- Plutarch (1939), 'On tranquillity of mind' in Plutarch, *Moralia*, volume VI, translated by W. C. Helmbold (Cambridge, MA: Harvard University Press).
- Pöppel, E. (1978), 'Time perception'. In *Handbook of Sensory Physiology*, Vol. VIII, ed. R. Held, H.W. Leibovitz and H.L. Teuber (New York: Springer).
- Reed, G. (1987), 'Time-gap experience', in *The Oxford Companion to the Mind* (Oxford: OUP).
- Richardson, D. (1979) *Pointed Roofs, Pilgrimage*, Vol 1 (London: Virago Press).
- Ruhnau, E. (1995), 'Time gestalt and the observer', in *Conscious Experience*, ed. T. Metzinger (Paderborn: Schöningh / Thorverton, UK: Imprint Academic).
- Russell, B. (1954), *The Analysis of Matter* (London: Allen and Unwin).
- Searle, J. (1992), *The Rediscovery of the Mind* (Cambridge, MA: MIT Press).
- Snowdon, P. (forthcoming), 'Personal identity and the unity of consciousness', in *Persons, Animals and Ourselves*.
- Strawson, G. (forthcoming), 'The sense of the self', in Crabbe (forthcoming).
- Strawson, P.F. (1966), *The Bounds of Sense* (London: Methuen).
- Updike, J. (1989), *Self-Consciousness* (London: Deutsch).
- Van Inwagen, P. (1990), *Material Beings* (Ithaca, NY: Cornell University Press).

Paper received February 1997