

2

The Experiential Self: Objections and Clarifications

DAN ZAHAVI

I. Introduction

Let me start with three quotes from Sartre's *L'être et le néant*—three quotes that conjointly articulate a view of consciousness that I think is widespread among phenomenologists, and which I personally endorse.

It is not reflection which reveals the consciousness reflected-on to itself. Quite the contrary, it is the non-reflective consciousness which renders the reflection possible; there is a pre-reflective *cogito* which is the condition of the Cartesian *cogito*.

(Sartre 2003: 9)

This self-consciousness we ought to consider not as a new consciousness, but as the *only mode of existence which is possible for a consciousness of something*.

(Sartre 2003: 10)

[P]re-reflective consciousness is self-consciousness. It is this same notion of *self* which must be studied, for it defines the very being of consciousness.

(Sartre 2003: 100)

What is Sartre saying here? First of all, on Sartre's view, an experience does not simply exist, it exists in such a way that it is implicitly self-given, or as Sartre puts it, it is 'for itself'. This self-giveness of experience is not simply a quality added to the experience, a mere varnish: rather for Sartre the very *mode of being* of intentional consciousness is to be *for-itself* (*pour-soi*), that is, self-conscious (Sartre 1967, 2003: 10). Sartre is, moreover, quite explicit in emphasizing that the self-consciousness in question is *not* a new consciousness. It is not something added to the experience, an additional mental state,

but rather an intrinsic feature of the experience.¹ When speaking of self-consciousness as a permanent feature of consciousness, Sartre is, consequently, not referring to what we might call reflective self-consciousness. Reflection (or higher-order representation) is the process whereby consciousness directs its intentional aim at itself, thereby taking itself as its own object. According to Sartre, however, this type of self-consciousness is derived; it involves a subject-object split, and the attempt to account for *self*-consciousness in such terms is, for Sartre, bound to fail. It either generates an infinite regress or accepts a non-conscious starting point, and he considers both options unacceptable (Sartre 2003: 8).

According to Sartre, the right alternative is to accept the existence of a pre-reflective and non-objectifying form of self-consciousness. To put it differently, on his account, consciousness has two different modes of givenness, a pre-reflective and a reflective. The first has priority since it can prevail independently of the latter, whereas reflective self-consciousness always presupposes pre-reflective self-consciousness. So to repeat, for Sartre pre-reflective self-consciousness is not an addendum to, but a constitutive moment of the original intentional experience.

In a subsequent move, Sartre then argues that consciousness, far from being impersonal and anonymous, is characterized by a fundamental selfness or selfhood precisely because of this pervasive self-givenness, self-intimation, or reflexivity. To quote the central passage from Sartre once again: ‘pre-reflective consciousness is self-consciousness. It is this same notion of *self* which must be studied, for it defines the very being of consciousness.’

2. The Experiential Self

One way to interpret Sartre’s final claim is as follows. Sartre (along with other phenomenologists) is drawing attention to a specific aspect of our experiential life, one that is so close to us, so taken for granted, that we tend

¹ Let me emphasize that the choice of the term ‘intrinsic’ is precisely meant to emphasize the difference from a higher-order or reflection-based account of self-consciousness, where self-consciousness is conceived in terms of a relation between two mental states. The term is not meant to indicate that we are dealing with a feature that our experiences possess in complete independence of everything else. To put it differently, to talk of self-consciousness as an intrinsic feature of experience is not to deny that the (self-conscious) experience in question is also intentional and world-directed.

to overlook it. As illustration, consider the following example. Imagine a situation where you first see a green apple and then see a yellow lemon. Then imagine that your visual perception of the yellow lemon is succeeded by a recollection of the yellow lemon. How should we describe the phenomenal complexity? One rather natural way to do so (which leaves out the fact and added complication that the whole scenario is played out in the imagination) is as follows: First, we have an intentional act of a specific type (a perception) which is directed at a specific object (an apple). Then we retain the intentional act-type (the perception), but replace the apple with another object (a lemon). In a final step, we replace the perception with another act-type (a recollection) while retaining the second object. By going through these variations, we succeed in establishing that an investigation of our experiential life shouldn't merely focus on the various intentional objects we can be directed at, but that it also has to consider the different intentional types or attitudes we can adopt. This is all trivial. But then consider the following question. If we compare the initial situation where we perceived a green apple with the final situation where we recollected a yellow lemon, there has been a change of both the object and the intentional type. Does such a change leave nothing unchanged in the experiential flow? Is the difference between the first experience and the last experience as radical as the difference between my current experience and the current experience of someone else? We should deny this. Whatever their type, whatever their object, there is something that the different experiences have in common. Not only is the first experience retained by the last experience, but the different experiences are all characterized by the same fundamental first-personal self-givenness. They are all characterized by what might be called a dimension of *for-me-ness* or *mineness* (Sartre uses the term *ipseity*—selfhood—from the Latin, *ipse*). It is, however, important to point to the special nature of this mineness. It is not meant to suggest that I own the experiences in a way that is even remotely similar to the way I possess external objects of various sorts (a car, my trousers, or a house in Sweden). Nor should it be seen primarily as a contrastive determination. When young children start to use the possessive pronoun, it frequently means 'not yours'. But as Husserl observes in one of his manuscripts, when it comes to the peculiar mineness (*Meinheit*) characterizing experiential life, this can and should be understood without any contrasting others (Husserl 1973b: 351), although it may form the basis of the self-other discrimination.

THE EXPERIENTIAL SELF: OBJECTIONS AND CLARIFICATIONS 59

Some might object that there is no property common to all my experiences, no stamp or label that clearly identifies them as mine. But this objection is misplaced in that it looks for the commonality in the wrong place. The *for-me-ness* or *mineness* in question is not a quality like scarlet, sour, or soft. It doesn't refer to a specific experiential content, to a specific *what*, nor does it refer to the diachronic or synchronic sum of such content, or to some other relation that might obtain between the contents in question. Rather, it refers to the distinct givenness or *how* of experience. It refers to the first-personal presence of experience. It refers to the fact that the experiences I am living through are given differently (but not necessarily better) to me than to anybody else. It could consequently be claimed that anybody who denies the *for-me-ness* or *mineness* of experience simply fails to recognize an essential constitutive aspect of experience. Such a denial would be tantamount to a denial of the first-person perspective. It would entail the view that my own mind is either not given to me at all—I would be mind- or self-blind—or present to me in exactly the same way as the minds of others.

Sartre's basic move, which is to link the notion of self to pre-reflective self-consciousness, is nicely captured in a formulation by another French phenomenologist, Michel Henry, who writes that the most basic form of selfhood is the one constituted by the very self-manifestation of experience (Henry 1963: 581, 1965: 53). But who or what is this self that has, or lives through, the experiences? The phenomenological account I favor can be seen as occupying a kind of middle position between two opposing views. According to the first view, the self is some kind of unchanging soul substance that is distinct from, and ontologically independent of, the mental experiences and worldly objects it is the subject of. According to the second view, there is nothing to consciousness apart from a manifold of interrelated changing experiences. We might, to adopt some traditional labels, speak of the self as the owner of experiences, and the self as the bundle of experiences, respectively. By contrast, the self currently under consideration—and let us call it the experiential core self—is not a separately existing entity, but neither is it simply reducible to a specific experience or (sub-)set of experiences. If I compare two experiences, say the perception of a green apple and the recollection of a yellow lemon, I can focus on the difference between the two, namely the respective object and mode of presentation, but I can also attend to that which remains the same, namely the first-personal self-givenness of both experiences. To put it differently, we can distinguish

a multitude of changing experiences from a *ubiquitous* dimension of first-personal self-givenness, and the proposal is that we identify the latter with the experiential core self. So on this view, the self is defined as the very subjectivity of experience, and is not taken to be something that exists independently of, or in separation from, the experiential flow.

When talking of first-personal self-givenness, one shouldn't think of self-reference by means of the first-person pronoun; in fact, one shouldn't think of a linguistically conditioned self-reference at all. Nor should one have an explicit or thematic kind of self-knowledge in mind, one where one is aware of oneself as a distinct individual, different from other individuals. No, first-personal self-givenness is meant to pinpoint the fact that (intransitively) conscious mental states are given in a distinct manner, with a distinct subjective presence, to the subject whose mental states they are, a way that in principle is unavailable to others. When saying 'distinct', the claim is not that the subject of the experience is explicitly aware of their distinct character: the point is not that the subject is necessarily attending to the distinctness in any way. But the first-personal self-givenness is distinctive even before, say, a child becomes explicitly aware of it, just as it is unavailable to others even prior to a child recognizing this.

Now, there are obviously various ways one might both elaborate on, as well as challenge this account. Does it fall victim to what Block has called the refrigerator illusion? Is our ordinary waking life characterized by an absorbed mindless coping, rather than by pre-reflective self-awareness? Is first-personal self-givenness and mineness a post hoc fabrication, something imputed to experience by subsequent mentalizing and theorizing? I don't have time to respond to these worries on this occasion (but see Zahavi 1999, 2005, 2009). Rather, I want to press ahead and directly engage with some of the criticisms that have been raised by defenders of a no-self doctrine. I will, more specifically, look at various objections recently made by Albahari and Dreyfus.

3. The Illusory Self

In her book *Analytical Buddhism: The Two-Tiered Illusion of Self*, Albahari's basic aim is to argue that the self is an illusion. What notion of self is she out to deny? She initially provides the following definition: The self should be understood as a unified, happiness-seeking, unbrokenly persisting,

ontologically distinct conscious subject who is the owner of experiences, the thinker of thoughts, and the agent of actions. What is interesting about Albahari's proposal is that, whereas many advocates of a no-self doctrine have denied that consciousness is characterized by unity, unbrokenness, and invariability, and taken the denial of these features to amount to a denial of the reality of the self, Albahari considers all three to be real features of consciousness, but she nevertheless considers the self to be illusory (Albahari 2006: 3).

To get clearer on why she thinks this is the case let us look more closely at a distinction she introduces among different forms of ownership, namely *possessive ownership*, *perspectival ownership*, and *personal ownership*. We can ignore possessive ownership, which in this context is of less interest, since it merely denotes the fact that certain objects (a car, a pair of trousers, etc.) can be regarded as mine by right of social convention. But what is the difference between personal ownership and perspectival ownership? Personal ownership is a question of identifying oneself as the personal owner of an experience, thought, action: it is a question of appropriating certain experiences, actions, thoughts, etc. as one's own, that is, a question of either thinking of them as being *mine* or apprehending them as being part of *me* (and this is something that can occur either pre-reflectively or reflectively). By contrast, for a subject to own something in a perspectival sense is simply for the experience, thought, or action in question to present itself in a distinctive manner to the subject whose experience, thought or action it is. So the reason I can be said to perspectivally own my thoughts or perceptions—if one will excuse this slightly awkward way of talking—is because they appear to me in a manner that is different from how they can appear to anybody else. When it comes to objects external to the subject, what will be perspectivally owned isn't the object, but the specific manner through which the object appears to the subject (Albahari 2006: 53).

Albahari argues that there is a close link between having a sense of personal ownership and having a sense of self. When the subject identifies certain items as being itself or being part of itself, it will harbor a sense of personal ownership towards the items in question. But this very process of identification generates the sense of a self-other distinction. It constitutes a felt boundary between what belongs to self and what doesn't. Thereby the self is cast as a unified and ontologically distinct entity—one that stands apart from other things (Albahari 2006: 73, 90). In this way, the subject

understood as a mere point of view is turned into a substantial personalized entity (Albahari 2006: 94). To put it differently, for Albahari, there is more to being a self than being a point of view, than having perspectival ownership.

One way to bring out the difference between perspectival and personal ownership is to point to possible dissociations between the two. Pathology seems to provide some examples. In cases of depersonalization, we can come across thoughts, feelings, etc. which are perspectively owned, that is, which continue to present themselves in a unique manner to the subject, without however being felt as the subject's own (Albahari 2006: 55). Thus on Albahari's reading, the process of identification fails in depersonalization, and as a consequence, no sense of personal ownership regarding the experience in question will be generated (Albahari 2006: 61).

Let us now consider Albahari's self-skepticism. What does it mean for the self to lack reality? What does it mean for the self to be illusory? On Albahari's account, an illusion involves a conflict between appearance and reality. X is illusory if x does not have any appearance-independent reality, but nevertheless purports to have such reality, that is, we are dealing with an illusion if x purports through its appearance to exist in a particular manner without really doing so (Albahari 2006: 122). One obvious problem, however, with such a definition is whether it at all makes sense to apply it to the self. Does the self really purport to exist outside of its own appearance, or is the reality of the self rather subjective or experiential? This consideration leads Albahari to redefine the notion of illusion slightly. If the self purports to be what she calls unconstructed, that is, independent from the experiences and objects it is the subject of, and if it should turn out that it in reality depends, even if only partially, on perspectively ownable objects (including various experiential episodes), then the self must be regarded as being illusory (Albahari 2006: 130).

Albahari also emphasizes the need for a distinction between self and sense of self. To have a sense of x , doesn't necessarily entail that x exists. Indeed, whereas Albahari takes the sense of self to exist and be real, she considers the self itself to be illusory (Albahari 2006: 17). Contrary to expectations, our sense of self is not underpinned by an actually existing ontologically independent self-entity. Rather, all that really exists is the manifold of thoughts, emotions, perceptions, etc. as well as a pure locus of apprehension, which Albahari terms *witness-consciousness*. It is the experiential flow in conjunction

with this locus of apprehension that generates the sense of self. But if this is so, the self lacks an essential property of selfhood, namely ontological independence (Albahari 2006: 72). In short, the illusory status of the self is due to the fact that the self does not have the ontological status it purports to have. Thoughts appear to be owned and initiated by an independently existing unified self, but rather than preceding the experiences, rather than thinking the thoughts, it is in reality the other way around. It is not the self that unifies our thoughts and experiences, they do so themselves with some help from the accompanying witness-consciousness (Albahari 2006: 130–132). To repeat, although it might seem to the subject as if there is a pre-existing self which identifies with various intentional states, the reality of the matter is that the self is created and constructed through these repeated acts of identification (Albahari 2006: 58).

As I mentioned at the beginning, an interesting aspect of Albahari's proposal is that she considers many of the features traditionally ascribed to the self to be real, it is just that they—in her view—become distorted and illusory if taken to be features of the self (Albahari 2006: 74). For instance, Albahari takes our conscious life to be characterized by an intrinsic, but elusive, sense of subjective presence, one that is common to all modalities of awareness, that is, one that is common to seeing, hearing, thinking, feeling, introspecting, etc. (Albahari 2006: 112, 144, 156). What does this subjective presence amount to? It includes the experience of being the perspectival owner of various experiences. It also includes diachronic and synchronic unity. Although we experience various objects, and although the objects we experience might change from one moment to the next, there still appears to be an unbroken consciousness that observes the change without itself changing (Albahari 2006: 155). Indeed, while from a first-person perspective it certainly makes sense to say that I have various experiences, we automatically feel them to belong to one and the same consciousness. For Albahari, all these features are properly ascribed to the witness-consciousness, and she is adamant that we have to distinguish witness-consciousness from self. Whereas the latter on her definition involves felt boundaries between self and non-self, the former doesn't.

Let me recapitulate. For Albahari, one can be aware without being presented to oneself as an ontologically unique subject with personalized boundaries that distinguishes a *me* from the rest of the world. One can be aware without being aware of oneself as a personal owner, a thinker of

thoughts, an agent of actions. Examples that come to mind are cases of pathology. Albahari asks us to consider both the real life case of epileptic automatism and the hypothetical case of global depersonalization. In both cases, the person or patient would be awake and responsive to the environment, so there would be awareness present. But there would be no sense of a bounded individual self; there would be a complete lack of personal ownership; there would be no sense of me or mine (Albahari 2006: 171, 177). Albahari suggests that such a state of mind might not only be encountered in pathologies, but also in newborn infants, and in primitive organisms. And as she then points out in the conclusion of her book, and this is of course where her Buddhist orientation becomes evident, if we were to attain enlightenment, we would move from consciousness-plus-self-illusion to consciousness-*sans*-self-illusion, and the latter condition, although strictly speaking not identical with global depersonalization—after all, it correlates with highly advanced cognitive capacities—might nevertheless be compared to it (Albahari 2006: 161, 207).

4. Self vs No-Self

The debate between advocates of self and no-self accounts is complicated by the fact that there is rather little consensus about what precisely a self amounts to, just as there is little agreement on what a no-self doctrine entails. Albahari's account in *Analytical Buddhism* constitutes a neat example of this. As we have just seen, Albahari basically denies the reality of the self and argues that it is illusory. To that extent, she should obviously count as a defender of a no-self account. At the same time, however, Albahari ascribes a number of features to what she calls witness-consciousness—features including invariance, unconstructedness, and ontological independence, features that many defenders of a traditional notion of self would consider essential and defining features of self. In fact, whereas I would suggest that we replace the traditional notion of a 'subject of experience' with the notion of a 'subjectivity of experience'—the first phrasing might suggest that the self is something that exists apart from, or above, the experience and, for that reason, something that might be encountered in separation from the experience, or even something the experience may occasionally lack, the

second phrasing, however, excludes these types of misunderstanding—Albahari wants to retain the former notion, since she considers the subject ontologically distinct from the experiences. Some might consequently claim that Albahari, despite her official allegiance to the no-self doctrine, is actually committed to a more robust notion of self than many contemporary defenders of the self, myself included.² But of course, one might also make the reverse move. I defend the reality of self, but according to Albahari, the notion of self that I operate with is so thin, and ultimately so revisionary in nature, that she in a recent article has claimed that my position ends up being very similar to that of the no-self theorists I am criticizing (Albahari 2009: 80). When I first encountered this criticism, I was somewhat puzzled, but I have subsequently come to realize that there is an obvious sense in which Albahari is right. It all comes down to what precisely a no-self doctrine amounts to. As Ganeri has recently pointed out, there is no simple answer to the question of whether the aim of the no-self doctrine—insofar as one can at all speak of it in the singular—is to identify and reject a mistaken understanding of self—one that perpetuates suffering—or whether the point is rather to reject and dispel all notions of self (Ganeri 2007: 185–186).³

In his paper in this volume, Dreyfus has explicitly defended the view that, although the no-self view does entail a denial of a self-entity, it shouldn't be read as entailing a denial of subjectivity. There is on his view no enduring experiencing subject, no inner controller or homunculus. Rather what we find is an ever-changing stream of consciousness. This stream is, however, to be conceived of as a process of self-awareness. Dreyfus consequently argues that consciousness is characterized by pervasive reflexivity, by a basic

² Consider that although Albahari denies unconstructedness of self, she ascribes it to witness-consciousness. As she puts it at one point, 'awareness must be shown to exist in the manner it purports to exist. Awareness purports to exist as a witnessing presence that is unified, unbroken and yet elusive to direct observation. As something whose phenomenology purports to be unborrowed from objects of consciousness, awareness, if it exists, must exist as *completely unconstructed* by the content of any perspectively ownable objects such as thoughts, emotions or perceptions. If *apparent* awareness... turned out to owe its existence to such object-content rather than to (unconstructed) *awareness itself*, then that would render awareness constructed and illusory and hence lacking in independent reality' (Albahari 2006: 162). This seems to commit one to viewing awareness as an ontologically independent region. It is not clear to me why one would want to uphold such a view of consciousness in the first place.

³ Needless to say there is also a rather significant difference between claiming that experience is fundamentally selfless and claiming that a dissolution or annihilation of self is an ultimate state we can (and should) seek to attain.

self-presencing that is part and parcel of our experiential life, and not to be conceived of as an additional or separate act of cognition. In opposition to some of the bundle theorists, Dreyfus consequently denies that experiences are fundamentally impersonal, as if the attribution of first-personal self-giveness to our experiential life is a post hoc fabrication. Rather, our experiences are from the very start intrinsically self-specified (Dreyfus, this volume, p. ??). But although Dreyfus, by implication, is prepared to accept the reality of subjectivity, he insists that distortion arises the moment we interpret this subjectivity as a bounded, unified self (Dreyfus, this volume, p. ??). In short, the undeniable presence of a transient flow of self-aware experiences doesn't entail the existence of an enduring self-entity, rather the latter is on Dreyfus' view an illusory reification (Dreyfus, this volume, p.??). More specifically, whereas Dreyfus wants to retain perspectival ownership and synchronic unity—and claims that both features are guaranteed by subjectivity—he argues that there is no diachronically unified self. There is no enduring entity that stays the same from childhood to adulthood.

Let me divide my critical rejoinder into three parts.

1. First of all, I reject the univocal definition of self provided by Dreyfus and Albahari. Both are very confident in spelling out what a self is, and after having defined it, they then proceed to deny its existence. In my view, however, the definition they provide is overly simplistic. There is no doubt that some people have defended the notion of self that Albahari and Dreyfus operate with, but I would dispute the claim that their notion is *the default notion*, that is, that it is either a particularly classical notion of self or that it is a particularly commonsensical notion, that is, one that is part of our folk psychology. Consider again the claim that the self—if it exists—is some kind of ontologically independent invariant *principle of identity* that stands apart from, and above, the stream of changing experiences; something that remains unchanging from birth to death; something that remains entirely unaffected by language acquisition, social relationships, major life events, personal commitments, projects, and values, something that cannot develop or flourish nor be disturbed or shattered. Frankly, I don't see such a notion as being very much in line with our pre-philosophical, everyday understanding of who we are. As for the claim that the definition captures *the* (rather than *a*) traditional philosophical understanding of self, this is also something I would dispute. Just consider, to take some (not entirely)

randomly chosen examples: the accounts we find in Aristotle or Montaigne (for informative historical overviews, cf. Sorabji 2006, Seigel 2005). In any case, when comparing the definition of self provided by Albahari and Dreyfus to the definitions found in contemporary discussions of self, it will immediately be evident that the latter discussions are far more complex, and far more equivocal, and that there are far more notions of self at play, including notions of ecological, experiential, dialogical, narrative, relational, embodied, and socially constructed selves. This complexity is ignored by Albahari and Dreyfus, and they thereby fail to realize that many of the contemporary notions of self—including those employed by most empirical researchers currently interested in the development, structure, function, and pathology of self—are quite different from the concept they criticize. To mention just one discipline that can exemplify this, consider developmental psychology and the work of developmental psychologists such as Stern (1985), Neisser (1988), Rochat (2001), Hobson (2002), or Reddy (2008). Thus, rather than saying that the self does not exist, I think self-skeptics should settle for a far more modest claim. They should qualify their statement and instead deny the existence of a special kind of self.

2. Albahari and Dreyfus both insist on distinguishing subjectivity and selfhood. Although Dreyfus doesn't say so explicitly, he would presumably agree with Albahari when she claims that my own notion of self is too thin and minimal, too deflationary and revisionary. In reply, let me right away concede that my thin notion of self is unable to accommodate or capture all ordinary senses of the term 'self'. In fact, it has some clear limitations, which I will return to in a moment. But although it certainly doesn't provide an exhaustive understanding of what it means to be a self, I think the very fact that we employ notions like *first-person* perspective, *for-me-ness* and *mineness* in order to describe our experiential life, the fact that the latter is characterized by a basic and pervasive immanent reflexivity, by *self-specificity* and pre-reflective *self-awareness*, is sufficient to warrant the use of the term 'self'. When arguing that an account of our experiential life that fails to include a reference to self is misleading and inadequate, I am to a large extent motivated by my opposition to the impersonality thesis, the no-ownership view, the strong anonymity claim (or whatever we want to call the position in question). I wish to insist on the basic (and quite formal) individuation of experiential life as well as on the irreducible difference

between one stream of consciousness and another stream of consciousness. Indeed, rather than obstructing or impeding a satisfactory account of intersubjectivity, an emphasis on the inherent and essential individuation of experiential life is a prerequisite for getting the relation and difference between self and other right. I consequently fail to see how a radical denial of the reality of self will ever be able to respect the otherness of the other. To put it differently, I don't think the question of whether there is one or two streams is a matter of convention. The fact that we frequently share the same opinions, thoughts, beliefs, and values doesn't change this. As Wilde put it, 'Most people are other people. Their thoughts are some one else's opinions, their lives a mimicry, their passions a quotation' (Wilde 1969: 97). But again, this observation targets a different issue.

Of course, some might maintain that first-personal givenness is just too formal to act as individuating principle. After all, all experiences, not only mine but everybody's, are characterized by first-personal givenness, so how could such givenness serve to pinpoint and define me? This objection is wrongheaded, however, in that it precisely fails to take the first-person perspective seriously. Consider two clones that are qualitatively identical when it comes to physical and mental characteristics. From a third-person perspective, it would indeed be hard to distinguish the two (except in terms of spatial location), and the presence of first-personal givenness would be useless as a criterion of individuation, since both of their experiential streams would possess it. Indeed, from a third-person perspective there would be no significant difference at all between the first-person givenness characterizing the experiential stream of clone A, and the first-person givenness characterizing the experiential stream of clone B. But compare, then, what happens if we instead adopt the first-person perspective. Let us assume that I am one of the clones. Although my mental and physical characteristics are qualitatively identical to those of my 'twin', there will still remain a critical and all-decisive difference between me and him, a difference that would prevent any confusion between the two of us. What might that difference consist in? It obviously has to do with the fact that only my experiences are given in a first-personal mode of presentation to me, whereas the qualitatively identical experiences of my clone are not given first-personally to me at all, and are therefore not part of my experiential life. As mentioned earlier, it is the particular first-personal *how* rather than some specific content which most

fundamentally distinguishes my experiences from the experiences had by others. This is why I wish to insist that the subjectivity of experience, its first-personal character, although quite formal, does individuate experiential life. This is why it can function as placeholder for features traditionally associated with the self.

Importantly, this emphasis on the first-personal character of experience does not entail an endorsement of the view that the self is unconstructed or unconditioned, in the sense of being ontologically independent of the experiences or the surrounding world, nor does it entail the view that the self is bounded in the sense of involving a strict division between self and world.⁴ As a case in point, consider and compare Neisser's and Heidegger's notions of self. According to Neisser's notion of ecological self, all perception involves a kind of self-sensitivity: all perception involves a co-perception of self and of environment (Neisser 1988). As for Heidegger, he explicitly argues that we should look at our intentional experiences if we wish to study the self. On his account, our experiential life is world-related, and there is a presence of self when we are engaged with the world, that is, self-experience is the self-experience of a world-immersed self (Heidegger 1993: 34, 250). But to deny that there is a strict boundary between self and world, to concede that self and world cannot be understood independently of each other and that the boundary between the two might be plastic and shifting, is not to question the reality of the difference between the two. To take an everyday example, consider the ever-shifting boundary between the sea and the beach. That the boundary keeps shifting is no reason to deny the difference between the two. To put it differently, contrary to the views of Albahari and Dreyfus, I would dispute that unconstructedness and boundedness are essential features of self, features that any viable notion of self must include. This is also why I reject the attempt to distinguish subjectivity and selfhood. As I see it, to reject the existence of self while endorsing the reality of subjectivity is to miss out on what subjectivity really amounts

⁴ Though, as already pointed out, I am committed to the view that there is indeed a firm boundary between self and other—as long as our concern is limited to the experiential notion of self. To quote James, 'Absolute insulation, irreducible pluralism, is the law. It seems as if the elementary psychic fact were not *thought* or *this thought* or *that thought*, but *my thought*, every thought being *owned*. Neither contemporaneity, nor proximity in space, nor similarity of quality and content are able to fuse thoughts together which are sundered by this barrier of belonging to different personal minds. The breaches between such thoughts are the most absolute breaches in nature' (James 1890: 226).

to—it is to pay lip service to the idea that we should take the first-person perspective serious.

3. My third comment concerns the metaphysical framework we are operating within. In recent years, quite a number of people have stressed the existence of convergent ideas in Western phenomenology and Buddhism. It has been claimed that both traditions represent serious efforts to nurture a disciplined first-person approach to consciousness (cf. Varela and Shear 1999), and some have even started to speak of Buddhist phenomenology (cf. Lusthaus 2002). I am not denying the truth of this, but when appraising Buddhist views of the nature and status of self, one should not overlook the fact that they are also driven and motivated by strong metaphysical and soteriological concerns, and that this occasionally leads to claims and conclusions that are quite far removed from phenomenology. As an example, consider the Abhidharmic view that billions of distinct mind-moments occur in the span of a blink of the eye (cf. Bodhi 1993: 156).

Dennett (1992) and Metzinger (2003) both deny the reality of the self, and part of their reason for doing this, part of the reason why they think the self is fictitious, is that a truly fundamental account of reality on their view can dispense with self. Some Buddhist metaphysicians would share this view (cf. Siderits' and MacKenzie's contributions to this volume). Although I have sympathy with the idea that we shouldn't multiply entities beyond necessity, I think the view in question is far too austere. It is hard to see why one shouldn't declare social reality fictitious on the same account. If there is no self, there can hardly be a *you* or a *we* either. In fact, it is hard to see why we shouldn't also declare the world we live in and know and care about (including everyday objects and events like chairs, playing cards, operas, or marriage ceremonies) illusory. Again, such a view is quite different from the phenomenological attempt to rehabilitate our life-world.

In the preceding, I have discussed a thin experiential notion of self, and have tried to present and defend this view. Ultimately, however, I favor what might be called a multidimensional account of self. I think the self is so multifaceted a phenomenon that various complementary accounts must be integrated if we are to do justice to its complexity. I consequently don't think that the thin notion I have defended above is sufficient. It must be supplemented by thicker notions that capture and do justice to other important aspects of self. More specifically, I think our account of human

reality is inadequate if we don't also consider the self that forms plans, makes promises, and accepts responsibilities, the self that is defined and shaped by its values, ideals, goals, convictions, and decisions. Consider as a case in point the issue of emotional investment: consider that we respond emotionally to that which matters to us, to that which we care about, to that towards which we are not indifferent. In that sense, one might argue that emotions involve appraisals of what has importance, significance, value, and relevance to oneself. Consider the extent to which emotions like shame, guilt, pride, hope, and repentance help constitute our sense of self. Consider in this context also the role of boundaries and limits. Your limits express the norms and rules you abide by; they express what you can accept and what you cannot accept. They constitute your integrity. To ask others to respect your boundaries is to ask them to take you seriously as a person. A violation of, or infringement upon, these boundaries is felt as invasive, and in some cases as humiliating. To put it differently, when it comes to these facets of self, I think boundaries, values, and emotions are extremely important, but I don't think an emphasis on boundaries has much to do with the endorsement of an enduring soul-substance that remains the same from birth to death. And I don't see why opposition to the latter should necessitate a rejection of the former as well. We are dealing with a culturally, socially, and linguistically embedded self that is under constant construction. But is this fact a reason for declaring the self in question illusory? I don't see why, unless, that is, one's prior metaphysical commitments dictate it.

5. Diachronic Unity and the Self

Let me end with a question that I quite on purpose have saved for last. It concerns the relation between self and diachronic unity. Is persistence and temporal endurance a defining feature of self?

We all have a direct experience of change and persistence. We can hear an enduring tone or a melody, just as we can see the flight of a bird. This phenomenological finding must be accounted for, and as a distinguished line of thinkers—including James, Bergson, Husserl, and more recently, Dainton—have argued, a mere succession of synchronically unified but isolated momentary points of experience cannot explain and account for our experience of duration. To actually perceive an object as enduring over

time, the successive phases of consciousness must somehow be united experientially, and the decisive challenge is to account for this temporal binding without giving rise to an infinite regress, that is, without having to posit yet another temporally extended consciousness whose task is to unify the first-order consciousness, and so forth ad infinitum. To account for the diachronic unity of consciousness, there is, however, no need for an appeal to some undivided, invariable, unchanging, trans-temporal entity. In order to understand the unity in question, we do not need to search for anything above, beyond, or external to the stream itself. Rather, following Husserl, I would propose that the unity of the stream of consciousness is constituted by inner time-consciousness, by the interplay between what Husserl calls primal impression, retention, and protention. Rather than being pre-given, it is a unity that is established or woven. This is not the right place to delve into the intricacies of Husserl's complex account (see, however, Zahavi 2003, 2004, 2007a), but on his account, even the analysis of something as synchronic as the conscious givenness of a present experience would have to include a consideration of temporality. For the very same reason, I would reject Dreyfus's attempt to make a sharp distinction between synchronic unity which he accepts, and diachronic unity which he rejects. You cannot have synchronic unity without some amount of diachronic unity (if ever so short-lived). To claim otherwise is to miss the fundamental temporal character of consciousness. Now, perhaps it could be objected that our experience of diachronic unity is after all 'merely' phenomenological and consequently devoid of any metaphysical impact. But to think that one can counter the phenomenological experience of unity over time with the claim that this unity is illusory and that it doesn't reveal anything about the true metaphysical nature of consciousness is to make use of the appearance-reality distinction outside its proper domain of application. This is especially so, given that the reality in question, rather than being defined in terms of some spurious mind-independence, should be understood in terms of experiential reality. For comparison, consider the case of pain. Who would deny that pain experience is sufficient for the reality of pain? To put it differently, if one wants to dispute the reality of the diachronic unity of consciousness, one should do so by means of more convincing phenomenological descriptions. To argue that the diachronic unity of consciousness is illusory because it doesn't match any unity on the subpersonal level is to misunderstand the task at hand.

Very well, the skeptic might retort, but accepting that our experiential life has a certain temporal density and extension is hardly the same as accepting the existence of a persisting self from birth to death. Quite right, I would reply, but even the former might be sufficient if you want to defend the existence of transient short-term selves, as Galen Strawson has consistently done (Strawson 2000), and I don't think Dreyfus has provided arguments against that specific notion of self.⁵ But more importantly, although Dreyfus denies that in his own case there is an entity that has endured from his childhood in Switzerland to his being a grown-up adult in the US, he does concede that we, in the case of episodic memory, do not have two absolutely different persons, the person remembering and the person remembered. In fact, on his account, we do have to keep first-personal self-givenness in mind, and although the remembering person is in some sense different from the one being remembered, the difference is certainly not as great as the one that separates me from other people (Dreyfus, this volume, p.??). My present act of remembering and the past act that is being remembered both share similar first-personal self-givenness. They consequently have something in common that distinguishes them from the experiences of others. As Dreyfus continues, when I remember a past experience, I don't just recall its content, I also remember it as being given to me. Now on his account, there is something distorting about this, insofar as the past experience is remembered as mine (thereby suggesting the existence of an enduring self), but as he continues, this isn't a complete distortion either (Dreyfus, this volume, p.??).

On my account, there is no experiential self, no self as defined from the first-person perspective, when we are non-conscious.⁶ But this does not

⁵ In his defense of a no-self account, Krueger claims that a Who is neither necessary nor sufficient for a How. In arguing for this view, Krueger repeatedly concedes that it might be legitimate to speak of minimal selves (rather than of a minimal self) (Krueger, this volume, p. ??). However, I find it quite hard to understand how the existence of a plurality of selves is compatible with, or might even count in favor of, a no-self theory, unless, of course, one stacks the deck by presupposing a quite particular definition of self.

⁶ This is also why I don't think the notion of experiential self will allow us to solve all relevant questions regarding personal identity and persistence over time. Consider, for instance, the case of a man who early in life makes a decision that proves formative for his subsequent life and career. The episode in question is, however, subsequently forgotten by the person. He no longer enjoys first-person access to it. If we restrict ourselves to what can be accounted for by means of the experiential core self, we cannot speak of the decision as being his, as being one he made. Or take the case where we might wish to ascribe responsibility for past actions to an individual who no longer remembers them. By doing that we postulate an identity between the past offender and the present subject, but the identity in question is

necessarily imply that the diachronic unity of self is threatened by alleged interruptions of the stream of consciousness (such as dreamless sleep, coma, etc.), since the identity of the self is defined in terms of givenness rather than in terms of temporal continuity. To put it differently, experiences that I live through from a first-person perspective are by definition mine, regardless of their content and temporal location. Thus, I don't think there is any mistake or distortion involved in remembering the past experiences as mine. Obviously this is not to say that episodic memory is infallible—I might have false beliefs about myself—I am only claiming that it is not subject to the error of misidentification (cf. Campbell 1994: 98–99). But does that mean that I take the first-personal self-givenness of the experiences as evidence for the persistence of an underlying enduring self? No, I don't, since the self I have been discussing in this paper is the experiential self, the self as defined from the first-person perspective—neither more nor less. I think this self is real and that it possesses real diachronicity, but as already mentioned I don't think its reality—its phenomenological reality—depends on its ability to mirror or match or represent some non-experiential enduring ego-substance. Having said this, let me just add that, although I don't think there is distortion involved in remembering a past experience as mine, there is admittedly and importantly more than just pure and simple identity. Episodic memory does involve some kind of doubling or fission; it does involve some degree of self-division, self-absence, and self-alienation. Episodic memory constitutes a kind of self-experience that involves identity as well as difference. At least this is pretty much phenomenological orthodoxy. As Husserl already insisted, recollection entails a self-displacement, and he went on to argue that there is a structural similarity between recollection and empathy (Husserl 1954: 189, 1966: 309, 1973a: 318, 1973b: 416). A related idea is also to be found in Merleau-Ponty, who wrote that our temporal existence is both a condition for, and an obstacle to,

again not one that can be accounted for in terms of the experiential core self. But given my commitment to a multidimensional account of self, I would precisely urge us to adopt a multilayered account of self. We are more than experiential core selves, we are, for instance, also narratively configured and socially constructed persons (cf. Zahavi 2007b). We shouldn't forget that our life-stories are multi-authored. Who we are is not something we exclusively determine ourselves. It is also a question of how we are seen by others. Even if there is no experiential self (no self as defined from the first-person perspective) when we are non-conscious, there are various other aspects of self that remain, and which make it perfectly legitimate to say that *we* are non-conscious, that is, that we can persist even when non-conscious. For a recent, very elaborate, and rather metaphysical discussion of these questions, cf. Dainton (2008).

our self-comprehension. Temporality contains an internal fracture that permits us to return to our past experiences in order to investigate them reflectively, yet this very fracture also prevents us from fully coinciding with ourselves (Merleau-Ponty 2002: 402).

But again, to some this answer might be dissatisfactory and evade the real issue: Is there or is there not an identical self from birth through childhood, adolescence, adulthood, and old age. Ultimately, I think this question is overly simplistic, since it presupposes that the self is one thing, and that there is a simple yes or no answer to the question. I disagree, and think the answer will depend on what notion of self we are talking about. But let me forego this complication and stick to the problem of whether or not the experiential self, the self as defined in terms of subjectivity, remains invariant across large time stretches. To put it differently, when remembering—from the first-person perspective – an episode that took place fifteen years ago, when remembering that past experience as mine, are we then confronted with a case where the experiential self has remained the same? Is the experiential self that originally lived through the experience 15 years ago, and the experiential self that today recalls the past experience, one and the same numerically identical self, or are we merely dealing with a relationship between two qualitatively similar selves, where the current self might stand, say, in a unique causal relationship to the former self?

I must confess my initial hesitancy when faced with these kinds of metaphysical questions: a hesitancy that probably stems from the fact that this simply isn't the way the self has traditionally been discussed in phenomenology. But here is my, perhaps surprising, reply. I find the idea that a stream of consciousness might start off as mine and end up being somebody else's radically counter-intuitive (cf. Dainton 2008, 18). Moreover, the moment one insists that the stream of consciousness is made up of a plurality of ontologically distinct (but qualitatively similar) short-term selves, one is inevitably confronted with the question regarding their relationship. I don't see any real alternative to the following proposal: their relationship is akin to the relationship between my self and the self of somebody else. And I find this proposal absurd.

But even if similarity doesn't amount to identity, surely—some might object—we need to distinguish an account claiming that the stream of consciousness involves some form of experiential *continuity* from an account claiming that it somehow involves diachronic *identity*. My response will be

to question the relevance and significance of that distinction in the present context. To put it differently, in my view the continuity provided by the stream of consciousness, the unity provided by shared first-personal self-giveness, is sufficient for the kind of experiential self-identity that I am eager to preserve. If you find this insufficient, I think you are looking for the wrong kind of identity.⁷

References

- Albahari, M. (2006), *Analytical Buddhism: The Two-Tiered Illusion of Self* (New York: Palgrave Macmillan).
- (2009), ‘Witness-Consciousness: Its Definition, Appearance and Reality’, *Journal of Consciousness Studies* 16/1: 62–84.
- Bodhi, Bhikkhu, editor (1993), *A Comprehensive Manual of Abhidharma* (Seattle, WA: Buddhist Publication Society).
- Campbell, J. (1994), *Past, Space, and Self* (Cambridge, MA: MIT Press).
- Dainton, B. (2008), *The Phenomenal Self* (Oxford: Oxford University Press).
- Dennett, Daniel Clement. 1992. ‘The Self as the Center of Narrative Gravity’, in Frank S. Kessel, Pamela M. Cole, and Dale L. Johnson (eds.), *Self and Consciousness: Multiple Perspectives* (Hillsdale, NJ: Erlbaum).
- Ganeri, J. (2007), *The Concealed Art of the Soul: Theories of Self and Practices of Truth in Indian Ethics and Epistemology* (Oxford: Oxford University Press).
- Heidegger, M. (1993), *Grundprobleme der Phänomenologie (1919/1920)*. Gesamtausgabe Band 58 (Frankfurt am Main: Vittorio Klostermann).
- Henry, M. (1963), *L’essence de la manifestation* (Paris: PUF).
- (1965), *Philosophie et phénoménologie du corps* (Paris: PUF).
- Hobson, R. P. (2002), *The Cradle of Thought* (London: Macmillan).

⁷ One might here mention Ricoeur’s careful distinction between two concepts of identity: Identity as sameness (*mémeté*) and identity as selfhood (*ipséité*) (Ricoeur 1990). The first concept of identity, the identity of the same (Latin: *idem*), conceives of the identical as that which can be re-identified again and again, as that which resists change. The identity in question is that of an unchangeable substance, or substrate, that remains the same over time. By contrast, the second concept of identity, the identity of the self (Latin: *ipse*), has on Ricoeur’s account very little to do with the persistence of some unchanging personality core. Whereas questions regarding the first concept of identity take the form of What questions, questions regarding the second concept take the form of Who questions, and must be approached from the first-person perspective.

- Husserl, E. (1954), *Die Krisis der europäischen Wissenschaften und die transzendente Phänomenologie: Eine Einleitung in die phänomenologische Philosophie*, Husserliana 6 (Den Haag: Martinus Nijhoff).
- (1966), *Analysen zur passiven Synthesis: Aus Vorlesungs- und Forschungsmanuskripten 1918–1926*, Husserliana 11 (Den Haag: Martinus Nijhoff).
- (1973a), *Zur Phänomenologie der Intersubjektivität: Texte aus dem Nachlass. Erster Teil: 1905–1920*, Husserliana 13 (Den Haag: Martinus Nijhoff).
- (1973b), *Zur Phänomenologie der Intersubjektivität: Texte aus dem Nachlass. Dritter Teil: 1929–1935*, Husserliana 15 (Den Haag: Martinus Nijhoff).
- James, W. (1890), *The Principles of Psychology I* (London: Macmillan and Co.).
- Lusthaus, D. (2002), *Buddhist Phenomenology: A Philosophical Investigation of Yogacara Buddhism and the Ch'eng Wei-shih Lun* (London: Routledge).
- Merleau-Ponty, M. (2002), *Phenomenology of Perception* (London: Routledge).
- Metzinger, T. (2003), *Being No One* (Cambridge, MA: MIT Press).
- Neisser, U. (1988), 'Five Kinds of Self-knowledge', *Philosophical Psychology* 1/1: 35–59.
- Reddy, V. (2008), *How Infants Know Minds* (Cambridge, MA.: Harvard University Press).
- Ricoeur, P. (1990), *Soi-même comme un autre* (Paris: Seuil).
- Rochat, P. (2001), *The Infant's World* (Cambridge, MA: Harvard University Press).
- Sartre, J. (1967), 'Consciousness of Self and Knowledge of Self', in N. Lawrence and D. O'Connor (eds.), *Readings in Existential Phenomenology* (Englewood Cliffs, N.J.: Prentice-Hall).
- (2003), *Being and Nothingness* (London: Routledge).
- Seigel, J. (2005), *The Idea of the Self: Thought and Experience in Western Europe Since the Seventeenth Century* (Cambridge: Cambridge University Press).
- Sorabji, R. (2006), *Self: Ancient and Modern Insights about Individuality, Life and Death* (Oxford: Clarendon Press).
- Stern, D. N. (1985), *The Interpersonal World of the Infant* (New York: Basic Books).
- Strawson, G. (2000), 'The phenomenology and ontology of the self', in D. Zahavi (ed.), *Exploring the Self* (Amsterdam: John Benjamins).
- Varela, F. and Shear, J. (1999), *The View from Within: First-Person Approaches to the Study of Consciousness* (Thorverton: Imprint academic).
- Wilde, O. (1969), *De Profundis* (London: Dawsons).
- Zahavi, D. (1999), *Self-Awareness and Alterity: A Phenomenological Investigation* (Evanston, IL: Northwestern University Press).
- (2003) 'Inner Time-Consciousness and Pre-reflective Self-Awareness', in D. Welton (ed.), *The New Husserl: A Critical Reader* (Bloomington, IN: Indiana University Press).

78 DAN ZAHAVI

- Zahavi, D. (2004), 'Time and Consciousness in the Bernau Manuscripts', *Husserl Studies* 20/2: 99–118.
- (2005), *Subjectivity and Selfhood: Investigating the First-Person Perspective* (Cambridge, MA: The MIT Press).
- (2007a), 'Perception of Duration Presupposes Duration of Perception—or Does it? Husserl and Dainton on Time', *International Journal of Philosophical Studies* 15/3: 453–471.
- (2007b), 'Self and Other: The Limits of Narrative Understanding', in D. D. Hutto (ed.), *Narrative and Understanding Persons* (Cambridge: Cambridge University Press).
- (2009), 'Is the Self a Social Construct?' *Inquiry* 52/6: 551–573.