

The Self restated

Jonardon Ganeri¹

Published online: 22 December 2016

© Springer Science+Business Media Dordrecht (outside the USA) 2016

Abstract This is a short summary of the book *The Self: Naturalism, Consciousness and the First-Person Stance* (Oxford University Press 2012). It introduced an “author meets critics” panel at the American Philosophical Association Pacific Division meeting in San Francisco 2016. I try to relate the discussion in the book to recent work that has appeared since its publication.

Keywords Self · India · Buddhist · Nyāya

“The function of consciousness,” Brian O’Shaughnessy observes in his marvelous book *Consciousness and the World*, “must be to link us attentively to the physical world that contains us...Attention ultimately functions as a sort of life-blood for a whole range of mental phenomena; or perhaps better expressed, as a kind of psychic space” (2002: 84, 277). O’Shaughnessy is sceptical about the role of the concept *self*, or, as he puts it, “the mythical S,” understood as something “that precedes and outlives its occupants” (2002: 288). His scepticism is echoed by John McDowell, who in his contribution to the volume *Mind, Reason, and Being-in-the-World*, cautions that “we should not pretend to find a detached self in all our experiencing and acting...we should discard the idea that mindedness implies the presence of a detached self” (2013: 41). When early Buddhists say that “here in this world there is no self that is something other than and apart from the mental factors” (Pm. 830 cited in Nāṇamoli 1991: 847; cf. Vasubandhu 1973: 1208, discussed in Ganeri 2007: 162–163) what is being rejected is precisely this notion of a self *detached* from

✉ Jonardon Ganeri
jonardon.ganeri@nyu.edu

¹ Faculty of Arts and Science, New York University, 19 Washington Square North, New York, NY 10011, USA

experiencing and acting. In *Self and Subjectivity*, Dan Zahavi already pointed out, however, that “one does not need to conceive of the self as something standing apart from or above experiences, nor does one need to conceive of the relation between the self and experience as an external relation of ownership... The self is not something that stands opposed to the stream of consciousness, but is, rather, immersed in conscious life” (Zahavi 2005: 125). In his magisterial new book, *Self and Other* (2014), he says that what he now calls “the experiential self” is the “for-me-ness,” the “dative of manifestation.” As Joseph Levine has put it, “The very phrase that serves to canonically express the notion of the phenomenal—“what it’s like for x to...”—explicitly refers to the phenomenal state being “for” the subject... Phenomenal states/properties are not merely instantiated in the subject, but are experienced by the subject. Experience is more than mere instantiation, and part of what that “more” involves is some kind of access” (Levine 2007: 514). We can call the access which is implicit in the very idea of the phenomenal *awareness-access*. Ned Block only echoes Zahavi when he suggests that it can be explained in terms of a phenomenal property like “me-ishness” (2007: 536).

In *Self and Other*, though, Zahavi acknowledges that the experiential self “is quite a minimal notion and that it is unable to accommodate or capture all ordinary senses of the term ‘self’” (2014: 49). He considers a criticism of his earlier view, “one that, like [him], embraces realism about self and which would deny the austere metaphysics of the Buddhist critics, but which emphasizes the normative rather than the experiential basis of selfhood” (2014: 51). There are, on the one hand, thinkers like Christine Korsgaard who “argue that human beings, qua rational beings, have a distinct form of identity, a norm-governed form of identity for which we are ourselves responsible,” and there are thinkers like Harry Frankfurt, according to whom thoughts and experiences are only properly one’s own insofar as one endorses them and commits oneself to them. Such an act of endorsement, Frankfurt says, “makes the desire on which he decides fully his own... The pertinent desire is no longer in any way external to him. It is not a desire that he ‘has’ merely as a subject in whose history it happens to occur, as a person may ‘have’ an involuntary spasm... even if the person is not responsible for the fact that the desire occurs, there is an important sense in which he takes responsibility for the fact of having the desire—the fact that the desire is in the fullest sense his” (Frankfurt 1988: 170). Zahavi adds that cases of mind-wandering, “thoughts that willy-nilly run through our heads, thoughts that strike us out of the blue,” though conscious events that happen to us, are not ones we have endorsed and so, in the normative sense, own (2014: 52).

So there is a normative strand in the notion of ownership, “linked to notions such as commitment, decision, responsibility, and reflection” (2014: 90). A thought is one’s own, in this new sense, if it carries evidential weight and for a state to be owned is therefore precisely for it to engage the whole of one’s being through its potential to make normative demands on any other owned state. This is not a consideration about agency, since one can consistently accept that one is the producer of an attitude or emotion but still think that one should dissociate oneself from it in the sense intended, and one can likewise give weight to attitudes which one does not have any sense of agency with respect to, for example simple perceptions. Given that the methods we need to bring about this dissociation are said

at the outset to be evidence-based and rational (Nyāya-sūtra 1.1.1), it is clear that the notion of ownership in play here is one of endorsement and commitment. Disowning a thought is breaking a commitment, specifically an entitlement to use that thought in the justification of one's beliefs or actions; likewise, if I disown a particular feeling of hurt or anger, I break any reason or right to act on it. Endorsing a belief is committing oneself to its being true, and so to resolving the question of whether it is one's own by referring to the very facts that determine if it is true or not (which is why, in Nyāya-sūtra 4.2.1–9, it is claimed that competence in the ways of gaining knowledge of the external world is the route to gaining self-knowledge).

My claim in *The Self* is that we need an account of selfhood which does not leave it inexplicable how experience can supply reasons for judgement, that is, saddle us with an unbridgeable gap between the experiential and the normative. My argument is that the self consists in the experiential relating itself to the normative, and I find myself here surprisingly in agreement with Kierkegaard, when he speaks of a relation which is the “synthesis of the infinite and the finite, of the temporal and the eternal, of freedom and necessity,” where “the self is not the relation but the relation's relating to itself” (Kierkegaard 1983). Attempts to pay heed to one factor without the other are the source of that alienation which he terms “despair”. Kierkegaard, I think, was right about the form of the proposal but wrong about the relata. What I claim is that this explanatory demand is best served by a double aspect theory of ownership. The single fact of ownership has an experiential aspect and a normative aspect: facing the space of reasons, ownership is identification as contrasted with mere occurrence; facing the space of experience, ownership is subjectivity as opposed to anonymity.

The fundamental project of *The Self* is to explore how we might understand the relationship between these two strands within the notion of ownership. Until very recently Zahavi had been content to rest his defense of the primacy of experiential self on the claim that normativity is tied with narrativity and with authorship, and that the diachronic unity of the self presupposes its synchronic unity. Yet Galen Strawson's episodic individuals are conceptually coherent transient selves; and in any case, to put the point in Sellarsian terms, an acknowledgement of a relationship that obtains within the space of reasons does not have the temporal structure of relationships in the space of causes. Zahavi has now restated his position as being that we need to provide a “multidimensional account” according to which “there is more to human selfhood than merely the experiential component. There is also an eliminable normative dimension, even if it doesn't necessarily take a narrative form. The important point is to recognize the need for a multidimensional account” (2014: 90). He now favors, he says, “a multifaceted and multi-layered model of self,” in which there are “two different approaches to self: a narrative approach, which highlights the importance of authorship, commitment, and normativity, and an experiential approach, which puts more emphasis on ownership, pre-reflective self-consciousness, and phenomenality...The question to ask is then whether we can make do with just one of them, or whether we need both accounts? Although I take the experiential approach to be the more fundamental of the two, I am happy to concede that the narrative approach, which I have treated as an example of a more

normatively oriented account, captures something important, something that might be specific to human selfhood” (2014: 90).

Yet just as no account of the present could be adequate which does not acknowledge that it is one time with a face towards the future and another towards the past (Husserlian protention and retention), so a merely multidimensional account of two approaches fails to do full justice to the fact that the how and the why of experience are intertwined, and the place of intersection is the ownership of thought. The point of calling the my view a “normative phenomenology” is that the phenomenology of subjectivity is not that of a raw “for-me-ness” but instead has the subjective character of what it is like to endorse a feeling or thought, to identify oneself with it, to participate in the normative structure it carries, and yet equally essential that this participatory endorsement is “lived”, that one is not a norm-governed machine. So a double aspect theory of ownership has implications both for the normativity of phenomenology (the dative of manifestation is not brute “for-me-ness” but how it is to identify with the normative demands of owning) and for the phenomenology of normativity (there is something it is like to identify oneself with an ideal or aspiration aside from the mere act of doing so).

Zahavi and I both want to find a naturalistically respectable version of realism about self. Here is where we part company with Kierkegaard, for Kierkegaard claims that the self-constituting synthesis of opposites must come from above, from something supernatural. In *The Self* I argue instead that subpersonal embodied mechanisms of attention, integration, monitoring and retrieval are fundamental to the binding of the experiential and normative stands within the notion of ownership. This entails that the two aspects of selfhood can come apart in pathological cases, and at the very end of the book I give an example of such double dissociation from the writings of Henri Amiel, who induced what he called “philosophical experiences” within himself. In one he retains access to an experiential self while losing access to the normative self, and he describes his resulting state as a condition of “depersonalization”, which we might think of a sort of virtual reality existence: “I find myself regarding existence as though from beyond the tomb, from another world; all is strange to me; I am, as it were, outside my own body and individuality; I am depersonalized, detached, cut adrift” (Amiel 1889: 352). In another of his “philosophical experiences” he seems to achieve a condition involving the retention of a grasp on normative endorsement but without any experiential immersion: [“My mind is the empty frame of a thousand vanished images. Sharpened by incessant training, it is all culture, but it has retained hardly anything in its meshes. It is without matter, and is only form... It is etherealized, algebraicized. Life has treated it as death treats other minds; it has already prepared it for a further metamorphosis. Since the age of sixteen onward I have been able to look at things with the eyes of a blind man recently operated upon—that is to say, I have been able to suppress in myself the results of the long education of sight” (Amiel 1889: 351).] The dissociation is that to which Nabokov gives voice when he says, “I have grown much too used to an outside view of myself, to being both painter and model, so no wonder my style is denied the blessed grace of spontaneity. Try as I may I do not succeed in getting back into my original envelope, let alone making myself comfortable in my old self” (1966: 19). With a phenomenal but no endorsing self,

one's desires and commitments would be as if that of a virtual avatar or simulacrum, experienced "from the inside" but without any normative pull or demand. With an endorsing but no phenomenal self, one would be as if afflicted by a disorder in which one's occupation of a first-person stance is devoid of phenomenal substance, and while one is not alienated from one's commitments and desires, they do not feel alive to one. Fully first-personal subjective consciousness is at once *grounded* (in "friction" with the world and subject to its constraint), *lived* (in experiential openness and presence to the world), and *engaged* (with the pulls and demands of emotion and intention on the world).

In executing this project I was fortunate to have at my disposal three bodies of philosophical writing of enormous richness and insight, all deriving from first millennial Sanskrit India. The Indian materialists, known as Cārvāka, provided me with a very helpful model of the emergence of mental properties from physical properties, which I dubbed "transformation emergence" (and I noted, in passing, that there was every possibility John Stuart Mill had read some of their writings, which were available in English and German translations as early as 1832). Buddhist philosophers belonging to the school of Abhidharma engaged in an extraordinarily sophisticated analysis of the experiential dimension of a first-person perspective, that is, of what Zahavi now calls the "dative of manifestation." I am still learning from them and indeed my new book (Ganeri 2017) is entirely about their contributions to the understanding of attention and consciousness. And finally philosophers in the twinned schools of Nyāya and Vaiśeṣika defended the self against Buddhist criticism primarily by stressing the ineliminability of the normative dimension, of commitment, responsibility, endorsement and reason in the synchronic constitution of selves. It is not as well-known as it should be that these thinkers went on to spearhead an early modernity in sixteenth and seventeenth century South Asia that was the philosophical equal of all parallel developments in Europe.

In particular, I drew on two ideas from the early Nyāya–Vaiśeṣika and the early Abhidharma Buddhist discussion, both ideas, ironically enough, denoted by their proponents using the same Sanskrit term *manas*, a cognate of 'mind' but using that term in completely different senses from each other. In the hands of Nyāya–Vaiśeṣika philosophers what this term denotes is a subpersonal comparator mechanism responsible for quite a range of subpersonal activities including selective attention to sensory stimuli and crossmodal monitoring, monitoring of internal states, and in particular, I suggested, subpersonal monitoring of desires and beliefs. The multitasking of this subpersonal mechanism for cognitive access enables it to create conditions of possibility for the unification of experience and thought.

One way it does this is through the multisensory integration of sensory contents into awareness of an objective world. A second way it achieves it is by searching for content incongruences, and I drew on Indian theories about the emotions to argue that emotions are non-cognitive signals of the existence of incongruences among beliefs and desires and other attitudinal states. According to an analysis also first advanced in detail by Praśastapāda the core attributes of an emotion are a felt appraisal, an inclination to action and a capacity to be felt. Pleasure, for example, is

elicited by the appraisal of something as likeable or pleasant, is exhibited in an inclination to maintain a continuing relationship with the object, and is felt as agreeable or gratifying or favorable. Desire is elicited by the appraisal of something not at hand as potentially so, is exhibited in an inclination to create a relationship with it, and is felt as a ‘calling out’. Emotions serve to indicate a relation of common ownership among the commitments, values, preferences and intentions of a single subject: they manifest incongruence and so enable rational revision. Our philosophers say that delight (*harṣa*) is the pleasure one has on fulfilling a longing for something desired, fear (*bhaya*) the distress when one is unable to rid oneself of the wish to flee in the presence of what will lead to something undesired, sorrow (*śoka*) the unfulfillable longing for something desired from which one is separated. Simone Weil defines joy as the feeling of reality; in our model it is the sign that one believes one’s beliefs to be true. The concept of self plays an essential role in this theory of emotions as outputs of unconscious comparator activity. The self is that single owner of the beliefs and desires among which comparisons are made: it is essential that the sure belief at t is a belief owned by the same subject as the negative belief at $t - 1$ if an emotion is to serve as a signal, that is to say as evidence of an incongruence. It is for that reason Amiel describes the unity of self as consisting in “intelligence working its way through love and pain.” It is the defining feature of a first-person stance that one occupies and endorses one’s states of mind and is not merely a spectator of them, and so it is from this stance alone that the relationships definitive of liberal naturalism, the *pulls* of reasons and the *demands* for action, come into view. I rejected a supplementary claim the early Nyāya–Vaiśeṣika thinkers make, which is that this unification of sensory contents, and of representations across time, and of the aspects of an object into a representation of a single object, leads to a conception of self as the common owner of multiple experiences that can exist in a incarnate state. I rejected this because sentience, agency, and emotion require embodiment.

The Buddhist analysis of *manas*, as articulated by Vasubandhu, is quite different. Vasubandhu’s claim is that *manas* is a pre-attentive mode of awareness which draws on a repository of semantic autobiographical information in order to introduce the phenomenal sense of mineness into experience. What it offers is thus precisely an account of the “dative of manifestation,” the phenomenal “for-me-ness” of experience. It seems to me that Vasubandhu’s analysis has an important virtue over Zahavi’s, which is that it does not treat phenomenal mineness as a primitive, but demonstrates how it is “rooted” in a first-person repository, as sort of first-personal mental file. The worry about Zahavi’s “minimal self” is that it is too minimal to display what makes one self distinct from another, and that worry is not present in Vasubandhu’s account. But Vasubandhu makes another move which I did not accept. Just as so-called dogmatists claim that perceptual experience is such as to provide prima facie reasons for judgement and belief, so one might suppose that phenomenal self-awareness should provide prima facie reasons for first-person avowals. Vasubandhu denies this, and argues even more strongly that the first-person pronoun is a device of disingenuous reference, meaning that it pretends to be a referring expression but in fact has no referent. If we reject that second move, then

we can say that he has provided materials for a very plausible analysis of phenomenal selfhood.

The view I eventually defend, then, is this: The concept *self* is the concept of something necessarily embodied, something whose existence is made sense of by the idea of unconscious mechanisms of comparison, monitoring, and information-retrieval, the states of which supervene on and indeed emerge from, mechanisms the character of which is determined by the very fact that they support the emergence of consciousness, underwrite the phenomenology of mineness, and explain the congruences and mismatches responsible for emotion and common ownership. To be self-conscious (in the intransitive sense) is to be occupying in the present moment a first-person stance, and since this implies the capacity of any one of one's owned states to engage the whole of one's being through its potential to make normative demands on any other—and so for that state to be “rooted” in the self—it entails that the single comparing, monitoring, selecting subroutine which is the unconscious mind or underself is in a state of characteristic activity, distinct from the more modularised state it is in when one is not conscious. The self is embodied because to *own* a mental state (that is to say, to inhabit and endorse it) is to be responsive in principle to its demands for action: that is why a discarnate state of being cannot be one in which there are any owned states of mind.

References

- Amiel, H. F. (1889). *Amiel's Journal*, trans. H. Ward (New York: A. L. Burt), 2nd edn.
- Block, N. (2007). Overflow, access, and attention. *Behavioural and Brain Sciences*, 30, 530–548.
- Frankfurt, H. (1988). *The importance of what we care about: Philosophical essays*. Cambridge: Cambridge University Press.
- 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.
- Ganeri, J. (2017). *Attention, not self*. Oxford: Oxford University Press.
- Kierkegaard, S. (1983). *The sickness unto death*. Princeton, NJ: Princeton University Press.
- Levine, J. (2007). Two kinds of access. *Behavioral and Brain Sciences*, 30(5–6), 514–515.
- McDowell, J. (2013). The myth of the mind as detached. In J. K. Shear (Ed.), *Mind, reason, and being-in-the world: The McDowell-Dreyfus debate* (pp. 41–58). London: Routledge.
- Nabokov, V. (1966). *Despair: A novel*. New York: G. P. Putnam's Sons.
- Ñānamoli, B. [Osbert Moore]. (1991). “Notes”. In *The path of purification: Visuddhimagga by Bhaddantācariya Buddhaghosa* (5th Ed., pp. 745–907). Kandy: Buddhist Publication Society.
- Nyāya-sūtra*. (2007). In A. Thakur (Ed.), *Gautamīya-nyāya-darśana with Bhāṣya of Vātsyāyana*. Delhi: Indian Council of Philosophical Research.
- O'Shaughnessy, B. (2002). *Consciousness and the world*. Oxford: Oxford University Press.
- Vasubandhu. (1973). Abhidharma-kośa with Abhidharma-kośa-bhāṣya. In D. Shastri (Ed.), *Yaśomitra's Sphuṭārtha*. Varanasi: Bauddha Bharati.
- Zahavi, D. (2005). *Subjectivity and selfhood: Investigating the first-person perspective*. New York: MIT Press.
- Zahavi, D. (2014). *Self and other*. Oxford: Oxford University Press.