

Book Review

Self to Self: Selected Essays, by J. David Velleman. Cambridge: Cambridge University Press, 2006. Pp. xiv + 385. H/b £64.00.

Good books, however abstract or remote their topic might be, somehow manage to make you think about yourself. Velleman's book should make you think about the very notion of the 'self'; and one of the things it might convince you of is that the self is not a thing—at least not if by 'thing' we mean some one entity that forms an essential part of a person's being. The primary goal, though, of Velleman's project is not negative; he is engaged in a meticulous critique of various ideas about the self, as a means of articulating a novel account of some important phenomena, including personal identity, autonomous agency, the experience of self-directed emotions, and the ability to judge—morally or otherwise—oneself.

Self to Self is a collection that combines virtues rarely exemplified nowadays in texts by a single author. On the one hand, it is an analytically-minded, systematic work, putting forward arguments for substantial claims about hot topics of current research. On the other hand, several of the essays read like running commentaries of other philosophers' work, both classic and contemporary, with each attribution of views scrupulously refereed, and each proposed interpretation properly grounded in knowledge of the original texts. As a result, the critical bar is set very high; whether or not the reader agrees with Velleman's conclusions, he will gain a lot from studying these essays, so as to explore the different aspects of a justly ambitious philosophical project.

What runs through Velleman's wide-ranging discussion is the apparently simple but explanatorily powerful thesis that 'the "self" merely expresses a reflexive mode or modes of presentation' (p. 1). Velleman distinguishes between three major modes of self-presentation, or, as he calls them 'three reflexive guises under which a person tends to regard aspects of himself' (p. 3): those of self-image, of self-sameness through time, and of self as a cause of autonomous action. The present review will focus on the second of those reflexive guises, that of sameness of oneself through time and on Velleman's concomitant views about past and future selves. I shall only make here a couple of remarks about the other two kinds of self-presentation.

A self-image is that 'by which a person represents which person and what kind of person he is'; it is an open set that might include a person's 'name,

address, and Social Security number, how he looks, what he believes in, what his personality is like, and so on' (p. 3). That is undeniably an important aspect of selfhood: it comprises elements by which a person identifies which of the world's inhabitants he is; it functions as a repository of one's evaluations of oneself; and it marks the point of reference for those attitudes by which we compare ourselves to others, and understand and possibly share their emotive and evaluative view of ourselves (p. 4).

What I find less apposite is the choice of terms Velleman employs for capturing the relevant phenomena. From the six items listed in Velleman's set, only the fourth, about 'how a person looks', has an imagistic content; the other five pick up numerical or other information of a propositional nature. And unless one assumes that mental content is not available in the absence of associated imagery, it is not clear why the kind of identificatory content Velleman wishes to capture is best conveyed by the notion of 'self-image'.

Velleman is perhaps not unaware of this difficulty, and thus it might not be an oversight that, a few pages down the text, he refers to exactly the same phenomena employing the different notion of 'self-concept' (p. 8). That term, though, is also not without its problems. A self-concept is something philosophers define or devise when they engage in conceptual analysis, whereas what Velleman purports to convey is the totality of characteristics, ideas, or facts through which a person is 'conceiving of who he is' (p. 4). Thus, I would suggest that the term 'self-conception' might be closer to the mark.

The other guise under which a person is presented to himself is as an autonomous agent. For Velleman, the locus of autonomous agency is that aspect of the person that is presented to him when his reasoning about the causes of his actions attributes their cause to itself: 'Insofar as a person's behavior is due to his causal understanding, its causes will appear to that understanding in reflexive guise, and the behaviour will appear as due to the self' (p. 7). Nearly half of the essays in the present volume (especially from Ch. 9 through to Ch. 14) bear upon this issue, and like all of Velleman's writings on agency, are required reading for anyone interested in the topic.

It might be worth stressing that Velleman's notion of 'autonomous agency', although it is offered as an elaboration on Kant's well-known theory, might not correspond to most people's idea of that phenomenon—especially if they approach autonomy from a Kantian perspective. In explicating, for instance, the idea of 'acting for a reason', Velleman begins by rightly pointing to a sharp distinction between autonomous and heteronomous behaviour:

Our behavior is autonomous when it is self-governed...it is heteronomous...when it is controlled by something other than ourselves.... A sneeze or a hiccup is not under our control; neither is a startle or an impulsive cry of pain; but all of these heteronomous behaviors originate within us. What makes them heteronomous is that... they don't originate with us: they aren't fully our doing. (p. 33)

I agree that a sneeze or a hiccup, is not an example of autonomous action; but that is not because it is heteronomous, but because it is not an action.

In general, Velleman appears to exclude heteronomy from the domain of reason-guided behaviour. ‘Acting for reasons makes us autonomous’ (p. 34) Velleman claims, and he intends that claim as a presentation of Kant’s own thesis. Recall, though, that when Kant introduces the principle of autonomy, he contrasts it to that of heteronomy as pertaining to actions aiming to satisfy an interest. However, acting out of ‘one’s own or another’s interest’ (*Groundwork* 4: 432) is not acting for no reason — and even less is it ‘not acting’ at all. Indeed, treating heteronomous behaviour — whereby one pursues, for instance, one’s own happiness — as behaviour that falls short of rational agency, would be rather peculiar for a philosopher who presents the hypothetical imperative as a practical rule of reason, which may ‘determine the conditions of the causality of a rational being as efficient cause ... with respect to the effect [of the action]’ (*Critique of Practical Reason* 5:20).

Issues in Kantian exegesis are far beyond the scope of this short notice. All I mean to point out is that it is not as obvious as Velleman’s approach appears to imply, that a Kantian is bound to relegate heteronomous behaviours to ‘mere happenings, or mere activities’, as opposed to ‘actions’ (the tripartite distinction comes from Velleman’s *The Possibility of Practical Reason*, Oxford: Oxford University Press, 2000, p. 4).

The other mode of reflexive presentation concerns the diachronic connections between selves. Self-sameness through time is the relation that connects a person to his past selves (which are past persons given, through experiential memory, in reflexive guise) and to his future selves (which are future persons, given in anticipation, under a reflexive mode of presentation): ‘Past and future selves are simply past and future persons whom the subject can represent as the “I” of a memory or the “I” of a plan — persons of whom he can think reflexively, as me.’ (p. 5).

The notion of self-sameness is intimately related to that of personal identity. In the title essay of the volume, Velleman offers a masterful analysis of what makes a person ‘self to himself’, as Locke famously put it. Locke, of course, was referring to the role of memory in the constitution of personal identity. Velleman wishes to untie ‘selfhood’ from ‘personal identity’, so as to gain a better grasp of what precisely each of those ideas amounts to, and, hence to avoid (what he considers as Locke’s) mistaking selfhood, which is ‘the relation borne to me by those whom I can think of first-personally’, for personal identity, which is ‘the relation among those who are one and the same person’ (p. 193).

Despite the enviable clarity of its prose, Velleman’s essay makes for dense reading, as it touches upon a large number of issues — not all of which are explicitly thematized: in methodology (what are the right questions to ask in this area, and what should count as a credible answer); in phenomenology (how best to describe an ordinary experience of sense-imaging or

propositional imagining); in epistemology (which of the items described are due to the experience as such, which are superimposed as we switch among various perceptual, theoretical, or practical perspectives, and how does this affect the veridicality of the corresponding accounts); in philosophy of language (concerning the function of indexicals or the expression of self-locating beliefs). Add to this list of issues some of the standard puzzles one encounters in the literature on personal identity, and one might get a sense of how rich Velleman's agenda is in that essay. Here, I will cite a few paragraphs, followed by brief remarks on some important points in Velleman's argumentation.

The first point concerns the future. Velleman begins by articulating a model of imagining being someone else—for instance, Napoleon Bonaparte viewing the desolate battlefield at Austerlitz—a model that he then applies, subject to adjustments, to intending, planning, or anticipating an event of my future selves:

When I think of [an] image as having a subject, it becomes a way of thinking about the person reflexively, as 'self'. And to think of a person reflexively, as 'self' is also to think of him as 'me'. If I think of the image as having a particular subject, such as Napoleon, the image becomes a way of thinking about Napoleon as 'me', and so it becomes a way of thinking that I am Napoleon. (p. 180)

That sequence of claims is attractive, but as a model for other mental acts might prove misleading. In ordinary cases of deliberation, or anticipation, it is the other way round: first there is me—in the sense that there is no question as to who might be the agent I am thinking about when I deliberate, plan, or anticipate what I will do; and then comes the issue of how things will be, feel, or look to me, from the standpoint that I will occupy. In Velleman's account, on the contrary, first comes the future standpoint, and then my self shows up as its occupier; indeed, that latter move is delivered in two stages, first as an answer to the question whether that future perspective is occupied by someone—and thus that we are not dealing with a case of a mere visualization but with the more robust project of imagined seeing (p. 169)—and secondly, whether that someone is me. Apart from sounding false in its phenomenology, that account faces the challenge of showing how that last question can receive a satisfactory answer. Judging that 'I am that future someone', cannot be a matching exercise, that would map an 'I'-description to a description of that future person, simply because there is no appropriate 'I'-description to begin with. (Here is how Simon Blackburn puts this thought: 'We are familiar with the idea that "I" cannot be substituted by any battery of descriptions, for that leaves open the question whether it is I who satisfies the descriptions. My parallel point is that when I reason about the future, the "I" cannot be substituted by something with some favorite battery of links to my present self, for that leaves open the question whether it is I who experiences one thing or another at the end of the links.' This passage is from p. 197 of Blackburn's 'Has Kant Refuted Parfit?' in *Reading Parfit*, ed. Jonathan Dancy, Oxford: Blackwell, 1997, pp. 180–202.)

A second point concerns the past:

If the image is indeed a copy of a visual impression ... then there is already a fact of the matter as to the identity of its notional subject: he is the person from whose experience the image was copied ... the image is just presented to me as having been copied from a visual impression, and it consequently represents things as seen by the subject of the impression from which it was, in fact, copied. Who he was is then determined by the image's causal history. This mechanism makes the reflexivity of my memory genuine ... in a way that the reflexivity of imaginings is not. In memory I really think of the notional subject as 'me'; in imagination, I only pretend to. (p. 188)

Again, Velleman's conclusion is correct but the way he reaches it is, to my mind, questionable. In remembering as opposed to imagining something, I of course think of the subject of the relevant memory content as 'me': but that is not because in the two cases I have applied the same test, and it gave 'me' in one case and 'not-me' in the other — in the case of remembering, at least, I do not go through any test at all. I do not claim that memory is infallible. What I assert is that if something is experienced as an instance of remembering, I may have doubts about the accuracy, the completeness, or the veridicality of the memory content; but it is not an issue for me whose memory I have, or whose past experience I try to recall.

Velleman appears to be exercised by precisely these kind of problems, in the following paragraph (p. 188):

Before I can frame an image that points to Napoleon at its center ... I must first frame another thought that picks him out, so that I can center the image on him ... I couldn't have picked out NB as 'me' without first picking him out as 'Napoleon', in order to stipulate that he was the notional subject of thought ...

On the contrary, Velleman continues

Genuinely reflexive thoughts don't rely on an antecedent specification of their target: they just point to the subject, at the center of thought. They are — to put it somewhat paradoxically — unselfconscious about their reference, in that they require no other thought about whom they refer to. (p. 188)

That is a crucial disanalogy between imagining being someone else, on the one hand, and remembering or anticipating events in one's life, on the other, and Velleman is right to stress that experiential memory, as well as anticipation involve 'genuinely reflexive thoughts'. Is he correct, though, in attributing that process to an unselfconscious pointing of reflexive thoughts to their reference? It depends on what precisely that means.

Recall, first, that Velleman is adamant in dissociating self-sameness through time from personal identity, even if the two might come together in some, perhaps most, ordinary cases. However, if 'the reference' of reflexive thoughts is a person other than my (present) self, calling the latching onto that person 'unselfconscious', sounds more like a title for what needs explaining, than a proper explanation of it. If, on the other hand, the one to whom I am referring is my (present) self there is no pressure to invoke

unselfconscious processes: my thinking about me as myself is as self-conscious as one may wish.

The source of this puzzle lies perhaps in a thesis offered at an early stage of the argument. Velleman contends that when I imagine that I am X, 'I am not imagining anything about my actual self' (p. 171), and hence — according to Velleman — in statements of the type 'I imagine that I am X', the latter 'I' refers to X, and not to me. That is a phenomenological claim I personally find difficult to absorb. I would maintain that when I imagine that I am Napoleon viewing the battlefield at Austerlitz, I put, so to speak, myself in Napoleon's boots looking out towards the desolate field. Hence, I find it hard to understand the idea that it is not myself that I put in that place, but someone else, and yet call that a mental act of imagining myself being someone else viewing a field.

The point about how best to describe that imagining act is significant, because on its ground Velleman will gradually erect his elaborate theory of how I supposedly access my 'past and future selves', that is by entertaining thoughts that 'just point to' their subject, and so are 'genuinely reflexive'. I have not, to be sure, presented any knock-down argument against that view, at most I have expressed a worry as to the explanatory power of Velleman's account — and it is perhaps nothing but a lack of imagination on my part that makes me fail to see the rightness in Velleman's starting point.

Let me pass to a quote that relates the future to the past: 'Just as a memory purports to represent the past from the perspective from which it originated in experience, so an intention purports to represent the future from the perspective at which it will arrive to guide action.' (p. 198). At a general level, that statement is unobjectionable. If it were set, though, as the mould on which all self-implicating, future-directed mental acts are to be modelled, it might prove misleading. That is especially so when we consider the standard psychological attitudes that are examined in the literature, and which are illuminatingly explored by Velleman, such as one's concern about one's own future. The problem, in my view, is that many of our attitudes about the future do not *ipso facto* require the specification of the perspective occupied by a future person, who will be our 'grammatical person-mate' as Velleman puts it (p. 193).

To be sure, a concern about sameness of myself through time may involve a concern about my future. Yet, that does not mean that I now have an articulate conception of what that means (what precisely a reference to my future includes), and, thus, that my concern about it (about my future), is a concern about whether that future, as now is supposedly envisioned, will be occupied by me. I may like to continue being around, without having a definite view for how long ('to infinity?', 'under any circumstances?', 'no matter what?'), or in what particular shape or form I would like myself to be. Moreover, it is not clear to me that it is a conceptual requirement for one's desiring oneself to be in the future, that one has another desire about

the particular things one will be engaged with in the future, to which the former desire is a means to that end. I may want to continue to exist, without wanting that, just so that another want of mine be realized. Being around is not something people do, or that, literally speaking, they may plan to do, so that some other thing gets done. Rather, being themselves around is a pre-condition for them themselves doing anything at all.

I have tried to voice some concerns and raise some queries about Velleman's analysis, which provides one of the best worked-out theories of one's relation to one's own past and future. I will close with a short quotation, and an even shorter comment: 'Not being just plain "me" to myself would be more than the loss of a pronoun; it would be the loss of self-intimacy that is part of what matters about having future selves' (p. 202). That is, I believe, a very important claim to which we should thoroughly subscribe—except, perhaps, for the last couple of words: what on earth is that talk about 'future selves'? Under normal circumstances, when I worry, hope, or deliberate, I do not do so with regard to my future selves (in the plural); not even about my future self (in the singular). Rather, what I think, hope, or worry about is: myself in the future. But what exactly that means, is still an issue that philosophers need to ponder.

Department of History and Philosophy of Science
University of Athens
 Athens 15771
 Greece
 hatzimoyisis@yahoo.com
 doi:10.1093/mind/fzs116

ANTHONY HATZIMOYISIS