

Book review

Tilo Kircher and Anthony David, *The Self in Neuroscience and Psychiatry*,
Cambridge University Press, 2003, Cambridge

This volume on the topic of the self offers a diversity of points of view, and it is the merit of the editors to welcome contributions from such diverse domains as the cognitive sciences, philosophy, theory and clinics of psychiatry and the neurosciences. Such an attitude fits the recent trend of research in a number of topics (mind-body, embodiment, emotions, intentionality, etc.), which require a multi-perspectival approach. The introduction lists the disciplines that have come to join the *philosophical* topic of the self: social sciences, cognitive sciences, the neurosciences and clinical sciences. It is to be expected that this does not only happen for the phenomenon of (self) consciousness, but also in the other domains mentioned above. Notwithstanding the diversity, the reader is kept from drowning, because, in general, each author is clearly situated within a particular domain of research and has succeeded in writing a clear and conveniently arranged chapter. In the exceptional case, where the author tries to cover too much ground, the quality of the contribution suffers from some vagueness and too many unwarranted assumptions.

The cognitive sciences are underrepresented in this volume, yet this is no shortcoming. In their classical guise, cognitive scientists are in the first place model-builders, making their designs based on computer simulations and experimental data that support a rather narrow theory regarding the self. Actually, there is only one explicitly cognitive-scientific contribution in the book. It is the one by Barnard, and in my view the least comprehensible chapter of the book. We find an all-encompassing model that wants to explain everything, although references to neurology are painfully lacking. It forms an all-round explanation in which little is said about the self; it explains many pathological symptoms while putting a whole number of syndromes on the same explanatory pile. The author focuses on the asynchronous processing exchanges between two levels of meanings, by which a whole range of pathological symptoms are explained. Clinical data are accounted for, but at the same time ignored because they are all put on the same footing. In fact, it is *interdisciplinarity* itself which is ignored here, by incorporating in a superficial way the findings of other disciplines.

This does not mean that interdisciplinarity involves abandoning specialization. It *does* mean, however, taking a particular attitude, in which the limits of one's own field of research are recognized and in which one attempts to use the confrontation with other perspectives in a positive way. Such a thing does not necessarily mean that we need to incorporate the other or alternative findings and points of view into our own perspective, but it does permit one's own presuppositions potentially to come to stand in another light.

In the contribution by Northoff and Heinzl three perspectives (first, second and third person perspective) on the self are reviewed. Inevitably, this involves an appreciation of the three corresponding disciplines of philosophy, psychotherapy/psychiatry and neuroscience. The authors tell us that theories are never absolutely true, but true only according to a particular perspective. We can easily agree with that. Unfortunately, this says nothing about the fact that in practice a certain perspective can *dominate* over other perspectives. Implicitly, however, the authors say that the first-person perspective is the 'most pure' one, because it allows access to the 'raw feeling' of experience. It is peculiar that - if one recognizes that truth is only valid due to and relative to a perspective, and thus constituted by a perspective - that the first-person perspective is *not* recognized in its constitutive power, and that the *experience* of direct access is considered an access to *pure* experience. One could go deeper into this myth of pure experience, but more important is what the author does with the issue of multi-disciplinarity. He explicitly recognizes that each perspective has its abilities and limitations. The consequences for the view on the self are indicated, but are all implicitly found to be legitimate, as no perspective is questioned regarding its value, adequacy or presuppositions. "This means that there is no superiority or inferiority among neuroscience, philosophy and psychiatry. All their theories of the self have a certain justification according to their perspective and also a certain limitation according to their perspective." (p. 53) That is only partially true, and it is not valid if the perspective itself is questioned. We do not plead for considering perspectives as inferior or superior, but we do plead for a questioning of the perspective itself. The reason for that is that a perspective has not merely epistemological but also ontological commitments. However, the authors claim that "(...) instead of discussing ontological considerations, the discussion should rather focus on linkages and relations between different perspectives and thus the different meanings of self in order to provide an integrative transperspectival theory of the self allowing for transdisciplinarity" (p. 53) That reveals their weak interpretation of the notion of perspective is at most an epistemological limitation/ ability. In contrast to that, we want to argue that a perspective also installs a particular ontology, which the authors deny in order to adopt a kind of integrating or integrated perspective. If they really considered the powers, and the limits, of a perspective, would they then maintain such an ontological armistice? The belief in *transdisciplinarity* expresses the desire to escape the

difficulties which an *interdisciplinary* approach inevitably encounters, due to a confrontation of perspectives. Such a confrontation is, however, not merely negative, and can lead to more profound insights and changes in position than a tepid and at the bottom impossible kind of transdisciplinarity.

Let us turn to the contribution by Berrios and Marková, which attempts an overview from a sort of zero-perspective, or a transcending perspective, of the ‘conceptual’ history of the self. From a bird’s-eye view the self is rejected as a construct, and they refuse to get involved with any of the concepts of self. But what does it mean if one claims that the self is ‘merely’ a construct that has served throughout history many different roles and functions? Rather little. “If there is one lesson to be learned from the history of the ‘disorders of the self’ before the Second World War, it was that such notions can only function properly in the medium of language: they are linguistic tropes, narrative yarns, modes of talking about people and their reasons for doing things, devices to capture meaning, even bridges over fences erected by human beings. They are *not* like stones, horses or weeds; even less should they be considered as semantically tantamount to anatomical structures or functions, not even to changes in blood flow or test scores” (p. 23) The argumentation for an elimination of the concept of the self is rather weak. Let us compare it to the history of the ultimate constituent of the world, the atom, or atomic particles. They enjoy even less than the self does a kind of substantive-like existence. But does that mean that the concept of atom or atomic particle is redundant, merely a way of speaking about something which does not nicely correlate with something observable or measurable? Has the concept of particle become superfluous? Or is it rather a construct necessary for another construction, which is the theory of physics? The criticism by Berrios and Marková is only valid for the self as a substance, not for the self as, for example, the psychical coherence of a subject.

More subtle is the argumentation by O’Brien and Opie, in their chapter on the multiplicity of consciousness and the emergence of the self, in which it is argued that consciousness is an *aggregate* of phenomenal elements. However, the distinction between the self as a unified *experience*, the functions underlying the self (e.g. subjectivity, agency) and the neural correlates are conflated in the argumentation. On p. 113, this confusion becomes apparent: “But the multitrack account is committed only to the claim that the various parts of experience are generated locally, at the very site where the information contents are fixed.” We can fully agree that there is no specific locus in the brain from which the self would rise as something monolithic. It is, however, something different to transpose *spatial* arguments (cf. locally) related to brain correlates to the phenomenon of consciousness, which is – precisely as experience – not spatial, not thing-like. Functional-spatial partitioning is no argument for a partitioning of consciousness as an experience. “The following picture emerges: pathological *losses* of phenomenology are associated with

brain lesions at specific sites, while transient but related *alterations* in experience (. . .) are found to covary with activity at these same sites” (p. 115). Yet, the model that claims a unity of consciousness is able to cope with *functional* partitioning, only if that model implies *physical* unity in the strong sense. Thus, it does not follow: “Since the contents and associated brain loci are multiple, phenomenal consciousness is multitrack.” The only thing which can be claimed is that there is no Cartesian theatre in the brain. But what does it tell us about consciousness as a unified experience? Therefore, the transition from a partitioning of consciousness to a partitioning of self-consciousness is too hasty a step. To argue that there is no Cartesian theatre is *not* the same as arguing that consciousness is multi-track. To suppose a Cartesian theatre starts from the *experience* of consciousness as unified, and to argue against a Cartesian theatre is to prevent an unwarranted transition from unified *experience* to a unified spot in the *brain*. To do the reverse, i.e. to argue for a multi-track consciousness because of an anatomical partitioning, is to commit the same mistake in the reverse sense. It is remarkable that the authors themselves have to make an appeal to a kind of unity: “[. . .] a self emerges when the multiple tracks of self-directed experience produced by the brain are sufficiently *representationally coherent*” (p. 117). Representational unity, however, does not solve the problem, but rephrases it.

Although the above remarks sound critical, the vast number of contributions are most interesting and solid. The volume is to be recommended to anyone who is interested in the self, from a philosophical, clinical or (neuro)psychological point of view. The diversity of chapters offers richness, not confusion. Zahavi’s chapter demonstrates to what extent a specialized philosopher can contribute to the debate. His contribution testifies to the importance of scholarship in the matter, and he does not choke by trying to digest too much at once. Instead, he offers a nuanced contribution to our understanding of the Kantian, hermeneutic and especially phenomenological philosophy about the self. The first-person perspective is firmly made a theme as self-givenness or self-manifestation (ipseity). That such an elaboration does not merely serve academic purposes, is shown by the implications in regard to schizophrenia and the perspectives for further research. Along with the chapters by Parnas and Sass, it forms a nice whole to read. Zahavi takes care of the theoretical-philosophical background, and Parnas and Sass show how clinical descriptions make use of phenomenological theory and help to make the experience of the patient more accessible to us. In that respect, Parnas offers and discusses case studies in a very instructive way. Sass is of particular importance for his conceptualization of hyperreflexivity in schizophrenia, in which phenomena that which would normally be inhabited or lived through transparently becomes object-like and items of focal awareness. Sass resists the view of Frith, who stresses deficits in high-level processes (i.e. self-monitoring) in schizophrenia, whereas Sass focuses on

relatively low-level processes implicated in the basic constitution of the act of awareness. A less committed contribution is made by Scharfetter, who distinguishes five basic dimensions in self-experience (ego vitality, ego activity, ego consistency/coherence, ego demarcation and ego-identity), and uses those dimensions as a means to make an empirical study of schizophrenia and other diagnostic groups. His contribution is qua theoretical insights rather modest, but certainly important for making the experiential indications operational for diagnosing a patient.

Another trio, although of more diverse theoretical background, commences with Stamenov's contribution from a linguistic perspective. He nicely points out the linguistic complexity of our use of the pronoun 'I'. His exposition is rather technical and complex, as he discusses several language-specific ways of identifying and tracking the self in conscious experience and thought. Actually, it forms a basis upon which the contributions by Philips and Gallagher, in the section on social psychology, can be read. They complement Stamenov's linguistic point of view with the concept of the narrative self. Philips introduces us into the debates of narrative theory, and after a somewhat thin philosophical start, he makes it up in what follows, namely, a set of clinical cases classified according to the kind of narrative self used. Gallagher continues where Philips stopped and adds technical aspects, distinguishing four capacities in the narrator (temporal integration, minimal self-reference, encoding and retrieving autobiographical memories, reflective metacognition). Moreover, he gives a state of the art demonstration of how narrative identity theory, the phenomenology of schizophrenia, and neuroscientific findings interrelate. Then we also have the chapter by Bentall, who attempts to assess the relationship between self, attributions (external-personal or internal-personal) for significant life events and paranoia and depression.

The reader himself has to compile such thematic bundles of articles, based on their relative coherence or complementarity, but it is well worth doing. It is, for example, interesting to see how three of the six neuroscientific chapters make up a unity. A first chapter, by Gallup et al., treats the role of the frontal cortex, self-recognition and the inference of mental states in others. Keenan et al. provides less experimental data, but stresses the connection of the prefrontal cortex with other brain structures and resists the idea of a center of the self in the right frontal lobe. Markowitsch, in his turn, stresses frontotemporal networks for auto-noetic consciousness and episodic memories. The three chapters nicely complement each other, are very accessible without simplifying and without putting wild hypotheses to the fore. It is up to the reader to make a detailed comparison and to detect the conflicting details in the overall account. To read Panksepp after the three preceding chapters is particularly agreeable. Panksepp's chapter is very balanced and with a subtle introduction. "Writing a short chapter on the neuropsychology of schizophrenia and the self is like navigating a vast ocean in a rowboat. The probability of being

swamped is much greater than a successful passage” (p. 197). Panksepp gives another account than the right-frontal-lobe approach; he stresses the often neglected motor approach to consciousness and the basic emotional circuits, the ‘gut-level’ of affective experience as decisive for the coming about and the functioning of the self. I.e. he stresses a rather low level of brain functioning in brain stem systems, instead of high cortical levels due to a view that focuses on perception. The reader is put on another track again, and complexity is re-introduced. Moreover, the reader should see connections with the contribution by Parnas and Sass, because the three contributions seem to resonate.

The part on clinical neuroscience forms the last part of the volume. Again, we have a kind of trio, formed by Jeannerod et al., Blakemore and Frith, and Fu and McGuire. The three of them are clear articles and they offer us what is to be expected from them: extended clinical-experimental situations and the results. Jeannerod explains the functioning of a ‘who-model’, responsible for the sense of agency, and Blakemore, Frith, Fu and McGuire explain further the internal-monitoring model of schizophrenia. Walter and Spitzer, however, resist the Frithian internal monitoring deficit and want to complement it with the dopamine hypothesis of schizophrenia. Vogeley also offers a complement to the self-monitoring deficit by including reality models in the explanation of misrepresentational phenomena, such as hallucinations. Kircher and David close the volume with a rather global concluding chapter, in which the optimism is striking. “We believe self-consciousness, and, in particular, the ‘feeling of selfhood’ (self-qualia) is a valid construct and its neurocognitive-emotional basis will be understood in the near future” (p. 466).

After reading the volume, the reader may not necessarily share the optimistic mood regarding progress and results. But one can certainly be optimistic about the publication of this volume, in which it becomes clear that several disciplines concentrating on the same topic need not exclude each other’s point of view. We can be optimistic that there is here a basis upon which a further dialogue can be started promising the potentiality of interdisciplinary cooperation. For anyone who only glances at the conclusion of this review: this interdisciplinary volume on the self is very readable and enjoyable. We hope the future shall bring more volumes compiled in the same spirit, but hopefully with a more firm view on the issue of interdisciplinarity itself.

HELENA DE PREESTER
Department of Philosophy
Ghent University
E-mail: Helena.DePreester@UGent.be