

Lost in Phenospace Questioning the Claims of Popular Neurophilosophy¹

Jan Slaby & Jan-Christoph Heilinger

1 Introduction

Philosophical critique can help clarify issues surrounding the burgeoning cognitive neurosciences and thereby contribute constructively to the field. 'Neurophilosophy' has become a fashionable chorus to some of the new experimental approaches to the neural foundations of human traits and capacities. Sometimes, self-declared neurophilosophers – a group that one might call the 'cheerleaders' of the new mind sciences – aspire to deliver over-arching visions of the human mind and its place in nature.² These visions almost rival older metaphysical approaches in their scope and generality. However, the impression that a broad *philosophical* conception follows directly from the *experimental results* of neuroscience is untenable. Unfortunately, quite often there is no proper distinction between what seems an obvious lesson derived from experimental findings and what is instead a philosophical assumption invested into the subject matter beforehand. Some of this is clearly the case with a recent, popular book on the subject, Thomas Metzinger's *The Ego Tunnel*.³

In the following, we will take issue with some of the claims that Metzinger makes in his book, in the hope of helping to clarify the specific relevance of new scientific approaches to the human mind. The main focus will be on claims concerning experience and intentionality, but we will also address the ethical aspirations that Metzinger expresses in the final part of his book. In a rather awkward way, Metzinger proposes a 'new approach to ethics' as a consequence of recent neuroscientific findings. Hereby, he enrolls in the camp of those commentators of neuroscience who claim a broad social relevance of the new science of the mind. Claims like these need to be viewed with some suspicion, as our arguments below will make clear.

Before we begin our discussion of Metzinger's book, we will briefly outline the motivation, aims and scope of an interdisciplinary initiative that goes by the name of 'Critical Neuroscience'. This project has been initiated and developed by graduate students and Postdocs in Cognitive Science, Philosophy, History of Science and related fields in Berlin, Montreal, and Osnabrück. It provides a framework for critical reflections like those developed in the following.⁴ The initiative brings together competences from different scientific and meta-scientific perspectives in order to help

¹An earlier version of this text that grew out of a close collaboration between the authors was published in C. Lumer/U. Meyer (eds.), *Geist und Moral. Analytische Reflexionen für Wolfgang Lenzen*, Paderborn: mentis 2011, pp. 35-53. We thank the publishing house mentis (Münster, Germany) for granting permission to re-use the material here. Part 5 below, in particular, contains ideas and arguments developed by Heilinger, who had not been listed as a co-author on the earlier version.

²Inaugurators of the movement and authors of canonical writings of neurophilosophical texts are of course Patricia and Paul Churchland; see, for instance, CHURCHLAND 1989 and CHURCHLAND 2002.

³METZINGER 2009.

⁴See CHOUDHURY and SLABY 2012.

understand, assess, and improve practices and theories in the cognitive neurosciences. Philosophical critique is a crucial part of the project's activities, but at the same time the philosophical perspective is significantly broadened and enriched by studies of the social and institutional realities of neuroscientific practice, its historical developments and anthropological contexts. Given today's widespread fascination with brain-related approaches in science and popular culture, a project like this seems to be increasingly called-for.

2 The Project of *Critical Neuroscience*

Critical Neuroscience attempts to better understand, explain, contextualise and, if necessary, criticize current developments in and around the cognitive – including the social, affective and cultural – neurosciences. One key aim of the initiative is to help create the competencies needed to deal responsibly with a range of concerns: the possible impacts of new brain-technologies; intellectual dominance of certain forms of 'neurocentric' thinking, including problematic, one-sided or outrightly mistaken philosophical theories; dangers of narrowly brain-oriented approaches to medical diagnosis and treatment to the detriment of other equally or more viable alternatives, methodological problems and limitations of technologies such as fMRI or PET, tendentious and exaggerated reports of neuroscientific findings in the popular press, and much else.⁵ The project addresses scholars in the humanities and, importantly, neuroscientific practitioners, young researchers and students.

Among the questions addressed by *Critical Neuroscience* are the following: Given the *de facto* shortage of groundbreaking results, why do the enthusiasts still so stubbornly announce that neuroscience will very soon have 'wide-ranging effects upon society'? Are we collectively overestimating its impact? Maybe the increased focus upon the brain and brain-related scientific approaches distracts attention away from other drivers of social and cultural change, such as economic or political developments – not to forget those of an increased commercialization of science and research.⁶ Thus, critics need to ask how and via what channels is neuroscience *in fact* interacting with contemporary conceptions of subjectivity, identity, and well-being?⁷ What are the dominant 'styles of thought' that have emerged in and around the neurosciences and in the new hyphenated 'neuro'-disciplines (such as neuro-economics, neuro-law, neuro-education, neuro-theology or neuro-aesthetics)?⁸ Not least: how is neuroscience institutionally and politically entangled with powerful social players such as pharmaceutical companies, funding agencies, policy makers, the emerging 'entrepreneurial University', or parts of the popular media?⁹ And why do we see concerted efforts to establish new fields such as 'neuroethics' that make a living out of alarmist, hyped-up neuro-talk?

As a further step, *Critical Neuroscience* strives to make the results of these assessments relevant to the practice of the cognitive, affective and social neuroscience themselves. What difference would it make to scientific practice if neuroscientists were

⁵A collection of essays on Critical Neuroscience, including two foundational essays as introduction to the field, has been published by Wiley-Blackwell, CHOUDHURY and SLABY 2012, see also SLABY 2010. The project is in full swing as an independent research and project group operating mostly from Berlin and Montreal. See www.critical-neuroscience.org.

⁶MIROWSKI 2011.

⁷On this, see DUMIT 2004, ROSE 2006 and JOYCE 2008.

⁸That it is no longer a style of blatant neuronal reductionism and methodological individualism is persuasively argued by PICKERSGILL 2009. See also ROSE and ABI-RACHED 2012.

⁹More on this in SLABY 2010 and in CHOUDHURY and SLABY 2012.

involved, from the outset, in the analysis of contextual factors, historical trajectories, conceptual and methodological difficulties and, not least, in assessments of the public dissemination of their results? There are several matters of concern that offer opportunities for cooperation between neuroscientists and scholars from other fields: for example, encompassing mixed-method studies of mental illnesses such as depression, the thorough investigation of social pathologies of various kinds (such as alienation in work and life environments, violence, attention problems), or ideas and popular conceptions of well-being, which are today often entangled with the soft but relentless imperatives of flexible capitalism.¹⁰ Our understanding of many of these phenomena might be significantly increased by neuroscientific findings, but it seems obvious that neuroscience cannot provide *all* that there is to know about these matters – far from it. Initial applications of the framework of *Critical Neuroscience* to empirical issues show promising results, for instance in approaches to addiction and adolescence.¹¹

In terms of its theoretical foundations, *Critical Neuroscience* draws on a variety of sources, among them Frankfurt School critical theory,¹² Bruno Latour's actor-network theory,¹³ Foucauldian approaches that analyse the entanglement of knowledge production and social power,¹⁴ and approaches to the increasing co-evolution and mutual articulation of venture capitalism and the biomedical sciences.¹⁵ The resulting approach provides a critical toolbox that helps observers to look beyond the appearances and official declarations of neuroscientific practitioners and institutions, enabling them to investigate the entanglement of research practice with broader developments within science, society, the economy and media. In line with its critical orientation, the aim is to provide the tools needed to look beyond the shiny facades of our latest technoscience and into the interests, agendas and structural factors that help shape its current outlook and direction. In its explicitly practical, hands-on orientation, *Critical Neuroscience* aims to close the gap between STS approaches and neuroscientific practice. For this reason, the project's activities are to an important part focused on teaching activities, workshops, and public events with the aim of mutual enlightenment and the creation of fruitful conflict between diverging perspectives. The present paper is broadly representative of the directions of the initiative; however it is predominantly focused on philosophical arguments and places less emphasis on the social and political climate and institutional structures that support current developments in the neurosciences. To be sure, especially Metzinger's aggressive plea for a new agenda of philosophical ethics is symptomatic of a recent trend towards a disguised pro-neuroscience 'advertising campaign' among a certain faction of humanities scholars.¹⁶

¹⁰MALABOU 2008.

¹¹For a *Critical Neuroscience* approach to addiction, see CAMPBELL 2010, for work that challenges reductionist approaches to the 'adolescent brain' while at the same time constructively engaging issues surrounding adolescence from various perspectives, see CHOUDHURY 2010. In-depth critical engagements with the neuroscience of sex differences are to be found in FINE 2010 and JORDAN-YOUNG 2010. A critique of premature neuro-appropriations by scholars in political theory in SLABY *et al.* 2012.

¹²HARTMANN 2012.

¹³See LATOUR 2005.

¹⁴See ROSE 1996, and ROSE 2006.

¹⁵SUNDER RAJAN 2006; COOPER 2008; DUMIT 2012.

¹⁶This stance has received its share of critical counter-reactions. For instance, see the recent manifesto by TALLIS 2011 and some of the essays in ORTEGA and VIDAL 2011. A still state-of-the-art philosophical critique of many neuro-inspired positions in the philosophy of mind is of course BENNETT and HACKER 2003.

3 Metzinger's Neurophilosophy of Mind and Self

In his popular science book *The Ego Tunnel – The Science of the Mind and the Myth of the Self*, Thomas Metzinger aspires to present an over-arching vision, the big picture towards which, in his opinion, the neurocognitive sciences have been driving in recent years. In the manner of the genre of popular scientific syntheses, he makes the full journey – from laboratory results, to philosophical theory, to social application and ethical reflection: neuroscience, neurophilosophy, and neuroethics are all covered. This is symptomatic of a recent trend: authors in this field seem not to be content with merely providing theoretical insights into the functioning of the brain; they are also extremely eager to spell out – and popularise – the field's alleged 'societal and ethical relevance'. So Metzinger's work is representative for the far-reaching ambitions of the neurosciences and their popularisation, particularly when it comes to determining the relevance of scientific and naturalist findings for explaining subjectivity and experience. It is for this exemplary status that we have chosen *The Ego Tunnel* as the primary subject of our critical discussion.

In the first part of this essay we will take issue with the central theoretical claim Metzinger makes. Afterwards we will have a look at Metzinger's case for a new 'consciousness ethics' as a version of 'neuro-ethics'. It is here, in Metzinger's ethical musings, that one can see a powerful new tendency at work: the push to claim social and ethical relevance, moreover a relevance of a 'revolutionary' kind, for the recent empirical and theoretical developments in neuroscience. As we hope to show, Metzinger works with a rather awkward conception of ethics, in which we find a focus on *states of consciousness* as the unit of ethical concern. This diverges significantly from common practice and introduces a number of difficulties. As regards the general trend to extend assessments of current neuroscience into the ethical and social realms, we will argue that this is a result of a gross over-estimation of the scientific and philosophical significance not only of the findings of neuroscience made so far, but also, and more importantly, of the potentials it has for future developments. While of potentially significant medical relevance, neuroscience simply does not operate on the level at which ethical guidelines are formulated or the fate of human societies is decided. It is important to correct this wrong impression as it threatens to negatively affect the self-understanding of practitioners as well as the public perception of neuroscience.

Of course, we cannot do justice to the full scope of the rich materials that Metzinger presents in his well-written and resourceful book. In particular, we have to omit the interesting discussion surrounding new research on out-of-body experiences, dreaming, and other new developments in neuroscience. Instead, we move straight to the heart of the matter, to the theory of the Ego Tunnel, Metzinger's neuroscience-inspired theory of the self and its relation to the world.

The book starts with a rather puzzling statement that immediately presents the reader with one of the key claims to be found in Metzinger's account:

In this book, I will try to convince you that there is no such thing as a self. Contrary to what most people believe, nobody has ever *been* or *had* a self. [...] to the best of our current knowledge there is no thing, no indivisible entity, that is us, neither in the brain nor in some metaphysical realm beyond this world. (ET, 1)¹⁷

Metzinger here denies the existence of 'a self', but he somehow fails to distinguish this seemingly radical claim from the rather trivial one that the self is 'no thing, no indivisible entity'. Only someone untrained in philosophy – someone unable to see the

¹⁷In all our references to Metzinger's book we will use the abbreviation 'ET' for *The Ego Tunnel*.

mistake in nominalizing the personal pronouns so as to literally assume the existence of *an entity* called 'the Self' (writ large, as it were) that is somehow distinguishable from the person as a whole – might find this claim really informative. Unfortunately, Metzinger indeed seems to play on this ambiguity, not merely for rhetorical reasons. For instance, in section 8 of his book, he again presents it as a substantial claim and as the result of his reflections, that there «is no self» and that «[w]e must face this fact: We are *self-less* Ego Machines» (ET, p. 208 – emphasis in the original). While much of the rest of what he says in the book indicates that he wants to argue for a substantial philosophical thesis, and moreover one that has the potential to shock naïve readers – all he in fact does with these initial statements is to claim that what we call 'self' is not a substantive entity. What of all those other construals of selfhood that do not depend on postulating a mysterious entity? For example, selfhood might be construed as a fluid process of reiterated agent-environment interactions in a formative social context, or as a matter of social attribution taken up and internalized by reflectively conscious agents, and so on. We suspect that it is no accident that Metzinger doesn't mention the names of any authors whom he is arguing against in this part of his theory: apart from the philosophically uninitiated or linguistically careless, there is just no opposition to the claim that what a person ultimately is, is «no little man inside the head» (ET, p. 208). In sum, this is a quite unsatisfying beginning, surely not excusable on the grounds that the book is aimed for popular readership and thus necessarily somewhat catchy and coarse-grained in style.

Metzinger goes on to outline a thorough representationalism pertaining to both world and self. In fact, most of the time, he even uses the word 'simulation' instead of 'representation', leaving no doubt about the radical nature of his thesis: «First, our brains generate a world-simulation, so perfect that we do not recognise it as an image in our minds. Then, they generate an inner image of ourselves as a whole.» (ET, p. 7). It is not just a claim to the effect that access to the world is mediated by representations, but rather the much more radical claim that what we perceive as the world and as 'ourselves' is in fact a *simulation* generated by the brain – a projection into 'phenospace', which one probably has to imagine as a kind of inner video screen.¹⁸ With this, Metzinger then elaborates upon the nature of the self-simulation thereby created:

The internal image of the person-as-a-whole is the phenomenal Ego, the "I" or "self" as it appears in conscious experience; therefore, I use the terms 'phenomenal Ego' and 'phenomenal self' interchangeably. The phenomenal Ego is [...] the content of an inner image — namely, the conscious self-model, or PSM. By placing the self-model within the world-model, a center is created. That center is what we experience as ourselves, the Ego. [...] We are not in direct contact with outside reality or with ourselves, but we do have an inner perspective. We can use the word 'I.' We live our conscious lives in the Ego Tunnel. (ET, p. 7)

The metaphor of the 'Ego Tunnel' is adopted from virtual reality engineering where we sometimes find the term 'reality tunnel'. This means that an artificial or highly selective environment is created, often in the context of computer or video games or simulation devices. So again, the point is this: according to Metzinger, the

¹⁸Again, just as in the case of the 'self', this is a substantial slip: By using the terms interchangeably and thus postulating a *de facto* equivalence between 'representations' (which might be accurate, after all) and 'simulations' (which suggest a more severe disconnection from reality), Metzinger obviously wants to operate in both registers at the same time. Theoretically rooted in the representationalist paradigm that is scientifically well-established, he tries to extract rhetorical shock value from the more radical simulation talk.

world as we perceive it is not in fact the real world, but rather a brain-generated phenospace, a highly selective projection onto a mental screen within each of us. It is quite refreshing to see the explicitness with which Metzinger formulates this position and draws consequences from it:

The conscious brain is a biological machine — a reality engine — that purports to tell us what exists and what doesn't. It is unsettling to discover that there are no colors out there in front of your eyes. The apricot-pink of the setting sun is not a property of the evening sky; it is a property of the internal model of the evening sky, a model created by your brain. The evening sky is colorless. The world is not inhabited by colored objects at all. It is just as your physics teacher in high school told you: Out there, in front of your eyes, there is just an ocean of electromagnetic radiation, a wild and raging mixture of different wavelengths. Most of them are invisible to you and can never become part of your conscious model of reality. What is really happening is that the visual system in your brain is drilling a tunnel through this inconceivably rich physical environment and in the process is painting the tunnel walls in various shades of color. *Phenomenal color. Appearance.* For your conscious eyes only. (ET, p. 20)

It is no accident that Metzinger invokes Plato's *Allegory of the Cave* at this point. According to his theory, we are in the same situation as the enchained prisoners in Plato's story, only that the cave is our mind, or rather our 'mental projection screen', and the 'shadows' dancing at the opposite wall are brain-generated simulations:

The wall [of Plato's cave] is not a two-dimensional surface but the high-dimensional phenomenal state space of human Technicolor phenomenology. Conscious experiences are full-blown mental models in the representational space opened up by the gigantic neural network in our heads—and because this space is generated by a person possessing a memory and moving forward in time, it is a tunnel. (ET, p. 23)

Everything we perceive, everything we know about, including ourselves, is in fact a part of 'virtual reality' generated by the brain.

Within the space of this essay, we can engage only one aspect of this view, but we hope to move close to the centre of the matter with it. To start things off, let us start from one rather telling phrase in one of the above quotations: «The world is just as your physics teacher in high school told you: an ocean of electromagnetic radiation, a wild and raging mixture of different wavelengths.» We will ask only one question about this: What justification is there for *prioritizing* this presumably physical conception of reality over all other conceptions, and especially above the manifest image of the world characteristic of pre-scientific common sense? More precisely, assuming the truth of Metzinger's story about brain-generated virtual reality, how could we ever *know* that this physicalist story is the true story about reality, if in fact all that we are ever in touch with is a reality simulation – *mere appearances*? If we could indeed be «brains in a vat», as Metzinger at one point happily suggests (ET, p. 21), how could we have ever managed to know with certainty what the world 'out there', outside the vat environment, is really like?

Thus, Metzinger's theory is exposed as incoherent. It places us in a position that, if we in fact were in it, we could not have any knowledge about. In this way, his story violates a Kantian condition (one that Kant, however, himself violated from time to time): that to delineate the limiting conditions of experience, we cannot move beyond an outer boundary of possible experience and draw the borderline from both sides, as it were.¹⁹ At the heart of the problem lies a kind of blatant scientism that

¹⁹*Locus classicus* for this kind of characterisation – and subsequently criticism for violating its own principle – of Kant's endeavour is of course STRAWSON 1996.

is an article of dogmatic metaphysics in disguise: it is a symptom of an exaggerated confidence in the epistemic power of the sciences. Somehow, according to Metzinger, the sciences achieve a kind of epistemic success – viz. epistemic access to reality – that by far exceeds the epistemic success that normal human subjects are ever capable of achieving. Given the fact that science obviously is a human endeavour and, as such, at any time tied to human epistemic capacities, it is highly doubtful whether science would be able to operate from such a privileged standpoint, *fundamentally* different from and superior to all and every nonscientific epistemic positionings. This might suffice as a short preview of a key argument against Metzinger's account. We will now turn to some of the relevant details.

The problem just outlined becomes most evident in the awkward theoretical status of the very organ around which Metzinger's theory revolves: the brain. Since the brain anchors the whole account as that which presumably generates all those simulations of a world and of ourselves, it has to be 'real' in the classical, non-representational sense of the term. But on the other hand, since it obviously also figures as an object of everyday perception, all Metzinger can claim is that it is also a simulation in phenospace – a factual (that is, transcendent) reality behind its mental appearance remaining unknown to us. Thus, when speaking about the brain as an object of scientific investigation, Metzinger has to assume that science has enabled us to leave the confines of our individual phenospaces and somehow access reality as it is in itself.

One can attack this thought from two sides. First, if indeed the brain as discovered by science is 'real' in the transcendent sense of the term, then it is hardly convincing that we stop there, claiming that of all we can see and perceive, *only* one single object, the brain, is 'truly real' and not just a representation, perceived as it is in itself.²⁰ Certainly, a working brain needs to be connected to a functioning body in order to function properly – but can we really say that a *real* brain can be connected to a *virtual* body? And so on – a functioning body needs oxygen and nourishment to survive and operate, so it has to be embedded in the right kind of environment; this environment, in turn, cannot just consist of simulations in phenospace because simulations are not nourishing. Thus, it has to be a *real* environment that is *really* nourishing the *real* body connected to the *real* brain. Conclusion: if the brain – as we (individually? collectively?) know it – is indeed real in the classical sense of the term, then pretty much everything else must be real in this sense as well.

Metzinger might block this rather crude line of argument by claiming, as he certainly does, that science does indeed have the capacity to discover reality as it is in itself. After all, *if* the physics teacher is indeed *right* about reality, then science, notably physics, must have succeeded in breaking free from the veil of subjective appearances, reaching out to transcendent reality. But now a number of pressing questions come up: which reality is the ultimate reality? If fundamental physics is right about wavelengths and electromagnetic radiation, isn't then the brain that current neuroscience investigates again relegated to the status of a 'mere appearance'? Wouldn't we be forced to stop taking neuroscience findings seriously until they are fully reduced to fundamental physics? – Which would be a reduction that, of course, might never happen just because of the enormous systemic complexity of brain organisation.²¹ Again, basing one's account on the reality of the brain as described by

²⁰This has been persuasively argued by LENZEN 2006.

²¹This brings us into the ambit of the debate about the autonomy of the special sciences; FODOR 1974 provides the classic kick-off argument to the debate about different disciplines operating at different scales using categories that will inevitably cross-classify reality and that are irreducible to physical categories. This thesis does not entail, to be sure, that physicalism is false from a metaphysical standpoint – it just

current neuroscience would be exposed as unfounded.

But a more fundamental question is this: how has science managed to achieve the remarkable feat of accessing 'true' reality in the first place? How have our scientists collectively managed to do what no individual and no non-scientific community ever achieved, namely, to break out of their phenospaces to access the world beyond subjective world simulations? It surely doesn't suffice to say, as Metzinger does at one point (ET, p. 8), that science is *intersubjective, communicative, and systematically organised*, thereby transcending individual perspectives. Because it seems evident enough that our pre-scientific lifeworld, our collective human practices, are also intersubjective, communicative, and systematically organised in intricate ways. Metzinger needs some factor that is *unique to science* (in a strong enough sense) that is intelligibly able to carry scientific practitioners beyond their internal representations, out of their Ego Tunnels into 'ultimate' reality.

Initially, it seems that every significant feature of science capable of achieving this epistemic feat has a corresponding feature in our pre-scientific lifeworld. Scientific experiments are an extension and refinement of meaningful human action; scientific theories are sophisticated, systematised belief systems; scientific instruments are highly refined tools; scientific observation and measurement is thoroughly dependent upon normal human observation capacities and skills; the scientific community is a well-organised subset of larger human communities; and finally, scientific rationality is not very different in principle from rationality at large, which is a human form of practice and a social institution that, to put it carefully, co-evolved with scientific practices in a long historical process. Thus, from this rather general vantage point, there seems to be no specific epistemic factor that singles out science as *radically* different.

But surely, science *is* significantly different from other human affairs in other ways. It is a relatively insulated domain of practice with its own rules and standards, it is effectively shielded – for the most part – from direct intervention by political, religious or other authorities (albeit increasingly less from commercial imperatives). The technological sophistication and the highly-specific skills and abilities of scientific practitioners in many fields also remove scientific practice from other forms of organized human conduct. Still, it is highly implausible that science could be *totally different* from the rest of human affairs in such a crucial respect as epistemic access to reality. The chief point of conflict here seems to us to be the blatant representationalist understanding of science that Metzinger presupposes. Is science's chief aim and modus operandi the generation of accurate representations of reality? This has been very thoroughly disputed.²² Science is essentially *practice* – intervening in the world, creating laboratory microworlds, designing and operating tools and technologies, attempting to construct and re-construct realities, and so on.²³ Could this endeavour so much as be conducted by beings whose life is led from the outset in a mental phenospace populated by world-simulations? At the very least, in order to establish science as we know it, one would need beings capable of engaging the world practically, acting intentionally in it, manipulating it and learning from their successes and failures. Those hands-on agents look rather different from Metzinger's Cartesian cavemen locked into their individual phenospaces.

In the end, Metzinger's conviction of a radical *representational* superiority of science

entails that physicalism has much less implications for actual scientific practice than advocates of the unity of science might suppose.

²²See HACKING 1983 for a pronounced argument to this effect.

²³See ROUSE 2002.

is exposed as a mere article of faith, unsupported by any of his own theoretical claims and thoroughly out of step with sensible theorising on the complex nature of scientific world-disclosure. The theory of the Ego Tunnel is based on an unacknowledged, ungrounded conviction that traditionally goes by the name of 'scientism'.

Almost needless to say, the actual claims of his theory are thereby exposed as utterly misguided. If one manages to see beyond the Cartesian heritage of radical mental 'simulationism', one may discover that a quite different picture of how brain, organism, and environment interact to produce human experience is currently emerging in the Cognitive Sciences. For instance, on most counts, agency is moving to the center of the picture. In many views currently debated, the mental is inextricably from agency – which of course means: various mode of engaging *the world* practically, so that a question of whether or not an agent is 'in touch' with the world does not even arise. Crucially, this alternative outlook is also supported by some findings from the neurosciences. The alternative conception goes by various names: 'embodied and embedded cognition' – 'enactivism' – 'the extended mind', and so forth.²⁴

In a book that appeared at about the same time, paralleling Metzinger's 'The Ego Tunnel' also in terms of broad accessibility and straightness of exposition, Alva Noë has outlined a version of this very different action-centred approach.²⁵ Noë explains the embodied, embedded theory of the mind – a theory that is construing the mental as thoroughly active, thereby presenting the mental agent as fundamentally 'open to the world'. This image invites us to embrace the natural-seeming thought of direct active and perceptual access to reality for all those cases where one is not manifestly misled by one's senses. Instead of being confined within one's skull and merely enjoying the brain-generated scam of an «out-of-brain-illusion»,²⁶ we can reinstate the world as by and large congruent with what it seems to us perceptually (where perception is intelligible only as itself a kind of agency). All those not under the spell of the Cartesian picture of mental simulacra on an inner presentation screen as the *only possible* mode of experience should grasp instantly that the image of an action-enabled 'openness to the world' is so natural and enthralling that it alone, without further considerations in its favour, could almost provide the basis of a *modus tollens* against Cartesian thinking.²⁷

We cannot here consider this further as our aim is to discuss Metzinger's view in its full scope, instead of exploring robust alternatives to it. Next, we will turn to a set of ideas that mark the turn, within Metzinger's book, from theoretical claims to ethical considerations.

²⁴See, for instance, VARELA *et al.* 1991; CLARK 1997; CLARK 2008; GALLAGHER 2005; THOMPSON 2007; Noë 2005, AND Noë 2009.

²⁵Noë 2009. Even the title of Noë's book tells much of the story: *Out of our Heads. Why you are not your brain and other lessons from the biology of consciousness.*

²⁶REVONSUO 2003.

²⁷An argument along these lines is developed in great detail by WILLASCHEK 2003. If philosophical considerations suggest that we are 'not in touch with the real world', then this first and foremost speaks against these philosophical considerations – simply because the common sense assumption of direct access to the real world is so overwhelmingly established in everyday practice. This style of reasoning as well as the 'image of openness to the world' have of course been famously developed and defended by McDOWELL 1994.

4 Postbiotic Ego-machines and the Ethics of Artificial Consciousness

One of the more spectacular sections of Metzinger's book is entitled «Artificial Ego-Machines» (ET, pp. 187-205). It is the section leading up to the plea for a new approach to ethics, which Metzinger calls a 'consciousness ethics'. The storyline is simple: first, Metzinger claims that it is possible, given what we currently know about the brain, to construct machines that are conscious and moreover possess a form of self-awareness; thus machines that can be said to possess what Metzinger calls an *ego* (pp. 190-95). Second, Metzinger emphatically warns us that we should do everything within our power to refrain from constructing those machines, because constructing them would likely increase the amount of suffering in the universe. Operating in a broadly utilitarian perspective, Metzinger assumes we should strive to minimise the amount of suffering in the world.

There are two quite remarkable claims here: first, that it is indeed so much as *technically possible* to construct an artificial consciousness that qualifies as a form of subjectivity. Second, that constructing conscious machines would very likely amount to constructing machines that would more or less continuously suffer. Metzinger thinks that at least the first generation of artificial Ego Machines would most probably be designed so imperfectly that, given their assumed level of cognitive and affective self-awareness, they would suffer most of the time. He illustrates the point by drawing the outrageous comparison to mentally retarded human infants (ET, p. 194). In effect, first-generation artificial Ego Machines would be as imperfectly developed as mentally retarded infants and for this reason more or less *constantly in pain* (that mentally impaired infants quite likely won't suffer constantly doesn't occur to Metzinger). An additional reason for the Ego Machine's plight is the fact that they are exploited by their human creators in the name of scientific advances, joining laboratory animals in the ranks of creatures having to suffer for the sake of human knowledge production. Given their alleged level of self-awareness, insight into this predicament adds to the suffering of the artificial Ego Machines (ET, pp. 195-6).

Since everything hinges on the first claim, we will only discuss this claim in the following. The second claim – that first-generation Ego Machines might be suffering more or less constantly – seems highly speculative to say the least.²⁸ Without knowledge of the technical details of Ego Machine construction it is impossible to assess it properly. Since we will show that the first claim is untenable, discussion of the second one becomes hypothetical anyway. Besides, the more than bizarre comparison of first-generation Ego Machines with mentally retarded children speaks volumes all by itself – we will not dwell on this any longer and let readers draw their own conclusions.

Of course, technoscientific science fiction is a perennial seller in the popular books department. Metzinger happily indulges in the sport of jumping from reports of some state-of-the-art science (bio-robotics, artificial life, and evolutionary robotics among others) to speculation about '*what is certainly possible*'. However, the gap that he thereby seamlessly attempts to close is one that is so wide that it is quite surprising that he makes these remarks almost casually and in passing, more or less hidden away in some back section of his book – instead of making them the core theme headlining his account. Artificial consciousness? What would be an outright scientific and

²⁸For instance, why couldn't it also be that rather simple minded but busy Ego Machines are the happiest beings, according to Gottfried Benn's – admittedly somewhat cynical – saying: „Dumm sein und Arbeit haben: das ist das Glück.“

philosophical sensation, by almost anyone's standards, is to Metzinger little more than a technical problem («the devil is in the details», ET, p. 189). Artificial self-awareness, genuine emotions and interests, insight into one's finitude and mortality, capacity to suffer – for Metzinger, if not outrightly 'easy', then at least only a matter of time before successfully implemented. The trick is done, not via any new theoretical insight, but by assuming that it is not a problem in principle to build (or evolve) an inclusive representation of the world – a world-model – into a machine or into a hybrid 'postbiotic' entity. Giving the right kinds of representations to the system – that is basically all we have to do. If a comprehensive representation of the outside world is in place, the decisive additional ingredient, according to Metzinger, is then just this:

If a system can integrate an equally transparent internal image of itself into this phenomenal reality, then it will appear to itself. It will become an Ego and a naive realist about whatever its self-model says it is. The phenomenal property of selfhood will be exemplified in the artificial system, and it will appear to itself not only as *being someone* but also as *being there*. It will believe in itself. (ET, p. 191-2)

For all intents and purposes, this is not *solving* the problem of consciousness and subjectivity, but moving into a state of denial. At best, it is a version of stating the task instead of showing how it can be accomplished. If Metzinger were right about this, we would face a rather grand intellectual riddle: why did generations of philosophers and scientists, among them the sharpest minds of our day, keep thinking that it is a substantial philosophical and scientific challenge to figure out how a physically implemented structure can give rise to consciousness of the world and of itself, how a subjective perspective on a world is possible within a naturalistic outlook, and how the experience of ownership of thought and sensation and authorship of action is realised in physical systems? More blatant still becomes the denial of the problems at hand when Metzinger moves from stating the problem in representationalist terms to the capacity for suffering in those self-conscious artificial systems:

Note that this transition turns the artificial system into an object of moral concern: It is now potentially able to suffer. Pain, negative emotions, and other internal states portraying parts of reality as undesirable can act as causes of suffering only if they are consciously owned. A system that does not appear to itself cannot suffer, because it has no sense of ownership. A system in which the lights are on but nobody is home would not be an object of ethical considerations; if it has a minimally conscious world model but no self-model, then we can pull the plug at any time. But an Ego Machine can suffer, because it integrates pain signals, states of emotional distress, or negative thoughts into its transparent self-model and they thus appear as someone's pain or negative feelings. (ET, p. 193)

By 'this transition', Metzinger simply refers back to the last step of his exposition, viz. the addition of an 'internal image of itself' to the artificial world-model. By this move alone, he now claims, we ensure the full implementation of the capacity for pain, negative emotions and suffering. But that is exactly the riddle: *why* is pain in fact *painful*, why does it *hurt*, indeed, why does it ultimately result in the existential predicament of *suffering*, when all there is to pain is the representation of some state of affairs as 'undesirable', that is, negative for the system? What does 'undesirable' mean here – what *can* it mean for such systems? Or to put it differently: where is the 'concernedness', the capacity to care, to have something matter to one? Without these, the idea of suffering is not so much as intelligible. The only move Metzinger has to offer is to accuse of obscurantism all those who keep expressing their puzzlement about how representing the world or processing information could give rise to a qualitative

dimension of experience.²⁹ It won't work to close a gap in understanding by denying its existence without actually showing that it is not there. Ultimately, the section on artificial Ego Machines turns out to be just another expression of Metzinger's convictions, not an argument in their favour.

5 The Call for a New Ethics

We can definitely increase our autonomy by taking control of the conscious mind-brain, exploring it in some of its deeper dimensions. This particular aspect of the new image of humankind is good news. But it is also dangerous news. Either we find a way to deal with these new neurotechnological possibilities in an intelligent and responsible manner, or we will face a series of historically unprecedented risks. That is why we need a new branch of applied ethics —consciousness ethics. We must start thinking about what we want to do with all this new knowledge — and what a good state of consciousness is in the first place.
(ET, p. 218)

Metzinger's argument for his suggested 'consciousness ethics' goes like this: Now that we are increasingly able, technologically or pharmacologically, to alter our brain states and our conscious experiences, we stand in need of a new paradigm for ethics in order to face the risks and challenges coming along with these new capabilities. With this, Metzinger aims at contributing to the heated contemporary debate about biotechnological human enhancements.³⁰ Within this broad debate he is particularly interested in the possibility of intentionally altering mental states, conscious experiences and also cognitive capabilities. Among the many examples for such 'phenotechnology' (ET, p. 222), Metzinger has decided to focus on the (experimental) use of psychoactive substances. Using such substances might have dangerous effects, he argues, because «many such experiments include the phenomenology of certainty and automatically lead to the conviction that one is *not* hallucinating» (ET, p. 220). The main problem hence seems to be that the result of using such substances blurs the distinction between what we only experience subjectively and what is truly out there, what is not a hallucination. Blurring this distinction and blending hallucinations with a feeling of certainty, will lead to, or so Metzinger seems to suggest, ourselves being mistaken about what we really are and what we really experience on the one hand, and what is only the phenomenal result of a psychotropic substance on the other. It seems questionable whether this allows to construct a specific argument against intentional phenotechnologies, given that on Metzinger's account we are nothing but Ego Tunnels anyways.

In any case, Metzinger claims that because of their impact on conscious states the emerging technologies are of both ethical and political relevance (ET, p. 222). Concretely, he calls for a new consciousness ethics to regulate which brain states and which conscious experiences we want to bring about. The key question of this novel approach to ethics, then, is this: «Which brain states should be legal? Which regions of the phenomenal-state space (if any) should be declared off-limits?» (ET, p. 229) Metzinger calls this the assessment of «the ethical value of various kinds of subjective

²⁹This accusation is implicit in the faked 'interview' with the 'first postbiotic philosopher' in part 7 of *The Ego Tunnel*, see ET, 201-05.

³⁰See HEILINGER 2010

experience *as such* [...] the rational search for a normative psychology or a normative neurophenomenology» (ET, p. 233).

In this last section of the paper we will briefly address the following questions: Are these phenotechnologies indeed novel means of intentionally altering conscious states of mind, or are they rather a gradual development from old-fashioned means with which humans have already experimented a long time? Are the basic assumptions of Metzinger's call for a 'new kind of ethics' convincing? We conclude with a remark about getting lost in phenospace.

5.1 The continuity of human enhancement interventions

There is a long tradition to use psychotropic substances in experimental and intentional ways to generate desirable mental states and conscious experiences among human beings. The (ab-) use of alcohol is only one, although prominent and wide-spread example. Clearly, there are obvious similarities between consuming alcohol and consuming more recently developed psychoactive substances, even though these similarities are obscured by the fashionable label 'phenotechnologies' for the latter. Much of the current debate about human enhancement interventions – be it cognitive or phenomenal enhancement, but also physical enhancement – tends to underestimate the long history of the human struggle for improvement or for realising desirable states of mind. Being aware of it allows one to judge the current options in a different light: They are gradually different, but not substantially new.

Those writings in the contemporary debate about human enhancement that forget this long history and continuity usually make overexcited calls for an urgently needed, completely new way of thinking. So does Metzinger as a protagonist in the currently popular 'neuro hype'.³¹ What would be necessary instead is a more calm and sober analysis of the continuities – and also the distinctive discontinuities – in the intentional use of available means to bring about improvements in human functioning or experience.³²

5.2 Metzinger's consciousness 'ethics'

Usually, ethics is primarily concerned with actions, not with states, be they physical or mental. This holds true even for the moral theory most interested in mental states: utilitarianism or broadly consequentialism.³³ In a utilitarian framework, mental states of happiness are the metric that helps identifying which option for acting is to be preferred; yet they are only the metric and not the end in itself. The utilitarian maxim to pursue the greatest good (e.g. happiness) for the greatest number of people ultimately aims at assessing the moral quality of different actions according to the different outcomes of happiness they would realise. Metzinger, however, is only apparently close to this utilitarian reasoning, instead he focuses on conscious states as the ultimate unit of ethical concern. This questionable shift becomes obvious in a telling example he offers in his chapter on 'a new kind of ethics'. He asks: «Should it be legal, for instance, to let children experience their parents in a drunken state?» (ET, p. 222).

Admitting that moral reasoning can indeed provide a justification for political or legal arrangements and that there can be a significant overlap between ethics and

³¹HASLER 2012.

³²HEILINGER 2010.

³³Also the Aristotelian virtue ethics approach is no exception: even though ethics in this understanding is not about providing rules for individual actions, it is about assessing the overall quality of a human life in assessing how the life is lived.

law, it is still not sure that this question is in any standard way an ethically relevant one. If Metzinger sees a potential moral problem (standing in need of legal regulation, usually reserved only for the most severe moral challenges) in children experiencing their parents being drunk, he would have to specify what exactly should be outlawed: that parents – or anyone else – lets children see their respective parents in a drunk state, or whether the children's experience would be illegal. Only in the former case it would be about an action, namely the omission to hide drunk parents from their children. Yet, why should it be illegal that children see their parents drunk? The justification Metzinger seems to offer is that the children get into a painful state of mind and make experiences they should be spared from.

Doubts are appropriate here as to whether the example describes a morally relevant situation. Only further assumptions – e.g. the drunk parent behaving aggressively or brutal – would make for a morally salient situation. Experiencing a tipsy parent seems to be per se not morally relevant, and even less a matter of legal regulation based on moral considerations. Standardly and convincingly, from a moral point of view one would be interested first and foremost in the persons' *acting* (e.g. the parents acting brutally or a third person showing the children an imbibed parent who has lost control in order to scare them). The consequences these actions have on others, e.g. children, may play an important role to assess the moral quality of the respective action in a consequentialist framework.³⁴ In any case, these experiences are in themselves seen neither as an object of moral evaluation nor of legal regulation.

There is another danger attached to making experiences a matter of ethical or legal assessment, namely that of an intolerant policing of thoughts. It is a major achievement of the Enlightenment that we have the political liberties to think and imagine whatever we want to. If those personal, truly subjective phenomenal states become matters of legal concern and control, the realm of privacy and individual freedom would potentially be in danger. Here again, what people *say* and *do* matters morally and legally. But what they happen to be willing to *think* and *experience* is up to their personal choice, as long as it does not generate immoral or criminal actions.

The shift in attention from actions to states of mind in Metzinger's suggested 'new kind of ethics' is problematic. For one, his proposition is not sufficiently specific to fully make clear how phenomenal states could be subject to moral or legal assessment. The main concern with his suggestion to make phenomenal states matters of moral concern is that it seems to interfere with the subjective realm of privacy. Furthermore, it assesses and potentially criminalises experiences that are, simply because they are experiences and no actions, to a significant degree beyond the individual's control and hence no proper subject for moral or legal assessment.

In order to meet this objection, Metzinger could and should claim that of course not the experiences but the actions leading to these experiences are the subject of moral and legal assessment. In that case, however, there would simply be no need and no justification for a 'new kind of ethics'.

5.3 Getting lost in phenospace

To substantiate his ethical re-orientation towards conscious experience, Metzinger provides three conditions for a «desirable state of consciousness» (ET, p. 233). These desirable states are to

- minimise suffering

³⁴They may also be unimportant in other moral theories like deontology.

- possess an epistemic potential (expanding insight and expanding knowledge)
- have behavioural consequences that increase the occurrence of future valuable types of experience (ET, p. 233)

Here we can see how Metzinger gets lost in phenospace. The first condition is a purely phenomenal criterion, as is the third which combines a maximising utilitarian thought with the idea of minimal suffering as the best form of experience. But also the second condition – seemingly intended to provide some kind of connection to the ‘outer world’ – remains completely self-related and fully confined within phenospace: as we know by now, the ‘bridge to the world’ he is looking for has already been destroyed beforehand. On the one hand, Metzinger has raised doubt about the fact that we can ever make any reasonable claims about a ‘real world’ and a ‘real self’. Thus, any kind of genuine insight and knowledge about the world would be impossible. On the other hand, he has claimed that phenomenal states can intrinsically be connected with a *feeling of certainty*, so that any distinction between ‘real’ insight and knowledge and its faked counterparts becomes indistinguishable within the realm of phenomenal experience. As Metzinger puts it: «Self-deception may feel like insight» (ET, p. 220). But if we cannot tell self-deception from insight, the three criteria of valuable or desirable states of consciousness confine us to the domain of subjective experience. We are lost in phenospace.

With this, Metzinger faces the challenge of Robert Nozick’s famous thought experiment, the ‘experience machine’.³⁵ Nozick ponders the possibility of a machine that produces in us whatever desirable or pleasurable experience we could possibly wish for. Future neuropsychologists, so the imagined scenario goes, have found a way to stimulate a person’s brain to induce all kinds of pleasurable experiences. Nozick then asks us the decisive question: given the choice, would we choose the experience machine over our real life?

While Nozick argues that we have several reasons to reject the temptation to hook up with the machine, Metzinger’s answer remains ambiguous. On the one hand, he claims: «If it makes any sense at all to speak about the value of human existence, we must concede that it depends on more than the conscious experience of happiness.» (ET, p. 201). He suggests that engagement with truth and creative activity might be «at least as valuable as happiness» (ET, p. 201). On the other hand, his account of ‘valuable conscious states’ fails to provide any way of distinguishing between the value of states achieved in using such a machine and those brought about without it. His theory of valuable conscious states would speak in favour of using the machine – at least it evaluates both states equally – even if Metzinger seems somewhat reluctant to admit it.

In light of this, it is not surprising that Metzinger reports in great detail the results of a study that shows how important and valuable an artificially induced ‘spiritual experience’ under the influence of psilocybin was to a group of subjects (ET, pp. 225-27). He speaks in favour of the controlled use of such drugs as ‘phenotechnologies’ and praises its enriching outcomes for human beings.

The question whether one wants to hook up with a phenotechnological machine or lead a real life is ultimately an existential choice to be made. One has to choose between isolating oneself into an – admittedly blissful – engagement with machines or just to live a life. The latter crucially consists in a real-world striving for ends that one deems worthy of pursuit. Knowledge to the effect that our lives have no real-world

³⁵Nozick 1974, pp. 42-45.

consequences – that our striving and experiencing is without effective contact with reality – would inevitably devalue one’s experiences, whatever their hedonic qualities might be. Can one really want to be constantly fooled about one’s existential situation?

Here we see once again how Metzinger’s consciousness ethics depends upon his theoretical position concerning the Ego Tunnel: if what appears to be the world we live in is in fact no more than a world simulation within a mind-immanent phenospace, then indeed it makes no difference whether we stick with our ‘natural’ world-simulations or whether we artificially produce different ones. If there is no real world in the first place, taking the world as we know it away is not such a big deal. Life, philosophy, and ethics would all only be matters of conscious experiences – idle plays of projections in phenospace, or on Plato’s cave walls. Seen in this light, Metzinger’s tacit inclination towards choosing the experience machine over living a real life is understandable. But those who do not conceive of themselves as Ego Tunnels might prefer not to willingly enter the confines of an artificially enhanced phenospace brought about by altered brain states. Instead, they would be much better off continuing to engage as full organisms in exchange with the world and with others and keep living their lives in the world.

Consciousness ethics, in the end, would amount to no more than mundane reflections about the dangers and possibilities of new neuro-technologies, asking how these should be integrated in individual human lives, who should be responsible for administering them, and how abuses can be avoided. Of course, Metzinger has a point when he says that, so far, there has been too little reflection upon the values and dangers of certain extreme states of consciousness. But does this warrant a complete re-orientation of ethics? If indeed scientific advances make spectacular and wide-ranging alterations to human experience and human capacities possible, one might ultimately also reflect upon how these new capacities alter our conception of human beings and of worthwhile human lives. But this is a task no different in nature from conventional ethics and philosophical reflection about the nature of man and the good human life. No revolution – in ethics or elsewhere – is called-for.

Acknowledgements

We thank Michael Kienecker from *mentis-Verlag*, Münster, for granting permission to re-use previously published materials. We moreover thank Alexander Brödner and Philipp Haueis for helpful suggestions leading to an improvement of the manuscript.

References

- BENNETT, M. R. and P. M. S. HACKER 2003, *The Philosophical Foundations of Neuroscience*, Blackwell, Malden MA.
- CAMPBELL, N. 2010, “Towards a Critical Neuroscience of ‘Addiction’”, in *BioSocieties*, 5, 2, pp. 89–104.
- CHOUDHURY, S. 2010, “Culturing the Adolescent Brain: What Can Neuroscience Learn from Anthropology?”, in *SCAN*, 5, pp. 159–167. doi: 10.1093/scan/nsp030.
- CHOUDHURY, S. and J. SLABY (eds.) 2012, *Critical Neuroscience: A Handbook of the Social and Cultural Contexts of Neuroscience*, Wiley-Blackwell, Oxford.

- CHURCHLAND, P. S. 1989, *Neurophilosophy: Toward a Unified Science of the Mind-Brain*, The MIT Press, Cambridge MA.
- 2002, *Brain-Wise: Studies in Neurophilosophy*, The MIT Press, Cambridge MA.
- CLARK, A. 1997, *Being There. Putting Brain, Body and World Together Again*, The MIT Press, Cambridge MA.
- 2008, *Supersizing the Mind. Embodiment, Action, and Cognitive Extension*, Oxford University Press, Oxford.
- COOPER, M. 2008, *Life as Surplus: Biotechnology and Capitalism in the Neoliberal Era*, University of Washington Press, Seattle WA.
- DUMIT, J. 2004, *Picturing Personhood. Brain Scans and Biomedical Identity*, Princeton University Press, Princeton.
- 2012, *Drugs for Life: How Pharmaceutical Companies Define Our Health*, Duke University Press, Durham, N.C.
- FINE, C. 2010, *Delusions of gender: How Our Minds, Society, and Neurosexism Create Difference*, W.W. Norton, New York.
- FODOR, J. A. 1974, "Special Sciences (Or: The Disunity of Science as a Working Hypothesis)", in *Synthese*, 28, 2, pp. 97–115.
- GALLAGHER, S. 2005, *How the Body Shapes the Mind*, Clarendon Press, Oxford.
- HACKING, I. 1983, *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science*, Cambridge University Press, Cambridge.
- HARTMANN, M. 2012, "Against First Nature: Critical Theory and Neuroscience", in *Critical Neuroscience. A Handbook of the Social and Cultural Contexts of Neuroscience*, ed. by S. CHOUDHURY and J. SLABY, Wiley-Blackwell, Chichester, pp. 67–84.
- HASLER, F. 2012, *Neuromythologie. Eine Streitschrift gegen die Deutungsmacht der Hirnforschung*, Transcript, Bielefeld.
- HEILINGER, J.-C. 2010, *Anthropologie und Ethik des Enhancements*, de Gruyter, Berlin, New York.
- JORDAN-YOUNG, R. M. 2010, *Brainstorm: The Flaws in the Science of Sex Differences*, Harvard University Press, Cambridge MA.
- JOYCE, K. A. 2008, *Magnetic Appeal. MRI and the Myth of Transparency*, Cornell University Press, Ithaca.
- KITCHER, P. 2001, *Science, Truth, and Democracy*, Oxford University Press, New York, Oxford.
- LATOUR, B. 2005, *Reassembling the Social. An Introduction to Actor-Network-Theory*, Oxford University Press, New York, Oxford.
- LENZEN, W. 2006, „Auf der Suche nach dem verlorenen „Selbst“ - Thomas Metzinger und die „letzte Kränkung“ der Menschheit“, in *Facta Philosophica*, 8, S. 161–192.
- MALABOU, C. 2008, *What Should We Do with Our Brain?*, Fordham University Press, New York.
- MCDOWELL, J. 1994, *Mind and World*, Harvard University Press, Cambridge MA.
- METZINGER, T. 2003, *Being No-One. The Self-Model Theory of Subjectivity*, The MIT Press, Cambridge MA.
- 2009, *The Ego Tunnel. The Science of the Mind and the Myth of the Self*, Basic Books, New York.
- MIROWSKI, P. 2011, *Science-Mart: Privatizing American Science*, Harvard University Press, Cambridge MA.
- NOË, A. 2005, *Action in Perception*, The MIT Press, Cambridge MA.
- 2009, *Out Of Our Heads. Why You Are Not Your Brain And Other Lessons From The Biology Of Consciousness*, Hill and Wang, New York.
- NOZICK, R. 1974, *Anarchy, State, and Utopia*, Basic Books, New York.

- ORTEGA, F. and F. VIDAL 2011, *Neurocultures: Glimpses into an Expanding Universe*, Peter Lang, Frankfurt.
- PICKERSGILL, M. 2009, "Between Soma and Society. Neuroscience and the Ontology of Psychiatry", in *BioSocieties*, 4, pp. 45–60.
- REVONSUO, A. 2003, "The Contents of Phenomenal Consciousness", in *Psyche*, 9, 2.
- ROSE, N. 1996, *Inventing Our Selves: Psychology, Power, and Personhood*, Cambridge University Press, Cambridge.
- 2006, *The Politics of Life Itself: Biomedicine, Power, and Subjectivity in the Twenty-First Century*, Princeton University Press, Princeton.
- ROSE, N. and J. M. ABI-RACHED 2012, *Neuro: The New Brain Sciences and the Management of Mind*, Princeton University Press, Princeton.
- ROUSE, J. 2002, *How Scientific Practices Matter. Reclaiming Philosophical Naturalism*, Chicago University Press, Chicago.
- SLABY, J. 2010, "Steps Towards a Critical Neuroscience", in *Phenomenology and the Cognitive Sciences*, 9, pp. 397–416.
- SLABY, J., P. HAUEIS, and S. CHOUDHURY 2012, "Neuroscience as Applied Hermeneutics: Towards a Critical Neuroscience of Political Theory", in *Neuroscience and Political Theory*, ed. by F. VANDERVALK, Routledge, New York, pp. 50–73.
- STRAWSON, P. F. 1996, *The Bounds of Sense. An Essay on Kant's Critique of Pure Reason*, Routledge, London.
- SUNDER RAJAN, K. 2006, *Biocapital: The Constitution of Postgenomic Life*, Duke University Press, Durham N.C.
- TALLIS, R. 2011, *Aping Mankind. Neuromania, Darwinitis and the Misrepresentation of Humanity*, Acumen, Durham.
- THOMPSON, E. 2007, *Mind in Life. Biology, Phenomenology, and the Sciences of the Mind*, Harvard University Press, Cambridge MA.
- VARELA, F., E. THOMPSON, and E. ROSCH (eds.) 1991, *The Embodied Mind: Cognitive Science and Human Experience*, The MIT Press, Cambridge.
- WILLASCHEK, M. 2003, *Der mentale Zugang zur Welt*, Klostermann, Frankfurt a. M.