

Book Reviews

J. Schloss & M. Murray (eds.)

The Believing Primate: Scientific, Philosophical and Theological Reflections on the Origin of Religion

Oxford: Oxford University Press, 2009

Reviewed by István Czachesz

Helsinki Collegium for Advanced Studies & University of Helsinki

This volume is the first collection of essays on the cognitive and evolutionary study of religion to encourage a dialogue between scientists and theologians. In the Introduction (pp. 1–43), Jeffrey Schloss surveys evolutionary explanations of religion and distinguishes ‘cognitive’, ‘Darwinian’, and ‘co-evolutionary’ accounts (p. 16). It appears that Schloss equates ‘cognitive’ theories with ones that regard religion as a by-product of Darwinian evolution — a somewhat misleading categorization, which excludes scholars arguing for an adaptive role of religion from the field of the cognitive science of religion.

Dominic Johnson and Jesse Bering (pp. 26–43) reformulate the frequently advocated hypothesis that believing in supernatural punishment has an adaptive value in human evolution. They suggest that individuals holding such beliefs will have a better chance to survive because they have a ‘lower probability of detection’ (that is, as transgressors; p. 39), considering arguments about the success of cooperative groups only as of secondary importance (p. 40). Whereas some causal link between religious beliefs and moral behaviour is quite plausible, as also argued in other essays in the book, one wonders whether moral feelings (guilt, shame, emotional reward) are not much simpler and cost-effective tools to constrain selfishness and avoid punishment in a group of humans equipped with an ‘intentionality system’ (pp. 34–36).

Joseph Bulbulia (pp. 44–75) elaborates on his ‘commitment-signaling’ model of religion, previously described in his other publications, in an attempt to offer a plausible alternative to ‘spandrelist’ conceptions of religion that have dominated the cognitive science of religion (p. 75). In terms of this model, hard-to-fake displays of religiously motivated emotions facilitate moral exchange. In this essay, Bulbulia especially describes the ‘special counterfactual scope operator’ @IMAGINE (p. 62) that decouples religious fiction from practical domains, nevertheless allowing it to be represented as true. Using this operator, religious people are able to confabulate events of religious experience, which give ‘emotionally powerful episodic support to religious commitment’ (p. 69).

The essay by Justin Barrett (pp. 76–99) reviews various findings in the cognitive science of religion (e.g., about minimal counter-intuitiveness, hypersensitive agent detection, intuitive theism, intuitions about death) to argue that cognitive and evolutionary explanations of religion are not hostile to Christianity but ‘compatible with orthodox Christian theology’ (p. 77). According to Barrett, cognitive sciences can enrich our understanding of how humans are made to ‘understand God sufficiently to enjoy a relationship with Him’ (p. 76).

Peter J. Richerson and Lesley Newson (pp. 100–117) use multilevel selection theory to examine the question whether religion is adaptive. They argue that given their biological and cultural complexity and diversity, ‘phenomena like religion are unlikely to support sweeping generalizations about adaptation versus maladaptation’ (p. 117). In the domain of cooperation, they consider historical cases rather than abstract models (cf. Bulbulia), and conclude that religion can have either negative or positive net fitness effect on a population (p. 110). Looking at different contemporary religious communities, Richerson and Newson suggest that small sects benefit the most from religion, being less at risk from penetration by selfish impostors (p. 113). At the level of cultural evolution, adopting new religious ideas seems to be beneficial for many people (p. 116).

Paul Bloom clearly states his position already in the title of his essay (pp. 118–127): Religion is an ‘evolutionary accident’. He identifies, in particular, three cognitive mechanisms that underlie religious beliefs: Agent-detection, intuitive creationism, and common-sense dualism (due to our separate cognitive systems dealing with material objects and social entities).

Two essays in the middle section of the book examine the possible consequences of the cognitive science of religion for Christian theology. In an answer to Bloom’s essay, Peter van Inwagen (pp. 128–138)

considers whether Bloom's explanation of religion, if it is correct, implies that supernatural belief is irrational or defective. He concludes that there are no such consequences. On this account, the cognitive enterprise is neutral with regard to the rationality and truth of Christian theology, but does not seem to support it particularly, as in Barrett's view. It has to be noted that in his criticism of the psychological universals mentioned by Bloom, van Inwagen seems to write about conscious and reflective judgments about agency, design, and dualism (pp. 130–131), whereas Bloom refers to implicit cognitive biases, measured by sophisticated, indirect tests rather than by straightforward questions.

Alvin Plantinga (pp. 138–178) examines the conflict between evolutionary psychology and Christian beliefs on a larger scale. He finds, on the one hand, that many of the cognitive and evolutionary theories of religion are incompatible with Christian belief. On the other hand, however, he concludes that these theories do not defeat Christian beliefs or should not be a source of cognitive dissonance for Christians, because they are based on evidence that is constrained by the Methodological Naturalism of (most) science. In this way, Plantinga sees more conflict between the cognitive science of religion and Christian faith than either Barrett or van Inwagen, and is concerned with finding an intellectual niche for Christian theology that does not exclude taking science seriously.

The next three essays contain critical reflections on the cognitive science of religion. Michael J. Murray (pp. 168–178), one of the editors of the volume, discusses why cognitive explanations make us think that religious beliefs are epistemically suspect. Murray also co-authored an essay with Andrew Goldberg (pp. 179–199), in which they argue that evolutionary explanations do not explain away religion. Charles Taliaferro's essay (pp. 200–214) is a response to Daniel Dennett's *Breaking the Spell* (New York: Viking Penguin, 2006). The author cites Dennett's view of consciousness as a root of his rejection of religion and argues for the superiority of a theistic frame of explanation. Taliaferro's criticism of Dennett deserves credit in several respects, especially when it comes to Dennett's rather narrow concept of religion, and the ambiguous category of the supernatural. Yet the theistic alternative suggested by Taliaferro is hardly viable for scientists: By definition, a theistic explanation (and any explanation based on intentionality, for that matter, see p. 208) cannot be subjected (even potentially) to controlled, empirical testing.

Del Ratzsch (pp. 215–245) takes up the perennial debate about the relation of science and religion, concluding that science cannot

undercut religion, or be hermetically separated from it, or take rational precedence over it, or be conceptually and intellectually autonomous (p. 244). Nancey Murphy argues in her essay (pp. 265–277) that the cognitive science of religion poses no threat to Christianity, but in fact can be used for theological purposes.

The last three contributions deal with the connection of religion and morality. Joseph Haidt (pp. 278–291) offers an overview of his theory of the five foundations of morality in his essay. Reasoning about harm/care and fairness/reciprocity is concerned with individual moral behaviour, and underlies liberal ethics. Conservative and religious values also incorporate an emphasis on in-group/loyalty, concerns about authority/respect, as well the idea of purity/sanctity. Haidt argues that the New Atheists, such as Richard Dawkins and Sam Harris, misrepresent the role of religion in human evolution when they emphasize factual beliefs (e.g., in creation or the afterlife) or literalist readings of sacred texts, forgetting about personal and societal benefits. Empirical research, Haidt argues, demonstrated such benefits beyond doubt, for example showing the greater willingness of religious believers to donate money, sacrifice time, and give blood to complete strangers. Christian Smith (pp. 292–317) argues that transcended monotheism was a crucial condition in giving rise to our current commitments to benevolence and human rights. David Sloan Wilson (pp. 319–338) conducts an interesting dialog with Smith in an attempt to build a bridge between evolutionary theory and social constructivism. After summarizing his own theory of multi-level selection, Wilson endorses many of Smith's critical remarks on evolutionary psychology and welcomes the idea of a universal, moral, and believing human nature — but suggests an evolutionary, rather than theistic, framework of explanation to study it.

With some knowledge of the field of the cognitive science of religion, one may be surprised by the predominance of adaptationist accounts of religion in the book, the only strongly spandrelist essay being that by Bloom. On the one hand, this can be seen as a welcome corrective to the overrepresentation of the spandrelist hypotheses in public discourse, especially by projects debunking religion, which have given the cognitive science of religion a bad reputation in religious circles. On the other hand, as Jeffrey Schloss emphasizes in the Introduction (p. 22), adaptationist accounts of religion do not prove religion right *per se* or justify it in a moral sense — even though the essays of Haidt and Wilson might give that impression. One can only wish that the cognitive and evolutionary study of religion remains as neutral as possible with regard to political and ideological battles and

continues to generate interesting, new scientific knowledge about religious thought and behaviour, these fascinating and puzzling characteristics of the human species.

Tim Bayne, Axel Cleeremans, and Patrick Wilken (eds.)

The Oxford Companion to Consciousness

Oxford: Oxford University Press, August 3, 2009, 672 pp., plus introduction

Reviewed by Bill Faw

Brewton Parker College bfaw@bpc.edu

These are 672 pages well worth reading. End of review!

Well, perhaps I should say a bit more to help you readers decide whether to spring for the \$135 list price — or even the \$100.85 that Amazon.com is charging for new or used copies, as of St Pat's Day, 2010.

I need to begin with three disclaimers: First, it is easy for book reviewers for big-league journals to urge others to spend their money on expensive books. We get the books free. This leads to my second disclaimer: I did not ask for the *JCS* freebie because I received two copies from OUP having written one of the entries in the Companion ('consciousness, modern scientific study of', if you must know). My third disclaimer is that I actually read the entire book, violating the accepted etiquette for reviewing long and complex books. The usual rule of thumb is: The longer the book, the less of it the reviewer has to actually read. It wasn't bad enough that I read the whole book. I even jotted down many of the article titles by category: Philosophical, psychological, neurological, biological, and the like.

The Oxford Companion to Consciousness (or *C2C* as it is affectionately known) has approximately 243 titles by 250 authors over 672 pages. Large complex books on Consciousness can take several forms. One form is for a single author to pull together a whole lot of material into a single book, like Koch's *Quest* or Baars' books on the scientific side and Dennett, Lycan, Searle, Chalmers, and many others on the philosophical side. Those books give coherent stories, but don't have a dazzling breadth of coverage. Another form is for a few leaders in the field to write long articles on central themes. That has the advantage of both coherence and breadth. But this 'C2C' style book has its own place, with dazzling breadth, but still articles authored by many of the greats in the field, and others by authors like me. However, the impact on the thorough reader is like picking up 243

separate jig-saw puzzle pieces — one by one — and trying to try to put the entire puzzle together, to see what the completed picture looks like.

C2C is a sheer treasure chest of just about everything you might want to know about modern consciousness studies — scientific and philosophical. In the Preface, the editors expressed their deep debt to the standing consciousness conference to which they have been closely associated: The Association for the Scientific Study of Consciousness (ASSC) — which had met at Oxford University in 2006, shortly before this project was launched. This book reflects ASSC's balance of western philosophy and science, with only occasional 'eastern' topics. As with ASSC conferences, the occasional 'eastern' articles tended to be produced by Susan Blackmore, who has credentials in all three camps. In this book, Susan wrote on *Altered States of Consciousness*, *Bicameral Mind*, *Consciousness Expansion*, *Extra Sensory Perception*, and *Grand Illusion*.

JCS readers who attend Tucson and Tucson-abroad conferences (under the moniker of *Toward a Science of Consciousness*) strictly because of their heavier dose of 'eastern' consciousness studies, may be disappointed with this book. However, ASSC attendees and those who attend Tucson because of its blend of science and western philosophy will be quite satisfied. Susan Blackmore is not the only multiple-article author in this book. The 'philosophy' editor, Tim Bayne, wrote on *Access Consciousness*, *Delusions*, *First/Third Person*, *Functions of Consciousness*, and *Homunculus*. I assume that he filled in gaps that were left.

For truly ambidextrous science/philosophy readers, the crown jewels in this book were the dozen or so article-pairs which had two articles back to back, the first titled (for instance) 'Action, Philosophical Perspectives' and the second titled 'Action, Scientific Perspectives' — and by different authors. The following topics have these pairs: Action, colour, dreaming, emotion, imagery, musical experience, orgasm, pain, self, and temporality. 'Perception' has a similar pair: 'Philosophical perspectives' and 'unconscious'. 'Correlates of Consciousness' has a triplet: Computational, Philosophical Perspectives, and Scientific Perspectives. These articles represent consciousness studies/science at its best! Readers who are convinced that assembling the jig-saw puzzle of Consciousness requires such ambidexterity might want to begin reading this book with those paired articles. Quoting John Denver, that is 'almost heaven'!

There are plenty of single articles for more mono-dexterous philosophers or scientists. However, let me warn you — many of these

philosophers have a decent grasp of the science, and *vice versa!* For instance, in addition to those mentioned above, you will find such philosophical topics as acquaintance, biological naturalism (by Searle), body image/body schema (Gallagher), Chinese Room argument, cognitive feelings, concepts of consciousness (Van Gulick), conceptual thought, contents of consciousness, dual aspect theories, dualism, eliminativism, emergence, epiphenomenalism (Robinson), epistemology of consciousness, explanatory gap (Levine), externalism, free will, hard problem (not by Chalmers), heterophenomenology (Dennett), higher order representative theories (Lycan), idealism, identity theory, intentionality, introspection, inverted spectrum, knowledge argument, multiple drafts model (Dennett), mysterianism (Kreigel), and on and on.

Those interested mainly in the scientific articles will be as excited as the philosophers. First, some of the neurological and psychiatric topics, most of which will not pass my computer spell-check: Achromatopsia, agnosia, akitenopsia, amnesia, anaesthesia, anarchic hand, anosognosia, apraxia, arousal vs awareness, autism, and automatism — all in the A's! Also, Balint's, blindness, blind sight (by Stoerig and Covey), Charles Bonnet Syndrome, commissurotomy (Trevarthen), confabulation, delirium, delusions, dementia, depersonalization, DID, epilepsy, hallucinations, neuropsychology and states of disconnected consciousness (Weiskrantz), and so on. Other brain research topics: EEG, forward models, frontal cortex, functional brain imaging, information integration theory (Tononi), gamma oscillations (Engel), hypnogogic experience, hypnosis, memory systems (Schacter), mirror neurons (Gallese), neural stimulation, and the like.

Just a few of the broader psychological topics: Global workspace theory (Baars), microconsciousness (Zeki), neuro ethics (Farah), neural global workspace (Dehaene), lucid dreaming, animal consciousness, infant consciousness, empathy, evolution of consciousness, inner speech, literature and consciousness, mind wandering, eye of origin, feeling of knowing, filling in, flash lag, fringe, inattentional blindness, and several articles starting with 'memory'. (The reader has perhaps figured out that I ran out of OCD motivation to categorize these around the 'N's'). Especially delightful were articles on solipsism, stream of consciousness, orgasm, and (Mr.) Spock. A bit strangely placed here were 'Wine' and 'Consciousness and Subliminal Tapes'. **Those who worry that Consciousness Science is focusing more and more on non-conscious processes while minimizing the role of conscious processing, will be confirmed in their worries**

by the fact that the book's listings range from 'absent qualia' to 'zombies'.

You should probably not read the book straight through. That would be hard to do. Many approaches would be more interesting. You might want to start with some of the many articles mentioned above. Or open the book at any page or thumb through the book, glancing at the titles and authors over the 672 pages. Flip until you see something intriguing and read that. Then read whatever articles that one refers you to. After you have run that rabbit trail, thumb through until you see another teaser. I ended up reading about half of the book using that method. Then I steeled my resolve and went *kiver to kiver* reading whatever I had missed. By that time I had volunteered to review this book. That might have helped steel my resolve.

You get the picture that this is a great book. And I know, from communications with the editors during the construction of this book, that it was a mammoth effort by Bayne, Cleeremans, and Wilken. But there are some non-content aids that are missing — which would have made it easier to put together jig-saw pieces. Because of its encyclopedia format, it has no table of contents or index. If you want to read something about 'free will' you look to see if there is an article title starting with those words, by working your way to page 296. There is a list of the contributors, so you can look to see that Dennett has at least one article, but that Chalmers does not. But you have to be lucky in guessing that there might be articles on heterophenomenology and multiple drafts model — and that they might be by Dennett. A table of contents with each article's name and author(s), would have been helpful. In fact, since the articles are in alphabetical order, the TOC should have been alphabetically by author.

Lacking that, the first thing I did when I received my copies was to read through the 5-page list of contributors and mark each name that I recognized. But then I wanted to see what my old pals John Bickle, Don Dulany, Tony Hudetz, Bill Lycan, Bill Robinson, and Bob Van Gulick had written. But I had no way to do that. So I spent another hour looking at each article's title and author (at the end of the article) and read articles by some of those people. I had gone through all of this well before I realized that I might write a review of this book. These are not criticisms as much as suggestions for when *C2C-The Sequel* comes out.

One of the editors of *C2C* has sent me (over the past week) reviews of this book by the journal *Brain* and by the American Psychological Association. It says a lot about this great book that these century-plus-aged organizations involved in brain research and psychology

would review this book. I will not read these reviews until I submit my review of the book — and then see what they caught that I missed!

Thomas Metzinger

The Ego Tunnel: The Science of the Mind and the Myth of the Self

New York: Basic Books, 2009, 276 pp.

ISBN: 978-0-465-04567-9

Reviewed by Carolyn Michèle Suchy-Dicey

Department of Philosophy, Boston University, cofounder of Neuphi

Ego tunnels, suggests Thomas Metzinger, are tunnels through reality: The ego, making up the walls of the tunnel, is constructed in the same way and at the same moment as its representation of reality, the content of the tunnel. Our experience of the ego tunnel is called ‘phenomenal subjectivity’, where the term ‘phenomenal’ picks out the experiential aspect of this subjectivity. In *The Ego Tunnel*, Metzinger attempts to explain phenomenal subjectivity while assuming naturalism, or the view that all existing things share a common ontological substrate. His big idea is that we experience phenomenal subjectivity because we possess transparent phenomenal self-models. These phenomenal self-models are transparent when we are unable to experience them as models. We look through them to the world and in doing so take self and world to be directly accessed through introspection.

In a recent review in *The New Scientist*, Owen Flanagan objects that *The Ego Tunnel* (henceforth, *TET*) is a waste of the paper it is written on because consciousness researchers already accept its conclusions. Flanagan has, however, misunderstood the book’s target audience. Metzinger says, ‘This book has not been written for philosophers or scientists. Instead, it is my first attempt to introduce a wider public to what I think are the truly important issues in consciousness research today’ (ix). This move is useful for those who found, like Alex Gamma, that *Being No One* (Metzinger’s earlier work) was ‘ungainly and complicated, exceedingly abstract and burdened with superfluous words that obstruct the reading flow’; *TET* is a welcome departure from that style to one that is both accessible and informative. Regarding content, Flanagan claims that *TET* simply rehearses old arguments establishing the illusion of selfhood. On the contrary, *TET* does not directly argue for this purported illusion, which was already amply discussed in *Being No One*, and instead assumes this result in attempting to provide a naturalistic explanation of the experience of

phenomenal subjectivity while accounting for the possible social and ethical side effects of such an explanation.

TET is divided into three parts. The first presents the main problems scientists face in attempting to explain conscious experience, ending with the problem of subjectivity, or the problem of providing a naturalistic explanation of subjectivity. The second part provides a solution to the problem of subjectivity in terms of the phenomenal self-model. The third and final part takes a look at the ethical implications that are entailed by Metzinger's account of subjectivity. The book offers clear, concise descriptions of empirical work, philosophical theories, anecdotes about the author's first hand experiences with lucid dreaming and out-of-body experiences, and interviews with three researchers and one imaginary philosopher — the first 'post-biotic' philosopher.

The first part of *TET* addresses what Metzinger sees as the six biggest problems in consciousness research today:

- (1) The unity of experience given the discontinuity of brain activity;
- (2) The experience of temporal indexicality given the uniformity of space–time;
- (3) The apparent transparency of experience to an external world given the evidence of perceptual construction;
- (4) The ineffability of particular phenomenal experiences;
- (5) The apparent lack of functional purpose for phenomenal experience;
- (6) The experience of a self–world border given that the self and world, if natural, are not separable phenomena.

Metzinger offers solutions to the first five of these problems in part 1, and the sixth problem is taken up in part 2. As a taster, consider how Metzinger solves the fifth problem, or the seeming lack of functional purpose for phenomenal experience. The idea is that in order to form strategies we require the capacity to imagine alternative scenarios. However, if we were unable to distinguish these imagined scenarios from reality, we might fail to recognize real dangers. The function of phenomenal experience is therefore to indicate the difference between imagined worlds and the real world, and the function of phenomenal subjectivity is to indicate the difference between our imagined perspective and our real self. Phenomenal experience and subjectivity function as reference frames that travel with us into these 'offline simulations'.

The second part of *TET* introduces the concept of a phenomenal self-model to answer what Metzinger sees as the most important problem of consciousness studies, the explanation of an experienced self–world border. The phenomenal self-model is a model of the self provided by the brain. The model represents more than the body–world boundary used by our immune system: The *self*–world boundary requires a sense of agency, or a model of ourselves as agents:

The decisive transition takes place when the system is already given to itself through minimal self-consciousness and then, in addition, represents itself as being directed toward an object. I believe this happens exactly when we first discover that we can control the focus of attention (p. 102)

That is, ‘epistemic control’ (knowledgeable control) arises when the brain tracks correlations between the intentions of the embodied self to perceive a particular object and the turning of its sensory organs to that object via attention. Thus, experience of agency arises as a result of temporal congruity between the intention to act and perceived action. Metzinger is careful to attribute this phenomena of ‘attentional agency’ to controlled attention, or what is called top-down attention in the neuroscientific literature and endogenous attention in the psychological literature; it is the effortful matching of intention and result that the brain ascribes to agency, much as the effortful matching of need and result is ascribed to bodily action.

One might wonder how this explanation of the natural development of agency could explain the ‘inwardness’ of phenomenal subjectivity. Metzinger argues that this is all there is to inwardness: Information that is reflexive by creating a model of itself. Thus, the problem of what it is like to be a bat becomes the problem of discovering the self-model of a bat (if there is one), and the problem of explaining why there should be what-its-likeness at all is just the problem of explaining why there should be phenomenal self-models, which is answered above.

The third part of *TET* explains Metzinger’s felt imperative:

Scientists and academic philosophers cannot simply confine themselves to making contributions to a comprehensive theory of consciousness and the self. If moral obligation exists, they must also confront the anthropological and normative void they have created (p. 215)

The normative void that Metzinger is alluding to is that created by the change in the public understanding of consciousness. A normative void is created when new understandings of consciousness enter

public dialogue but the public remains uncertain about the exact nature of these changes and of their social and ethical implications. When public understanding is altered, socio-political institutions require guidance on how to apply this new knowledge in making normative decisions. Thus, the normative void requires a clear communication of both the empirical and conceptual findings and suggestions for application. Metzinger attempts to do both: Most of the book is dedicated to a clear explanation of the empirical and conceptual findings relating to the self, and the final part of the book is his attempt at providing a moral groundwork for the new field of consciousness ethics.

An Epicurean tone resonates throughout this third part of the book, as Metzinger claims ‘if there is such a thing as forbidden fruit in modern consciousness research, it is the careless multiplication of suffering through the creation of artificial Ego Tunnels without a clear grasp of the consequences’ (p. 197). *TET* tries to motivate an argument against work on artificial consciousness until we have maximal knowledge by asserting that suffering is a natural evil that must be prevented and that, furthermore, ‘we must make sure always to err on the side of caution’ (p. 193). The risk that Metzinger alludes to here is that of the unforeseen needs of artificially conscious beings, such that our lack of understanding results in their suffering. One can imagine here an alien baby, helpless and alone, in anguish over felt needs that we are unable to understand or secure. This picture is surely bleak, but whereas Metzinger’s assertion that suffering is a natural evil probably won’t find many detractors, he offers little support for the claim that the *risk* of suffering is likewise to be avoided at all cost. We regularly and rightly risk suffering for greater goods, and the research on artificial consciousness is not a clear exception. To avoid all risk is to shut down the research program on artificial consciousness, which is what Metzinger is proposing. Without further argument, this call for caution is not only unwarranted but is also impractical in a way that threatens to render the entire third part of the book at risk of marginalization.

In sum, *TET* is a bold and innovative sketch at a naturalistic explanation of the self that uses the most recent empirical and conceptual work available to make its argument. At the same time, it offers a preliminary understanding of consciousness ethics and the kinds of questions that researchers should be asking themselves as they proceed in this area. All of this should be valuable for consciousness researchers who keep in mind that the book is exploratory, rather than definitive.

References

- Flanagan, Owen (2009), 'Review: *The Ego Tunnel* by Thomas Metzinger', *New Scientist*, 20 March, 2009.
- Alex Gamma (2003), 'Review of Thomas Metzinger's *Being No One: The Self-Model Theory of Subjectivity*', *Brain and Mind*, 4 (3), pp. 385–393.
- Metzinger, Thomas (2009), *The Ego Tunnel: The Science of the Mind and the Myth of the Self* (New York: Basic Books).

Peter Goldie (ed.)

The Oxford Handbook of Philosophy of Emotion
Oxford: Oxford University Press, 2010, 722 pp., £85

Reviewed by Giovanna Colombetti
Department of Philosophy, University of Exeter

This weighty new Handbook is a very welcome addition to the Oxford Handbooks of Philosophy series. There are already some volumes that bring philosophical approaches to emotion together, but this collection is much more comprehensive than any other published so far (and that probably ever will), and has something for all researchers interested in emotion — mainly philosophers, but not necessarily only. The succinct introduction by the editor, Peter Goldie, usefully summarizes each contribution in one paragraph, providing a map of the volume. The Handbook itself is divided into six parts, each containing four to six articles: Part I, 'What Emotions Are' (articles by Deigh, Ben-Ze'ev, Cowie, De Sousa); Part II, 'The History of Emotion' (Price, Gill, King, Abramson, Hatzimoysis, Charland); Part III, 'Emotions and Practical Reason' (Elster, Döring, Helm, Tappolet); Part IV, 'Emotions and the Self' (Ratcliffe, Pugmire, Morton, Stocker, Rorty, Hobson); Part V, 'Emotion, Value and Morality' (Mulligan, Neu, Prinz, Greenspan, Roberts, D'Arms and Jacobson); and Part VI, 'Emotion, Art, and Aesthetics' (Matravers, Feagin, Robinson, Kieran).

For the purposes of this review I have chosen to focus mostly on the contributions in Part I and Part IV — which, I believe, are most likely to attract readers of *JCS* — and a few other articles in sparse order. Looking at the volume as a whole, Part II and Part VI in particular differentiate it from other collections in the philosophy of emotion. Part II is especially useful, with original contributions showing the richness and complexities of past emotion theories, from Plato to Sartre via the Stoics, Aquinas, the Sentimentalists, and nineteenth century French psychopathologists — just to mention some; this section really fills an important gap (even though a lot remains to be done in this

area) and a more historically informed approach will hopefully influence future debates and conceptualizations of emotion.

Readers of *JCS* should be warned, however, that the Handbook contains mostly ‘armchair’ philosophical analyses, with little reference to affective-scientific empirical and theoretical work (this is the case also for Part I and Part IV). One apparent missing link in the collection is the philosophical-scientific approach to emotion; there is no section dedicated to the philosophy of biology (think for example of the contribution that Griffiths, 1997, has made to the philosophy of emotion) and/or the philosophy of neuroscience. The question whether emotions are natural kinds, for instance, is one currently debated among philosophers but also psychologists (e.g. Barrett, 2006), and it would have been useful for interdisciplinary purposes to collect some philosophical discussions of this complex issue. Related questions of whether or not emotions are ‘modules’, of whether it makes sense to look for the ‘neural correlates’ of so-called ‘basic’ and/or ‘discrete’ emotions, and of whether or not the latter notions should be kept, remain largely marginal. This choice, I think, is unfortunate, for it loses one opportunity to show how philosophers can usefully contribute to research in other disciplines. It is ironic that De Sousa’s paper in Part I invites philosophers to embrace and consider scientific results in a volume that ends up minimizing contributions from philosophy of science. Goldie himself in the Introduction remarks that one of the reasons why so many philosophers are now interested in emotion is the increasing attention paid by philosophers of mind to empirical work in cognitive science. Yet there is no detailed philosophical discussion of current affective science in the Handbook. ‘Interdisciplinarity’ consists only in the inclusion of a few papers by psychologists. In Part I Roddy Cowie talks about current research in emotion-oriented computing, and how difficult it is to make computers react appropriately to human emotions, not least because humans appear to be expressing emotions most of the time, and more than one emotion at once (unfortunately, however, the article does not give the details of any specific project in this research area, so the reader is left to wonder exactly how the difficulties mentioned by Cowie arise, and what they tell us about the nature of human affectivity). In Part IV Peter Hobson provides a useful overview of developmental findings of the place of emotion in intersubjectivity, and of his own recent studies on autism that corroborate his claim that certain emotions and modes of affective interactivity are necessary for the development of the notions of self and other. He also uses his findings to question some of Goldie’s (2000) points about how we

understand others' emotions, such as Goldie's claim that sympathy does not require sharing the other's affect. This is all very welcome, but also calls for a similar attitude from the philosopher of emotion to engage with the scientist's work (not just the other way round). Papers that mention and partly rely on empirical work in the rest of the volume are those by Tappolet (who criticizes the notion of 'motivational modularity' and discusses works in evolutionary psychology among other things), Prinz (who draws on the experimental literature on specific emotions, such as disgust, to assess the place of emotion in morality) and Robinson (who interweaves aesthetic theories of music with a variety of psychological and neuroscientific works to argue that music can arouse non-intentional moods without the intervention of a cognitive appraisal). Note however that these papers mainly take empirical research for granted, and primarily use it to support theoretical claims with philosophical import; what is also needed, I think, are philosophical approaches that question and/or clarify some of the conceptual apparatus used by affective scientists, and that can engage critically and knowledgeably with their methodology.

Another area that is underrepresented in this Handbook is phenomenology. Goldie (p. 1) seems appreciative of this tradition: 'Philosophy of Mind in the Anglo-Saxon tradition was for a long time (and in some way still is) preoccupied with the mind-body problem, ... and had little truck with the work of the phenomenologists, much of which included insightful discussions of the emotions'; yet again this appreciation is not really reflected in the choice of readings for the volume. Phenomenology here fares just a little bit better than philosophy of science, with Hatzimoysis' focused and clear overview of some of Heidegger's and Sartre's most relevant points on affective phenomena, and Ratcliffe's original elaboration of Heidegger's notion of the 'depth' of moods.

Otherwise however most papers do belong to the Anglo-Saxon analytical tradition. Although they are not much concerned with the mind-body problem — which, incidentally, I think is a pity given the still very controversial status of the place of the body in emotion — many of them endorse the related distinction between 'intentional' and 'feeling' theories of emotion. (This distinction largely overlaps with the one often drawn by psychologists of emotion between 'cognitive' and 'Jamesian' theories. Roughly, according to the former emotions contain and/or are caused by a cognitive evaluation of some event in the world, and according to the latter emotions necessarily involve feelings of bodily changes). In spite of various recent attempts to overcome this dichotomy (for Goldie, 2000, feelings have their

own form of intentionality which is not reducible to the one of propositional attitudes such as beliefs and desires; for Ratcliffe, 2008, and Slaby, 2008, even bodily feelings can be part of the structure of intentionality), Deigh and Ben Ze'ev (and seemingly De Sousa too), for example, still take the phenomenal aspect of emotion to consist in 'mere feelings' enclosed within themselves, dislocated from any meaningful action and interaction with the world. Contrast this view with the phenomenological approach, according to which all conscious states (feelings are no exceptions) are intentional or world-oriented — in the narrow sense of oriented towards specific objects, or in the broad sense of being 'open' toward otherness (for a clear introduction, see Thompson and Zahavi, 2007). In this approach, even bodily feelings can be intentional. Indeed, as Husserl and many others have emphasized, the body as experienced by the subject (known as 'the lived body') is part of the subject's awareness of the world. Bodily self-awareness is not just, and not even typically, the perception of one's own body via e.g. proprioception; it importantly includes a 'non-reflective' (or 'pre-reflective') awareness of one's own body as that through which one experiences the world. From this standpoint, it becomes possible for bodily feelings in affective experience to be not just perceptions of physiological changes detached from the world, but bodily ways of appraising a situation as e.g. dangerous, enthusing, maddening, and so on.

Even though the dichotomy between feeling and intentional theories of emotion is not likely to go away easily, a sincere interest in phenomenology should lead to a more informed engagement with this tradition. Instead, phenomenology is still all too often taken to be merely synonymous with some kind of more or less disciplined introspection, with all the negative connotations that this term carries with it. De Sousa's paper is representative of this attitude. He offers two arguments against phenomenology, 'interpreted as the doctrine that skilled introspection can give reliable access to the character and meaning of one's own mental states or dispositions' (p. 98). The first argument appeals to established empirical findings which show that we are often mistaken about e.g. what we think will make us happy, and how we will behave in a specific situation. The second argument embraces content externalism and states that because meaning depends on factors outside the subject's knowledge, it cannot be fully disclosed by introspection. Hence, 'what emotions feel like cannot give us full access to their nature' (p. 100). These arguments mischaracterize the phenomenological enterprise. It suffices here to point out that Husserl's *epoché* is meant to consist in a change of

attitude from a naïve realistic belief in the natural world, to a careful attending to how the world (including oneself) is experienced, or better 'given' to the subject in experience. This careful attending is not meant to disclose 'the full nature of mental events' where that would include unconscious processes that are beyond experience. Phenomenologists are interested in attending to the phenomena themselves, for what they are; the 'meaning' they are interested in is the one experienced by the subject.

The consequence of De Sousa's interpretation is that he ends up inserting a sharp wedge between phenomenology and science: Because science has shown that we do not have privileged access to our mental states, we should look at science and not phenomenology to understand what emotions are. Yet as some readers of *JCS* will know already, and as De Sousa himself briefly acknowledges in a footnote (in which he also admits that his interpretation of phenomenology is 'narrowly focused', see p. 98), there are at present various attempts to integrate phenomenological methods with empirical data generated by the cognitive sciences. Although not all phenomenologists approve of this partnership (see discussion in Zahavi, 2004), the relevant point is that 'skilled introspection' need not be at odds with the scientific enterprise, but can complement and illuminate it (as well as *vice-versa*, but this is another part of the story; see Gallagher and Zahavi, 2008, for an introduction to the project of naturalizing phenomenology).

As for Part IV ('Emotions and the Self'), note that with the exception of Hobson's paper mentioned earlier, this section does not address questions such as whether and how affectivity makes up the self, and/or whether some sense of self is always implied by affective experience; nor does it address existing scientific hypotheses that posit a constitutive link between minimal or core selfhood, affectivity, and the body (see e.g. Panksepp, 1998, and Damasio, 1999). However once one comes to terms with the fact that this section is about emotions and the self in the broad sense of how emotions fit together with other aspects of a person's mental life, then one can enjoy some interesting and novel contributions. I have already mentioned Ratcliffe, who emphasizes that affective phenomena can be more or less 'deep' depending on the degree of specificity of their intentional objects; thus sadness for the loss of one's favourite pet is shallower than sadness about one's inability to engage meaningfully with other people, which is shallower than sadness about the status of human rights in many countries (the examples are mine). Pugmire offers a skilful and elegant discussion of what language does to feelings. Feelings can be

ineffable; yet, once put into words, they can change in various ways. Sometimes language distorts experience and distances one from it (describing and representing turn *my* feelings into objects), but other times — think of poetry — it augments and completes it. Stocker's and Morton's papers are both about the relationship of emotion to knowledge, and should be read back-to-back. Stocker elaborates on his previous suggestion that there are 'intellectual emotions', such as intellectual interest and intellectual courage; he argues that they are instrumentally valuable (although perhaps not necessary) for good intellectual work, and thus need to be recognized and encouraged. Morton goes further, and argues that what he calls 'epistemic emotions' (such as epistemic curiosity and epistemic worry) are *essential* to the acquisition and maintenance of knowledge. Finally, aside from Hobson's paper mentioned above, Part IV includes a paper by Rorty on the creative and even virtuous character of ambivalence. Her paper is not 'strictly speaking' about emotion (I do not think she mentions the word anywhere in the paper), yet ambivalence is clearly closely related to affectivity — we often have 'mixed feelings', and our attitude towards ambivalence is not neutral either, as we may feel uncomfortable or even guilty about our own ambivalence.

In sum, even though I would have liked to see more phenomenology in this Handbook, and more attention to current affective-scientific research and what philosophy can contribute to it, still it cannot be denied that this volume will be an irreplaceable research tool for any emotion researcher for quite some time. It shows clearly that emotion enters the philosophical inquiry in many different ways, and it does the emotion researcher a big favour by bringing together so many papers representative of these different ways. I myself will consult it often, and use it especially as an entry door into those topics with which I am less familiar. There is little unity in the philosophy of emotion, not just in the sense that existing accounts vary widely from one another (which is to be expected) and that philosophers from different traditions seem to have quite different emotions (!), but in the sense that there are few definite and agreed upon questions and approaches that constitute a common target for discussion and analysis. The major strength of this Handbook, I think, is that in putting all this material together it will contribute to a clearer overall picture, and thus to the emergence of a more unified field — one, however, that will doubtlessly always include many different styles, perspectives and accounts, consistent with the multi-faceted nature of its subject.

Acknowledgements

Many thanks to Julian Kiverstein for his advice. This work is funded by the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013), ERC grant agreement nr. 240891.

References

- Barrett, Lisa (2006), 'Are emotions natural kinds?', *Perspectives on Psychological Science*, **1**, pp. 28–58.
- Damasio, Antonio (1999), *The Feeling of What Happens: Body, Emotion and the Making of Consciousness* (London: Heinemann).
- Goldie, Peter (2000), *The Emotions: A Philosophical Exploration* (Oxford: Clarendon Press).
- Griffiths, Paul (1997), *What Emotions Really Are* (Chicago, IL: University of Chicago Press).
- Panksepp, Jaak (1998), 'The periconscious substrates of consciousness: affective states and the evolutionary strategies of the self', *Journal of Consciousness Studies*, **5** (5–6), pp. 566–582.
- Ratcliffe, Matthew (2008), *Feelings of Being: Phenomenology, Psychiatry and the Sense of Reality* (Oxford: Oxford University Press).
- Slaby, Jan (2008), 'Affective intentionality and the feeling body', *Phenomenology & the Cognitive Sciences*, **7**, pp. 429–444.
- Thompson, Evan & Zahavi, Dan (2007), 'Philosophical issues: phenomenology', in: *The Cambridge Handbook of Consciousness*, eds. P.D. Zelazo, M. Moscovitch & E. Thompson, pp. 67–88 (New York: Cambridge University Press).
- Zahavi, Dan (2004), 'Phenomenology and the project of naturalization', *Phenomenology and the Cognitive Sciences*, **3**, pp. 331–347.

Raymond Tallis

Michelangelo's Finger: An Exploration of Everyday Transcendence

London: Atlantic Books, 2010, 166 pp., £18.99 (hbk)

ISBN: 978-1-84887-119-9

Reviewed by Richard Moore

Max-Planck Institute for Evolutionary Anthropology, Leipzig

In *Michelangelo's Finger*, a non-academic philosophy book written for a non-specialist audience, Britain's foremost gerontologist-cum-philosopher examines the role of the index finger in hominid development. He sees a 'deep symbolic connection between this finger and the special nature of humans' (p. xv), which is illustrated in Michelangelo's depiction of the hand of God reaching out, forefinger extended, to Adam. It is this scene from the ceiling of the Sistine Chapel from which Tallis' book takes its name.

The function of the human forefinger that is of particular interest to Tallis is its ability to be extended and deployed in communicative pointing. He proposes an explanation of why humans alone can point, and sketches an account of the growth of civilization in which the act of pointing is seed: 'This cognitive universe does not begin with words... [but conceivably] with that family of gestures... whose paradigm example is pointing' (p. 26). It is the ability to point that separates us from our nearest ape ancestors, because pointing requires uniquely human abilities, but in turn permits the discovery of a 'deperspectivalised' objective world.

Tallis writes in a Merleau-Pontyesque jargon that can be eloquent but is often clumsy. This sometimes makes his arguments difficult to follow, but I will try to reconstruct them. His central thesis is that two conditions are necessary for pointing. First, those who point must be explicitly aware of their own bodies in a manner that enables them to use their bodies as signaling tools. This is a prerequisite of being able to produce intentional pointing gestures. Second, pointers must have an awareness of the different viewpoints of others, and a correlated ability to reflect on what others don't know. This is necessary because 'the fundamental occasion for pointing' (p. 11) is to inform. Without grasping that others can fail to know what we know, we would have no motive to point for them. It's because apes lack these two characteristics that they do not point, says Tallis.

Humans but not apes satisfy the first requirement on pointing because we alone know 'that we *are* our bodies' (p. 19). What this mysterious claim means is never quite clear: reading Tallis, one often has the sense of importantly different phenomena being lumped together. However, it seems to include at least that we understand ourselves as embodied agents situated in a world of mind-independent objects with which we can interact. From this we acquire 'the sense of agency, of the *explicit* sense of using one's own body to achieve *explicitly* entertained ends' (p. 21; my italics) that a pointer exploits in extending her finger communicatively. Because humans have this developed sense of agency, we can grasp the possibility of fashioning tools to aid us in the pursuit of our goals — first the possibility of using our own bodies as tools; and second the possibility of creating external tools with which to shape our environment to our will.

Tallis' empirical claims about apes' failure to grasp that others' perspectives can differ from their own, and about their lacking sense of agency, are frustrating. He states his thesis with confidence, but supports it with only a selective use of empirical studies — references are drawn almost exclusively from Povinelli (2003). Povinelli's views

about the cognitive abilities of our nearest ancestors are congenial to Tallis' own, but his claims have been contested with experimental data (some of which is reviewed in Tomasello & Call, 2006) and the existence of challenges is nowhere acknowledged.

Despite what Tallis says, existing evidence suggests that apes have a relatively developed sense of agency. Certainly, the phenomenon of tool use is well-established (Whiten *et al.*, 1999; van Schaik *et al.*, 2003). This includes evidence that apes use and fashion a variety of external tools. Furthermore, great apes often gather them before they need them (Mulcahy & Call, 2006) — suggesting that they plan ahead, a seemingly explicit act. They also produce gestures in pursuit of goals (Tomasello *et al.*, 1985; 1989) — using their own bodies as tools in purposive activity. We don't yet know how to characterise the mental processes that accompany these activities, but whether or not they are explicit in the manner that Tallis demands is an open question that cannot be answered with his confident 'no'.¹ If I've properly understood what Tallis means by a 'sense of agency', then apes might well possess one sufficient to enable them to point.

Tallis' claim that apes don't point because they lack an understanding of other minds is also unsatisfactory. For a start, while it's largely agreed that chimpanzees don't know, or care to know, the mental lives of their peers in the sophisticated ways that we do, there is good evidence that they do know that others cannot always see what they can see (Tomasello & Call, 2006; Call & Tomasello, 2008). To give just one example, subordinate chimpanzees will eat food that's hidden behind an opaque barrier, and so invisible to a dominant chimpanzee, but not food hidden behind a transparent barrier.

Since chimpanzees possess at least rudimentary forms of the abilities that Tallis thinks necessary for pointing, his answer to the question 'Why don't apes point?' cannot be right. A plausible alternative is that a further condition, missing in apes, is necessary for pointing. A candidate here is that while they do have some basic grasp of themselves as agents, and of the different perspectives of others, apes typically lack both the motivation to cooperate with others, and any expectation that others would do the same for them. After all, informative pointing implies a desire to help one's interlocutor, and

[1] Tallis also argues that dogs cannot understand pointing. His evidence for this claim is entirely anecdotal. However, there exists robust empirical evidence that some dogs excel at understanding (if not producing) pointing behaviour, even when chimpanzees cannot. Additionally, he thinks it unsurprising that dogs do not grasp pointing, since apes cannot. However, unlike apes, dogs have co-evolved with humans — such that grasping human communication would be selectively advantageous for them (Hare *et al.*, 2002).

grasping an informative point requires recognizing that another might be acting with a helpful motive. Absent a culture of cooperation, these phenomena might be both absent and difficult to comprehend. ('I'm hungry, but why's he holding his arm out?') says the anthropomorphised chimp to itself, unaccustomed to the possibility that anyone would show it where to find food.) This holds for imperative pointing too: Why point for out-of-reach food if there's no prospect of someone handing it to you? The comparative lack of a cooperative instinct in apes is consistent with the possibility that they lack any interest in engaging with others in the ways that humans do. As Tallis notes, infants of 12 months will often point to an object that its caregiver can see, so that the two of them might enjoy attending to it together (Liszkowski *et al.*, 2004). Similar desires seem to be absent in chimpanzees (Tomasello & Carpenter, 2005).

These ideas have all been discussed at length (for example, Tomasello, 2006; 2008) in a burgeoning literature on pointing with which Tallis barely engages. This is disappointing, because his timely book would have been a perfect opportunity to present to a wider audience some of the exciting work now being done by developmental and comparative psychologists. Much of this work would be congenial to the over-arching claims that Tallis wants to make about the role of pointing in the origins of human society.

In the second half of *Michelangelo's Finger*, following a speculative and insubstantial chapter on autism, Tallis turns to the role of pointing in language-acquisition. He wrestles with the idea that pointing might facilitate infant word-learning, through making possible the ostensive definitions of words. Recognizing that words don't map on to bits of the world to which speakers could unambiguously point, he argues against a purely Augustinian account of word-learning. However, he doesn't give up on the idea that pointing is an 'enormously important aid to the process of inducting a child into the common human cosmos' (p. 86). It is a shame that Tallis doesn't say more about how pointing might contribute to this process, since this chapter ends without any real positive claim.

After an entertaining but angry chapter on the rudeness of pointing *at* someone and a less entertaining discussion of signposts comes a final chapter in which Tallis addresses the 'transcendence' of the title. In his words, 'pointing has a key role in opening up a realm of pure possibility, of generality, beyond sense experience' (p. 127). There are two parts to this claim. First, says Tallis, pointing to an object creates a shared public space between the pointer and his interlocutor. This 'fundamental action of... making a world in common' (p. 132)

constitutes a first step in the building of a public sphere in which man finds himself, but of which he is not the centre. This idea is compelling, but it's far from obvious that pointing is the only action that could serve this function, even in the absence of language. Second, Tallis suggests that pointing makes possible conceptual thought — by which he means our ability to see objects as instances of general categories. There is no clear statement of this argument, but what Tallis seems to envisage is that — quite plausibly — informative pointing is a pre-cursor of assertion. It is not clear how he thinks the act of assertion relates to the possession of concepts, though. Additionally, I failed to understand how the two claims about transcendence relate to one another.

Michelangelo's Finger is not a satisfying book. Good science writing can be clear and precise without being inaccessible, but Tallis' claims are often under-theorised and his knowledge of the empirical studies relevant to his subject is weak. Nonetheless, he is passionate about his subject matter and writes with a contagious sense of intellectual curiosity. This, along with its wealth of interesting ideas, makes for a book that is fun. For all its flaws, I enjoyed grappling with it.

References

- Call, J. & Tomasello, M. (2008), 'Does the chimpanzee have a theory of mind? 30 years later', *Trends in Cognitive Sciences*, **12**, pp. 187–192.
- Liszkowski, U., Carpenter, M., Henning, A., Striano, T., & Tomasello, M. (2004), 'Twelve-month-olds point to share attention and interest', *Developmental Science*, **7** (3), pp. 297–307.
- Hare, B., Brown, M., Williamson, C., & Tomasello, M. (2002), 'The domestication of social cognition in dogs', *Science*, **298** (5598), pp. 1634–1636.
- Mulcahy, N.J. & Call, J. (2006), 'Apes save tools for future use', *Science*, **312**, pp. 1038–1040.
- Povinelli, D.J. (2003), *Folk Physics for Apes: The Chimpanzee's Theory of How the World Works* (Oxford: Oxford University Press).
- Tomasello, M., George, B., Kruger, A., Farrar, J. & Evans, A. (1985), 'The development of gestural communication in young chimpanzees', *Journal of Human Evolution*, **14**, pp. 175–186.
- Tomasello, M., Gust, D. & Frost, T. (1989), 'A longitudinal investigation of gestural communication in young chimpanzees', *Primates*, **30**, pp. 35–50.
- Tomasello, M. & Carpenter, M. (2005), 'The emergence of social cognition in three young chimpanzees', *Monographs of the Society for Research in Child Development*, **70** (1).
- Tomasello, M & Call, J. (2006), 'Do chimpanzees know what others see — or only what they are looking at?'. in: *Rational Animals*, eds. S. Hurley & M. Nudds, pp. 371–384 (Oxford: Oxford University Press).
- Tomasello, M. (2006), 'Why don't apes point?', in: *Roots of Human Sociality: Culture, Cognition and Interaction*, eds. N. Enfield & S.C. Levinson, pp. 506–524 (Oxford: Berg Publishers).

- Tomasello, M. (2008), *Origins of Human Communication* (London: MIT Press).
- van Schaik, C.P., Ancrenaz, M., Borgen, G., Galdikas, B., Knott, C.D., Singleton, I., Suzuki, A., Utami, S.S., & Merrill, M. (2003), 'Orangutan Cultures and the Evolution of Material Culture', *Science*, **299** (5603), pp. 102–105.
- Whiten, A., Goodall, J., McGrew, W.C., Nishida, T., Reynolds, V., Sugiyama, Y., Tutin, C.E.G., Wrangham, R.W. & Boesch, C. (1999), 'Cultures in chimpanzees', *Nature*, **399** (6737), pp. 682–685.

Alva Noë

Out of our Heads: Why You Are Not Your Brain, and Other Lessons from the Biology of Consciousness

New York: Hill & Wang, 2009, Hardcover

ISBN-13: 978-0-8090-7465-5

Reviewed by Julian Kiverstein

School of Philosophy, Psychology and Language Sciences, University of Edinburgh

Alva Noë's latest book calls for nothing short of a revolution in how scientists and philosophers think about consciousness. What Noë seeks to unseat is the idea of the brain as the seat of consciousness, the dominant view in the science of consciousness today. He proposes we put in its place a picture in which the biological basis of consciousness resides in a living body in dynamic interaction with its environment. Brains are of course necessary for consciousness — nothing could be conscious without one, but Noë argues that brain processes aren't sufficient for consciousness. Brains enable animals to interact dynamically with their environment, and it is in this interaction, not in the brain, that consciousness is to be found.

Noë's writing is elegant, imaginative and colourful. The metaphors he crafts are illuminating, and the stories he recounts, insightful. Some philosophers reading the book might crave more precision. Indeed if you are looking for careful, patiently crafted philosophical arguments this book is almost certainly not for you. However, Noë is clear in his preface that he intends to reach beyond the narrow confines of academia to readers fascinated by the wonders of consciousness, and daunted by attempts at understanding and explaining it (p. xiii). In this objective Noë succeeds admirably. In this short review, I've chosen to focus my attention on Noë's positive account of consciousness. Noë also mounts a number of arguments against reductive approaches to consciousness that would have us identify conscious experience with information processing occurring in the brain. It seems to me, however, that the plausibility of these arguments largely

rests on the case Noë's makes for rethinking our concept of consciousness.

Out of Our Heads continues a paradigm-shift taking place in contemporary cognitive science. In the early days of cognitive science minds were thought of as cold, calculating logic engines. Today this idea of mind is undergoing gradual replacement by an alternative vision of minds as embodied, embedded, extended, enactive and affective. There remains much disagreement in embodied cognitive science today over the exact details of this alternative vision, not least of all concerning the details of how embodiment should be understood. Noë is among the main players in this debate, and *Out of Our Heads* offers an accessible overview of his understanding of the embodied mind. In chapter 2 for instance we find the claim that an animal's mind consists in the ways in which it 'actively meshes with its environment and gears into it' (p. 42). Mind is identified with the bodily capacities an animal employs in engaging with the world according to its interests and needs. Noë goes on to tell us an animal's 'mind is its life' (*op cit*). This has the somewhat counterintuitive consequence that mind reaches all the way down the tree of life to the simplest of creatures. A bacterium, for instance, qualifies as an agent insofar as it is a living being with its own interests and needs. It has a relationship with its surroundings in which certain features stand out as salient, and this, according to Noë, suffices for it to have a mind.

Noë's view of mind as continuous with life is given further flesh in chapter's 3–5. Chapter 3 contains a wonderful discussion of the brain's plasticity, in which Noë explains how cognitive development can be understood as the process by which we come to be at home in the world, rooted in a practical environment. In early development, between birth and early adulthood, the functional and structural organisation of the brain is strongly dependent upon sensory and social stimulation. Brains grow and shape themselves in response to recurring features of their ecological and social setting; they are literally constructed to fit the environments with which we regularly deal. Indeed, Noë thinks that such is the fit between the internal neurocognitive structures of the brain, and the external environment that the boundary separating the internal from the external, the brain from the non-neural body and environment, no longer holds up.

A critic might well object that it is the intrinsic structural and functional organisation of the brain that constitutes the biological basis of cognition and experience. (See for instance, Ken Aizawa's, 2010, discussion of Noë on neural plasticity.) The history of the animal's causal interaction with the environment may be important for explaining

how a brain comes to have the right kind of functional organisation to support experience. This however establishes only a causal dependence of the brain on the environment, and doesn't on its own warrant any relocation of the mind–world boundary.

Noë might reasonably respond that this misses the tightness of the fit between the internal physiology and the external environment. As an illustration, he briefly discusses the experience of immigrants to a new culture who find themselves in an unfamiliar environment that radically differs from the one in which they grew up. (Noë is drawing on Bruce Wexler's fascinating work here, see chapter 4 of Wexler, 2006.) Immigrants find themselves equipped with internal neural structures that no longer match their external environment, and this mismatch necessitates a change in those internal structures that can be lengthy and painful. The children of immigrant families (because of the plasticity of their brains) often find it easier to adapt to their new environment, creating conflict and additional distress for their parents. Noë writes: 'At a very basic level, we are *involved* — that is to say, tangled up — with the places we find ourselves. We are *of them*' (p. 69). We are entangled with our environments in a way that undermines the distinction of inner from outer because our brains have a functional organisation that is moulded to match the environment in which they've developed.

The theme of mind as life continues in chapter 4 where Noë takes up the question of where the boundaries of the self end, and the rest of the world begins. He begins with an interesting discussion of the rubber hand illusion and phantom limbs. Both of these phenomena illustrate how a person can feel a sense of ownership for a body part that is not attached or connected to their body. Noë suggests it is the integration of my body parts into my sensory and motor interactions with the world that gives me the feeling that a body part is mine. Phantom limb patients continue to feel an amputated limb is a part of their body because their 'behavioural, environment-involving attitudes and engagements outlive the loss of the limb' (p. 76). In a similar fashion a tool that is located in my external environment can be incorporated and feel like it is a part of me when it becomes integrated into my body schema. The latter comprises my practical knowledge of what I can do with my body. A tool becomes part of me when my knowledge of how to use the tool is integrated into the range of action possibilities I bring to bear in my dealings with the world.

Tools don't just extend the range of possibilities for action open to us. They can also act as props that scaffold our thinking enabling us to solve problems, which without the aid of 'mind-tools', might well

prove beyond us. Noë is in full agreement with Andy Clark that the causal processes that support human cognition aren't confined to the processes that take place inside of our skulls. Consider in this light his discussion of language: Of the tools that augment our thinking, language is arguably the 'ultimate artefact' (Clark, 1997, ch.10). Noë offers a convincing argument that we shouldn't think of language as a system of rules that tell us how to use words to make true statements about the world. Rather, understanding language is a matter of being able to participate in a social practice. Linguistic knowledge is, Noë suggests, a set of abilities. These abilities are manifest in the flows of activity that make up our lives, which Noë takes to be definitive of mind.

The central claim of chapter 5 is that human beings are creatures of habit. Adult humans spend much of their everyday lives dealing with situations expertly. Examples include talking, reading, cooking, parenting, driving and walking. These actions are for the most part immediate, unreflective responses to a situation. They are radically unlike the behaviours of a computer or robot that is designed to first build up a model of the world, form a plan of action based on this model, and finally work out how to carry out its plan. These behaviours are the actions of experts that have built up practical knowledge of how to fluently and effortlessly cope with situations regularly and repeatedly dealt with in the past. Consider again language as an illustration of the place of habits in our lives, and contrast the fluent speaker of a language with someone starting out on learning a new language. When learning a language we learn the rules that govern the use of words, their declension, conjugation and combination with other words. We devote attention to the language itself. Once we can fluently speak a language our attention is no longer focussed on language, but on what we are using the language to say. We no longer need to concern ourselves with the rules of the language because they have been mastered, and we can follow them as a matter of habit. A fluent speaker of a language will very rarely face the problem of having to assign a meaning to an utterance. As we become fluent speakers of a language so we become familiarised with the situation-specific uses of language until this becomes second nature.

It might be thought that Noë has moved a long way from the topic of consciousness with his excursus on the place of habits in human life. Habits are after all behaviours that can be performed unthinkingly and so we might naturally suppose unconsciously. Noë, however, takes there to be an important connection between habits and consciousness: It is thanks to our habitual ways of acting that the world shows

up for us in consciousness. In the book's epilogue Noë tells us: 'Our relation to the world is not that of an interpreter. The meaningful world is there for us, understood, before interpretation gets its start' (p. 184). Our consciousness opens out onto a meaningful world, however, only because of our skills and habits. We inhabit a familiar environment in which we feel at home because of our skills and habits. A creature without habitual ways of acting would be alienated from the world and would be forced to adopt an interpretative stance to make sense of its environment. This is, however, not the position we find ourselves in.

This leads me to what is perhaps a more familiar theme in Noë's philosophy for those who know his previous writings. Noë defends what is called an *enactive* theory of conscious experience as a skilled activity. According to the enactive theory perceptual experiences aren't passive occurrences that occur inside the brain. Perceptual experience is an active process in which the perceiving animal uses its sense organs to probe and explore the world. This exploratory activity is *skilful* insofar as the perceiver draws on his familiarity with patterns of dependence that hold between bodily movements and changes in sensory stimulation. A simple example is looming: As the perceiver approaches an object, so the area of the visual field the object occupies will increase. Insofar as all perception involves this kind of familiarity every perceiver is an expert, and we exercise our expertise in perceiving the world.

In chapter 3, Noë explains how the enactive theory might go about explaining the qualitative character of experience. The character of experience is due to the ways in which 'sensory stimulation varies as a function of movement in relation to the environment'. The brain clearly has a central role to play in enabling this sensorimotor interaction with the world, but it doesn't generate experience of the world all on its own. Noë offers a useful analogy when he compares the brain to a musical instrument: 'Instruments don't make music or generate sounds on their own. They enable people to make music or produce sounds' (p. 64).

It should by now be becoming clear how much of the science of consciousness is challenged if we buy the story about conscious experience Noë recounts in *Out of Our Heads*. A prominent strand of research in the science of consciousness aims at identifying the neural correlates of (visual) consciousness: 'The minimal set of neuronal events and mechanisms jointly sufficient for a conscious percept' (Koch, 2004, p. 16). If Noë is right, there aren't any neural correlates of consciousness, since there are no neural events and mechanisms

that suffice for conscious experience. A *leitmotif* that occurs at regular intervals throughout the book is the claim that ‘consciousness isn’t something that happens inside us: it is something that we do, actively, in our dynamic interaction with the world around us’ (Noë, 2009, p. 24). One might reasonably wonder in response to such a strong claim about dreams. Surely in this case it is not the animal’s dynamic interaction with the world that explains the experiences we undergo, but the neural events and mechanisms instantiated in the dreamer’s brain at the time. What of the many anecdotal reports one frequently encounters of neuroscientists eliciting perceptual experience through the direct stimulation of cortical areas? Surely in these cases it is direct stimulation in conjunction with various background conditions that is explanatory of experience? Noë calls this the *Foundation Argument* because it seems to provide the strongest support for the view that conscious experience is exhaustively a neural phenomenon. Clearly Noë is not persuaded, and he offers a number of arguments aimed at undermining the Foundation Argument in chapter 8 of the book. The success of *Out of Our Heads* will largely depend on how persuasive readers find the arguments of this chapter.

Out of Our Heads is a delight to read, no matter what you make of its arguments. True to his vision of consciousness, Noë offers an integrative view of what a science consciousness must look like: It is one that effortlessly weaves together philosophy, psychology and developmental neuroscience. Many readers will no doubt fail to be persuaded by the book’s radical message. However, even the most sceptical of reader will find it hard to resist rethinking some deep-rooted assumptions about what consciousness is, and how science might go about explaining it.

References

- Aizawa, K. (2010), ‘Consciousness: Don’t give up on the brain’, in: *The Metaphysics of Consciousness*, eds. P. Basile, J. Kiverstein & P. Phemister (Cambridge: Cambridge University Press).
- Clark, A. (1997), *Being There: Putting Brain, Body and World Together Again* (Cambridge, MA: MIT Press).
- Koch, C. (2004), *The Quest for Consciousness: A Neuroscientific Approach* (Englewood, CO: Roberts & Company).
- Wexler, B. (2006), *Brain and Culture: Neurobiology, Ideology and Social Change* (Cambridge, MA: MIT Press).