

# **The Subjective Brain**

Pete Mandik

Draft 6.0

## **Table of Contents**

Ch. 0. Introduction: Consciousness and the Invisible Brain

### Part One: Metaphysics Groundwork

Ch 1. The Metaphysics of the Neuron

Ch 2. Introspecting Brain States as Such

Ch 3. Beware the Unicorn: Consciousness, Intentionality, and Inexistence

### Part Two: Explaining Consciousness

Ch 4. The Neurophilosophy of Consciousness

Ch 5. Transcending Zombies

Ch 6. The Allocentric-Egocentric Interface Theory of Consciousness

### Part Three: Objective and Subjective Representation

Ch 7. Animate Semantics

Ch 8. The Neural Accomplishment of Objectivity

Ch 9. Consciousness Without Subjectivity

Ch 10. Swamp Mary's Revenge

List of Figures

References

## **Ch. 0. Introduction: Consciousness and the Invisible Brain**

We don't need science in order to know that we are conscious or aware. But science suggests that we are conscious largely due to our brains. And science perhaps further suggests that our conscious states are identical to certain states of our brains. This later idea—that conscious states just *are* certain brain states (hereafter, the Identity Thesis)—is the main aim of this book to defend.

It is worth pausing a bit to appreciate what's surprising or problematic about the Identity Thesis. If seeing is believing, then we can expect that the Identity Thesis suffers a problem of credibility, since we seldom have experiences that we would naturally describe as experiences *of* our brains. However, being locked in an opaque container behind our eyes is the least of the causes of the problem.

### **1. Three Sources of Invisibility**

There are three main ways that the topic of conscious states may be introduced, and each of the three ways arguably raises problems for the Identity Thesis. The three ways of introducing conscious states are

(way 1) as states for which there is *something it's like* to be in them,

(way 2) as states *of* which we are conscious, and

(way 3) as states *with* which we are conscious.

The three problems for the Identity Thesis raised in turn are

(problem 1) it is not clear how brain states can be the sorts of states for which there is something it's like to be in them. Worse, for each possible brain state, it seems possible to be in that brain state without there being anything it's like.

(problem 2) states of ourselves of which we are conscious don't seem particularly brainy. While we are conscious of obviously *bodily* states like a swollen ankle and obviously *mental* states like a desire to get a drink of water, we are not conscious of any states obviously *brainy*.

(problem 3) states with which we are conscious are typically states largely characterized in terms of properties in an environment external to, among other things, our brains.

I will spell out these three ways and three problems in further detail in the next three sections.

## **2. What it's Like**

The first way of introducing the topic of conscious states is the one that has raised the most discussed and perhaps the biggest problems in the recent literature. This way of raising the topic of consciousness is in terms of states for which there is *something it's like* to be in them. There is something it's like for me to see roses as well as something it's like for me to taste lemons. However, there is nothing it's like for a rock to be warmed by the sun or soaked by the rain (yet alone for a rock to see roses or taste lemons). The so-called "hard problem of consciousness" that has taken up so much of the recent literature is the problem of explaining why there should be anything at all it's like to be a creature with brain states like mine (Chalmers 1996).

Related to the problem of explaining why certain of my brain states are like anything at all is the problem of explaining why they are like *this*. If we can indeed conceive of inverted spectra—cases in which someone physically similar to me differs in that what it’s like for them to see red is more like what it’s like for me to see green and vice versa—then we can conceive of beings for whom there is something it is like to be, but despite their physical similarity to us, what it’s like to be them is very different from what it’s like to be us. Why, then, is the way some mental state phenomenally appears to me *that* way and not some other?

Insofar as the reader can grasp the discussion of the previous two paragraphs, they are in a position to grasp what it means to introduce the topic of consciousness in terms of *what it’s like*. Introducing consciousness in such terms involves commitment to what I shall call ‘the what it is like principle’ or just WIL:

(WIL): A conscious state is a state of a creature in virtue of which there is something it’s like to be that creature.

WIL is very closely related to Nagel’s view that “an organism has conscious mental states if and only if there is something that it is like to *be* that organism” (1974 436). The core notions involved in WIL are also closely related to the notions of the phenomenal character of conscious states or their *qualia*.

It is difficult to say uncontroversial things about qualia; however, there are several points of agreement among many of those philosophers that believe that mental states have such properties (though it should be noted that I do not agree with such philosophers

on all of the following points). These philosophers describe qualia as (i) intrinsic properties of conscious states that (ii) are directly and fully knowable only by that subject and (iii) account for “what it is like” for a subject to be in that state. More briefly, qualia are (i) intrinsic, (ii) subjective, and (iii) such that there is “something it’s like” to have (states with) them. Less briefly, we can start with (iii) and work our way to (i) as follows. When I have a conscious perception of a cup of coffee there is, presumably, something it’s like for me to have that perception and, for all I know, what it’s like for you to have a conscious perception of a cup of coffee is quite different. Further, for all that you can tell me about your experience, there is much that cannot be conveyed and thus is subjective, that is, directly and fully knowable only by you alone. The supposition that qualia are intrinsic properties of conscious states serves as a possible, though questionable, explanation of their subjectivity. The inference from subjectivity to the intrinsic nature of qualia may be articulated as follows. If something is defined by the relations that it enters into, then it is fully describable by the relations it enters into and if it is not fully describable by the relations it enters into it must not be defined by the relations it enters into.

One way of putting a point on the problems that qualia, subjectivity, and what it’s like raise for the Identity Thesis is through the Knowledge Argument of Jackson (1982) which is itself built on certain features of conscious experience remarked upon by Nagel (1974), who famously posed the question of what it’s like to be a bat. The answer that Nagel urged is that bat experience must be insufficiently similar to our own for us to know what it’s like to be a bat. Jackson's knowledge argument attempts to use the subjectivity of conscious experience as a premise in an argument that certain aspects of

conscious experience must be non-physical. In brief, the knowledge argument begins by supposing it possible for someone to know all of the physical facts without ever having had a conscious experience of seeing red objects. It is further supposed that such a person would not, then, know what it's like to see red. Assuming further that such knowledge—knowledge of what it is like to see red—is factual knowledge, it follows then (from this assumption and the previous suppositions) that this must be knowledge of a non-physical fact.<sup>1</sup>

### **3. What one is conscious of**

Another way of introducing what conscious states are is as mental states that we are conscious of having. We can put this in terms of what has been called the Transitivity Principle, and what I shall call simply, Transitivity.

(TRANSITIVITY): A state is conscious only if one is conscious of this state.<sup>2</sup>

---

<sup>1</sup> The Knowledge Argument is one of several “epistemic gap” arguments against physicalism (and thus, the Identity Thesis) that involve introducing consciousness in terms of “what it’s like”. For a review, see Chalmers, D. (2003). *Consciousness and its Place in Nature*. Blackwell Guide to Philosophy of Mind. S. Stich and F. Warfield. Oxford, Blackwell.

<sup>2</sup> See, e.g., Rosenthal, D. (2005). Consciousness and Mind Oxford, Clarendon Press. p. 3.

Transitivity is closely associated in the literature on consciousness with Higher-Order Representation theories of consciousness—HORs—which are theories that seek to explain a mental state’s being conscious in terms of the mental state’s being represented by another mental state (see, e.g., (Lycan 1996) , (Rosenthal 2005)).

One of the main lines of reasoning in favor of HORs is that such theories constitute proposals for how Transitivity is implemented. HOR theorists regard representation as a way in which someone can be conscious of something. Being conscious of something is (or is a kind of) representing something. Thus Transitivity’s requirement on conscious states that their possessors be conscious of them is implemented by HOR’s proposal that conscious states are ones that are appropriately represented.

The main problem that Transitivity raises for the Identity Thesis is that it is far from clear that when we are conscious we are thereby conscious of our brain states. However, at least we are conscious of something inside of us according to Transitivity and this gets us close to the brain. We move even farther away from the brain when we focus on the third way of introducing consciousness.

#### **4. What one is conscious *with***

Yet another way of introducing consciousness is in terms of the Transparency Thesis or, as I shall call simply, Transparency.

(TRANSPARENCY): When one has a conscious experience all one is conscious of is what the experience is an experience of.

Transparency is closely associated in the literature on consciousness with First-Order Representation theories of consciousness—FORs—which seek to explain a property’s being phenomenal in terms of the property being represented in experience (see, e.g., (Dretske 1995) and (Tye 2000)). Tye and Dretske embrace the wide-spread view that phenomenal properties are those properties in virtue of which there is something it’s like to have conscious states. Like Rosenthal, Tye and Dretske interpret ‘conscious of’ as indicative of representation: being conscious of something involves mentally representing something. Thus, according to FOR, the properties determinative of what it’s like to be in an experiential state are the properties represented by the state. When experiences are veridical, the properties determinative of what it’s like just are the properties of the objects as they are correctly perceptually represented (Tye 2000). So, for example, as Dretske (1995) puts it:

[Q]ualia are supposed to be the way things seem or appear in the sense modality in question. So, for example, if a tomato looks red and round to S, then redness and roundness are the qualia of S’s visual experience of the tomato. If this is so, then ... if things ever *are* the way they seem, it follows that qualia, the properties that define what it is like to have that experience, are exactly the properties the object being perceived *has* when the perception is veridical (pp. 83-84).

Thus are qualia a certain kind of “represented properties,” that is, qualia are defined as “phenomenal properties—those properties that...an object is sensually represented...as having” and as properties *not* of the experience itself (Dretske 1995 73).

Regarding this latter point, that phenomenal properties are *not* properties of experiences, Tye writes:

Visual phenomenal qualities or visual qualia are supposedly qualities of which the subjects of visual experiences are directly aware via introspection. Tradition has it that these qualities are qualities of the experiences. Tradition is wrong. There are no such qualities *of experiences* (Tye 2000 49).

What FORs are theories of, then, is the second-order property of being phenomenal. What distinguishes phenomenal properties from non-phenomenal properties is that only phenomenal properties are represented in a certain way. A ripe tomato has lots of properties, but when one of them gets represented in a certain way, it goes from being a mere property to being a phenomenal property. When I correctly represent in experience the redness of a red tomato, the property determining what it’s like to have this experience is a property *of* the tomato—the redness—and it (the redness) takes on the second-order property of being phenomenal by being represented in a certain way. More precisely, for FORs, being phenomenal just is the property of being represented in a certain way.

We can distill from the above remarks the following key ideas of FORs: (1) Some properties have the property of being phenomenal (that is, the property of being

determinative of *what it's like* to have a conscious experience). (2) It follows from Transparency that the only properties we are conscious of when we have a conscious experience are properties that the experience is an experience *of*, not properties of the experience *itself*. (3) One comes to be conscious of properties in ways relevant to *what it's like* by having an appropriate experiential representation of those properties.

The problem posed by Transparency for the Identity Thesis is quite clear (pun intended): if phenomenal properties aren't in the head, then they can't be identical to any neural properties.

## 5. Three Questions of Consciousness

One way of viewing the different ways of introducing conscious states and the different theories that typically rely on them is in terms of three core questions concerning consciousness:

- (1) **The question of phenomenal character**—the question of what makes a conscious state have “something it’s like”-ness,
- (2) **The question of state consciousness**—the question of what it is that makes a mental state a conscious mental state, and
- (3) **The question of transitive consciousness**—the question of what it is that we are conscious *of* when we have a conscious mental state.<sup>3</sup>

---

<sup>3</sup> The terms “transitive consciousness” and “state consciousness” are due to David Rosenthal. For discussion, see Rosenthal, D. (1993). "State Consciousness and Transitive Consciousness " Consciousness and Cognition 2(4): 355-363.

Different theories of consciousness may place different emphases on each of the three questions. But various theories can agree that all three questions need to be answered.

The question of state consciousness concerns in what consists the difference between mental states that are conscious and mental states that are unconscious. Mental states vary with respect to whether they are conscious. Consider, for example, your memory of your mother's name. You may have had that memory for years but until you read the previous sentence it is unlikely that it was a *conscious* memory for the entire time between its initial acquisition and its current retrieval. In what does the difference between conscious and unconscious mental states consist?

The question of transitive consciousness concerns what it is that we are conscious *of*. When one has a conscious state, typically, if not always, one is conscious *of* something, as when I am conscious of a buzzing insect. Things may vary with respect to whether I am conscious of them, as when I am only intermittently conscious of the conversation at a nearby table in a restaurant. What does it mean to be *conscious of* something?

The question of phenomenal character concerns the so-called qualia of conscious states. Conscious states have certain properties—their phenomenal character—properties in virtue of which there is “something it is like” to be in that state. When I have a conscious perception of a cup of coffee there is, presumably, something it is like for me to have that perception and, for all I know, what it is like for you to have a conscious perception of a cup of coffee is quite different. What makes a conscious state have “something it's like” to be in that state?

Given the centrality of these questions, we will have several occasions to return to them throughout the book. In brief summary they are:

**The Question of State Consciousness:**

In what consists the difference between mental states that are conscious and mental states that are unconscious?

**The Question of Transitive Consciousness:**

When one has a conscious mental state, what is one thereby conscious of and how? (That is, *how* does one come to be *conscious of* something?)

**The Question of Phenomenal Character:**

When one has a conscious state, in what consists the properties in virtue of which there is something it is like for one to be in that state?

## 6. Neurophilosophy

Traditional philosophical issues that consciousness raises involve the relation of consciousness to the rest of the world, especially as that world is conceived of by the natural sciences. Thus much philosophical discussion concerns whether the world as conceived of by physical theory can adequately accommodate consciousness or if instead we are left with a dualism that cleaves reality into, for example, a non-physical consciousness and a physical everything else. Even among philosophers that agree that consciousness is consistent with physicalism, there is much disagreement, for there are several proposals for how best to spell out the consistency of a physicalistic world-view that makes room for phenomenal consciousness. One way of portraying this cluster of issues is in terms of which natural science is best suited to study phenomenal

consciousness and how to conceive of the relation between that science and the sciences involving the most basic aspects of reality (the physical sciences). One major view is that psychology is the proper science for understanding phenomenal consciousness and further, that psychological investigation of consciousness should be regarded as autonomous from sciences such as the neurosciences. In opposition is the view that the proper science is neuroscience and whatever contributions come from psychology are only valid insofar as psychological theories are reducible to neuroscientific theories. Increasingly, proponents of the latter view identify themselves as practitioners of *neurophilosophy*.

Neurophilosophy is a sub-genre of naturalized philosophy—philosophy that embraces Quine’s (1969) vision of philosophy as continuous with the natural sciences—wherein the natural science in primary focus is neuroscience. It is perhaps worth addressing here in further detail what is distinctive of neurophilosophy as opposed to other kinds of naturalism. The role that neuroscience plays is, of course, key, but not just any mention of the brain in a philosophical theory will suffice to make it neurophilosophical. Neurophilosophical appeals to neuroscience involve explicit and detailed use of contemporary neuroscientific literature. Further, neurophilosophy is not to be distinguished from other forms of naturalism by the philosophical *conclusions* that might be reached but by the role that contemporary neuroscience plays in the *premises* of the arguments for those conclusions. These points about different styles of naturalistic philosophizing may be illustrated in terms of some recent examples. For example, Jaegwon Kim is a kind of naturalist and even advocates a reduction of mental state types to physical state types. However, he is not thereby a neurophilosopher. His identification

of the relevant physical state types makes no explicit reference to contemporary neuroscientific findings. The state types in question involve no familiarity with the typologies specific to either neurophysiology or neuroanatomy. In contrast, the research of a neurophilosopher like Kathleen Akins makes explicit reference to contemporary neuroscientific findings in the arguments for various naturalistic conclusions. In her (1996) she argues against traditional views of the role that sensory states play in grounding the contents of intentional states. Crucial to her arguments are detailed examinations of the neurophysiology of thermoreception.<sup>4</sup>

Some authors draw a distinction between neurophilosophy and philosophy of neuroscience wherein the former involves the application of neuroscientific results to topics of philosophical concern (usually in the philosophy of mind) and the latter is a sub-discipline of the philosophy of science.<sup>5</sup> Though often neurophilosophers are also philosophers of neuroscience (see for example the Churchlands, Akins, Grush, Mandik), the current book involves the activities distinctive of the former group.

The term “Neurophilosophy” entered philosophical parlance with the publication of Patricia Churchland’s *Neurophilosophy* (1986), the aims of which were to introduce neuroscience to philosophers and philosophy to neuroscientists with an emphasis on the former.

---

<sup>4</sup> For a longer discussion of examples of neurophilosophical work such as Akins’s, see Bickle, J., P. Mandik, et al. (2006) *The Philosophy of Neuroscience*. [The Stanford Encyclopedia of Philosophy \(Spring 2006 Edition\)](#), .

<sup>5</sup> See Bickle, Mandik, & Landreth, 2006.

Patricia Churchland and husband Paul Churchland are paradigmatic examples of neurophilosophers. It is worth emphasizing that their work is primarily philosophical. Their professional training is primarily philosophical, their appointments are in philosophy departments, and they publish in philosophy journals. Because of this, neuroscience and philosophy do not have equal influence over neurophilosophy, at least in its paradigmatic instances. Instead, the primary forces that drive its development as an academic pursuit emanate from conventions of philosophical institutions. Thus, neurophilosophical work on consciousness proceeds largely by bringing neuroscientific theory and data to bear on philosophical questions such as the three questions of consciousness.

The question arises, of course, of what motivates the neurophilosophy of consciousness. The primary answer is that neurophilosophy has a certain appeal to those with an antecedent belief in physicalism in that neurophilosophy seems especially well suited to bridge the gap between consciousness and physical phenomena. Attempting to bridge the gap by reducing consciousness all the way down to chemistry or microphysics may strike many as too far a distance to traverse. More plausible is to seek a connection to a higher-level physical set of phenomena as offered in biology. Of the biological phenomena, the most plausible candidates are neural. Of course, such claims need the support of further argument. The remainder of the book is to fulfill this need.

## **7. The Plan of the Book**

In chapter 1, I argue from general physicalistic premises to a conclusion favoring a form of reductive physicalism. I argue further that the physical reduction base for the

mental will be neural. The rest of the book takes up the project of showing how such a view can handle consciousness as well as intentionality (at least insofar as intentionality is relevant to consciousness). The main challenge to Identity Theory posed by consciousness is that consciousness is alleged to have certain epistemic properties that neural states do not. The main challenge posed to Identity Theory by intentionality is that intentionality is alleged to essentially involve relations to things external to the nervous system in ways that neural states do not.

At least two of three main intuitive principles for introducing the topic of consciousness, Transparency and Transitivity, seem to indicate that consciousness is to be explained in terms of intentionality. The third, WIL, leaves the relation between consciousness and intentionality somewhat open.

In chapter 2, I address the question of what sorts of properties introspection gives us access to. Irreducible properties having been ruled out in chapter 1, the options remain that the properties in question are either neural, or, as Transparency and the closely related FORs are alleged to show, external. I argue that, contrary to Transparency, brain states may be introspected *as* brain states. Such a view is compatible with Transitivity and the closely related HORs.

In chapter 3, I examine the relation of intentionality to consciousness by presenting an argument—the Unicorn Argument—against both HORs and FORs. The Unicorn Argument raises problems for representational theories of consciousness based on our ability to consciously mentally represent things that do not exist. I conclude chapter 3 with the suggestion that the appropriate theory of consciousness, if

representational, must distance itself from both Transparency and Transitivity and rely instead on WIL for introducing the topic of consciousness.

The next three chapters, chapters 4,5,and 6, involve the development of a neurophilosophical theory of consciousness that satisfies the desiderata set forth in the preceding chapters. In chapter 4 I canvas the prospects for a neurophilosophical theory of consciousness and review the main features of some current proposals. In chapter 5, I sketch the outlines of an epistemological theory of consciousness adequate for dealing with various epistemic features of consciousness, namely the sorts cited in discussions such as those surrounding the topic of subjectivity and the Knowledge Argument. In chapter 6, I present my own neurophilosophical theory of consciousness: the Allocentric-Egocentric Interface theory of consciousness (AEI). The crux of AEI is to identify conscious states as states that involve causally interacting pairs of representations where one member of the pair is a subjective or egocentric representation and the other member is an objective or allocentric representation.

The discussion of consciousness in chapters 4, 5, and 6 make use of a not-fully-spelled-out notion of representation, and the question remains open whether this notion is consistent with the constraints on physicalism and consciousness spelled out in chapters 1, 2, and 3. The aim of chapters 7, 8, and 9 is to further flesh out an adequate account of representation. Chapter 7 addresses the project of understanding the simplest and earliest forms of representation as well as the project of relating such forms to the types of representation present in normal adult humans. Such early forms are egocentric and action-oriented. I address the issues of more advanced, that is, objective or conceptual

representations, in chapter 8. Chapters 7 and 8 flesh out the accounts of egocentric and allocentric representations brought up in connection with chapter 6's discussion of AEI.

Such accounts also set the stage for the final two chapters, wherein I return to the topic of consciousness and, building on the material from previous chapters, address the topic of whether the subjectivity of consciousness bars an objective understanding of consciousness in terms of brain states. I argue that, with an appropriate understanding of the neural bases of experiential and conceptual representation, we can see that conscious experience is *not* essentially subjective and thus does not constitute an ultimate obstacle for the Identity Thesis. The subjective brain—the mind as viewed from both inside and out—turns out not to be subjective in the strongest sense of that term.