

# Reflecting on Neural Self-reflection: The Churchlandish Introspection of Brain States

Pete Mandik, William Paterson University

One of the most intriguing ideas surfacing at various places in Paul Churchland's work, *Matter and Consciousness* counting among them, is the suggestion that we might one day achieve an ability to introspect our own brain states as brain states. The suggestion is that upon this achievement, when one introspects, one's own brain states will *seem* like brain states. In *Matter and Consciousness*, the suggestion appears in the book's final section, "The Expansion of Introspective Consciousness," and builds on material from the book's fourth chapter, especially the section "The Problem of Self-Consciousness."

I have two aims in what follows, each of which I will try to be quite brief about: (1) spell out the view and (2) spell out some of its key consequences.<sup>1</sup>

## 1. The View

Central to the view is that there are certain core analogies that can be drawn between perception and introspection. Further, the view is highly influenced by Churchland's view of perception, so let's start there.

Perceiving is perceiving-as. For instance, I perceive the object to my left as my iPhone. What I perceive things as is a matter of how I conceive of them, what concepts I apply to them. And this in turn is a matter of what theories I subscribe to. (What subscribing-to amounts to aside from believing is interesting, but I set it aside. Churchland is famously eliminativist about beliefs, but I don't see this as essential to the present view. So subscribing might as well just be treated as believing for present purposes.)

Perception is conceptual. However, it is distinguishable from other exercises of conceptual capacities. One distinguishing mark is how, in perceiving this thingie as an iPhone, my concept of an iPhone (and supporting theory of what iPhones are) is applied *automatically*. Getting to a point where this is done automatically is a matter of acquired skill. Initially, there's a period of non-automatic, deliberate application. I see a dark rectangular thingamabob, and consciously work through a series of hypotheses of what it might be. A sequence of hypothesis testing against a background theory never goes away, but does eventually become something that can be done nondeliberately and

---

<sup>1</sup> A much longer, though not as up-to-date, discussion along these lines may be found here: Mandik, Pete. (2006) [The Introspectability of Brain States as Such](http://www.petemandik.com/philosophy/papers/introspect.pdf). In: Keeley, Brian (ed.) <i>Paul M. Churchland: Contemporary Philosophy in Focus</i>. Cambridge: Cambridge University Press. pp. 66-87.  
<http://www.petemandik.com/philosophy/papers/introspect.pdf>

unconsciously. Perception is the automatic application of a concept in response to a stimulus, an application that exploits information carried by the occasion of stimulation.

Part of what makes some application of a concept *perceptual* is what the concept is a concept of. If the concept concerns external world entities, then my conceptual response upon being stimulated is perceptual. In contrast, if the concept concerns my own state of response, a state of my mind or brain, then the conceptual application is a candidate for being an occasion of introspection. Carrying through with the analogy, the introspective application of the concept needs to be automatic. It doesn't count as introspection if, after a chain of deliberate and conscious thinking, I arrive at the conclusion that I'm enjoying a certain activation-pattern in a neural population in cortical area M4. Such a conclusion needs to be automatic. Just as musicians and taste-testers can be trained to be expert perceivers, automatically applying bodies of theory to tease out sensible subtleties of their environment that the rest of us may be merely *cognitively* aware of, so can we mere mortals eventually become expert introspectors, armed with a powerful body of theory concerning our inner worlds.

## 2. Key Consequences of the View

### 2.1. Theory-ladenness and non-neutrality.

The Churchlandish view of introspection is part of a larger view in the philosophy of science that is highly suspicious of any kinds of alleged neutral arbiters between competing theories. Perceptual and other probes of nature are themselves laden with theory, so there's no neutral foundation upon which theorizing takes place. What counts as what's perceived is as much open to revision and replacement as any other part of Neurath's boat. This sort of point is no less true when applied to the question of what's introspected. If, upon introspecting one's own conscious sensations, the sensations don't seem like brain states, if they seem instead, for instance, like intrinsic properties irreducible to structural or functional aspects of the brain, then all this indicates is one's prior tacit acceptance of a theory that says conscious sensations are in that way irreducible. What counts as the correct way to introspect one's own states is answerable in a holistic way to a wide range of theoretical considerations. There's no theory-neutral way of isolating, even in one's own case, "what it's like" to be in such-and-such state.

### 2.2. Fallibility.

It is only in contrived and uninteresting cases that theory is infallible. Perhaps examples include the theory that there is at least one theory, or the theory that something or other has happened. (This latter example calls to mind the C.S. Peirce quote, "It is easy to be certain. One only has to be sufficiently vague.") Part and parcel of the theory-ladenness of introspection is the view that there's no *general* reason to expect the deliverances of introspection to be infallible. Of course, this leaves open the possibility that there might

be some specific examples that are infallible. But clear examples are trivialities, like my introspective judgement that I'm judging, or totally vacuous, like my introspective judgement that "this is this".

### 2.3. Transparency?

It is sometimes claimed that perceptual experience is transparent in the sense that in perceiving a tree, one cannot be aware of any aspect of the experience itself, but can only be aware of aspects of the tree. The Churchlandish model of introspection is clearly incompatible with this sort of transparency claim, for it holds that, in perceiving a tree, I might have a co-occurring introspective awareness not of the tree, but of the neural activation patterns constitutive of my awareness of the tree. In a [recent discussion on this blog](#), Wayne Wu usefully distinguishes the standardly discussed transparency claim (which asserts a commonality of the targets of perception and introspection) from a transparency claim that asserts a commonality in the processing mechanisms engaged in perceiving and introspecting. I think this latter view, that there's a common mechanism, is one that Churchland can be read as endorsing, while nonetheless denying the common-targets transparency thesis.

### 2.4. Directness?

Churchland sometimes describes the introspection that he's interested in as "direct." See, for instance, his 1985 J. Phil paper, "Reduction, qualia, and the direct introspection of brain states." I must confess that it isn't fully clear what he has in mind by "direct", but I'll end with a few speculations. One use of "direct" that may perhaps be pertinent here comes from the philosophical literature on perception that contrasts, for example, indirect realist views with direct realist views such as intentionalism or disjunctivism. As I understand it, the crucial contrast upon which "direct vs. indirect" hinges has to do with a contrast between whether, in being perceptually aware of a tree, one does so only in virtue of first being aware of some mental intermediary, a sense datum or mental representation. The indirect theory is not so-called merely for positing a representational intermediary between perceiver and perceived, but instead for asserting that one must be *aware* of the intermediary and that any awareness of external objects depends on this prior awareness. Intentionalism holds the perceptual state to be representational, but counts as a direct theory for its denial that one must be aware of the representational state in order to be aware, for instance, of the tree that it represents. It's instructive to regard Churchland's view of perception I sketched earlier as direct in this sense. The conceptual capacities brought to bear in perceiving are applied automatically, and this can be read as being applied without deliberation or awareness. That is, one need not be aware of the concepts--a kind of representation--deployed in perceiving. Of course, it's part of Churchland's account of introspection that what one need not be aware of is nonetheless something that one can be aware of. And the model of introspection, following a close analogy to Churchland's account of perception, is analogously direct.