From “Neural correlates of consciousness” to “Neural causes of consciousness”
A commentary on “Consciousness, biology and quantum hypotheses”, by Bernard J. Baars and David E. Edelman

Jan R. Wessel

1. Introduction

As any other psychological phenomenon, consciousness must have tangible, measurable, and quantifiable underpinnings in the brain of any species capable of experiencing it.

As Baars and Edelman point out in their review [1], the empirical body of neuronal correlates of consciousness is growing rapidly, not least due to the work of the authors themselves. Here, the authors address a theory that has been proposed decades ago, but never gained much traction in the mainstream of cognitive neuroscience: The Orch-OR theory of consciousness by Penrose and Hameroff. Notably, theories that neglect most mainstream evidence before ultimately being proven right or wrong usually do not endure multiple decades in cognitive neuroscience.

What is different here? One big reason, in my opinion, is the biological nature of the commonly proposed neuronal underpinnings of transitive consciousness, which makes it unusually hard to take the ultimate step for a neuroscientific theory: to go from a number of neuronal correlates to a definitive set of biological causes.

2. Establishing causation in cognitive neuroscience

For a biological mechanism to be neural correlate of a psychological event, the mechanism must covary in magnitude or extend with the magnitude of the event. In order to establish a neuronal cause (in the narrow sense; the term “neuronal correlate” is sometimes used to imply causation), it must be proven that without the biological process in question, the psychological process cannot exist.

There are three main ways to establish causation in (human) cognitive neuroscience: Lesion quasi-experimentation in patients with structural brain damage, virtual lesion experimentation using transcranial magnetic stimulation (TMS), and artificial induction of neuronal activity (using e.g. TMS or deep brain stimulation). Despite known problems associated with every single one of these methods, they’ve proven very effective in establishing neuronal causation for several psychological phenomena.

DOI of original article: http://dx.doi.org/10.1016/j.plrev.2012.07.001.
E-mail address: jwessel@ucsd.edu.
In the field of consciousness, however, I think the nature of the brain processes proposed as causes make it particularly difficult to use these methods.

3. The biological nature of the neuronal correlates of consciousness

As the authors describe in detail, both theoretical and empirical work in the cognitive neurosciences demonstrates wide-spread cortico-(thalamo-)cortical (CT) phase-locking of oscillatory processes as a main neuronal correlate of the emergence of transitive consciousness [2]. This has not only been shown for perception of stimuli, as discussed by the authors. Even more complex “higher-order” contents, like hidden environmental regularities that can become consciously available over time, are accompanied by high-frequency oscillatory coupling in the EEG when they do [3,4]. Similar mechanisms have also been hypothesized to underlie the conscious perception of action errors [5].

The reason why it is problematic to ultimately make the point that phase-locking of oscillations is necessary – and, in fact: causal – for consciousness is that it is next to impossible to investigate with one of the three methods outlined above: The critical regions that need to be “in-sync” for a stimulus representation to be broadcast on a “global workspace” [6] might be variable depending on stimulus content, and might in fact not always be focal at all. The authors cite several studies in which focal damage to brain areas abolish conscious stimulus representations, but one could ask the question: Is it the oscillatory activity that is decisive for consciousness in these cases, or is it the specific representation within the damaged brain areas? If the neuronal cause of the emergence of consciousness is indeed a ubiquitous mechanism (like wide-spread CT-phase-locking) more so than a specific brain network, lesion experimentation might not necessarily provide definitive answers.

With respect to the induction of neuronal activity as a method of establishing causation, similar arguments apply: While it is certainly possible to induce simple neuronal activity within distinct brain networks in order to prove causation, with the current set of methods it is hard to artificially generate and induce the sort of phase-locked oscillations that potentially underlie the emergence of consciousness (of note, the same arguments hold true for Penrose and Hameroff’s theory).

4. Conclusion

Many neuronal correlates of consciousness are evident, as the authors describe. They have been well established in dozens of studies, are present across modalities and content, and are explained within comprehensive theoretical frameworks. Thus, they are very likely candidates to be causes of consciousness. However, until there is “ultimate proof” for a unified, general mechanism underlying consciousness, without which consciousness cannot emerge, and which inevitably leads to consciousness when present, alternative theories might stay harder to rebut than in many other fields of cognitive neuroscience.

References