Consciousness and the brain: deciphering how the brain codes our thoughts

Article · December 2014

DOI: 10.1080/15294145.2014.956209

1 author:

Terence Rogers
Icahn School of Medicine at Mount Sinai
6 PUBLICATIONS **132** CITATIONS

SEE PROFILE

Neuropsychoanalysis: An Interdisciplinary Journal for Psychoanalysis and the Neurosciences
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/rnpa20

Consciousness and the brain: deciphering how the brain codes our thoughts

Terence Rogers Ph.D.

Accepted author version posted online: 22 Aug 2014.Published online: 01 Oct 2014.

To cite this article: Terence Rogers Ph.D. (2014): Consciousness and the brain: deciphering how the brain codes our thoughts, Neuropsychoanalysis: An Interdisciplinary Journal for Psychoanalysis and the Neurosciences, DOI: 10.1080/15294145.2014.956209

To link to this article: http://dx.doi.org/10.1080/15294145.2014.956209

Neuropsychoanalysis, 2014
http://dx.doi.org/10.1080/15294145.2014.956209

BOOK REVIEW

Consciousness and the brain: deciphering how the brain codes our thoughts, by Stanislas Dehaene, New York, Viking (The Penguin Group), 2014, 266 pp., US$32.95 (hbk), US$19 (pbk), ISBN 9780670025435

If you read only one book on consciousness this year, make it this one.

It is entertaining, thoughtful, educational, and beautifully written, and provides an excellent introduction to the latest mainstream thinking in the neuroscience of consciousness. It is suitable for the general reader and the expert, and I especially recommend it to those who wonder what science has to say about consciousness.

The book has three major parts, sandwiched between an introduction and some final speculations. The first part is the single best exposition of the scientific study of consciousness that I have seen or heard. It sets a standard for clarity and rigorous thinking that will be hard to beat. The second part, comprising Chapters 4 and 5, captures the two main results of Stanislas Dehaene's main body of research: his proposal for "signatures" of a conscious thought and his theory of consciousness, the Global Workspace Theory. The third part is Chapter 6, The Ultimate Test, describing his initial attempts to use the signatures and the theory in clinical practice. Each part individually makes the book worth studying, and combined they are a tour de force, a compelling account of a subject long deemed beyond the reach of science.

Studying consciousness in the lab

In Chapter 1, he introduces the three basic ingredients of the experimental study of consciousness: focusing on conscious awareness, manipulating that awareness, and carefully recording introspection.

After clearly delineating the difference between vigilance (wakefulness), attention, and conscious awareness, Dehaene simplifies the problem of studying consciousness experimentally by deferring issues such as self-awareness and metacognition. In the wrong hands such simplifications often lead to data, but no insight. However, Dehaene's track record shows his outstanding scientific judgment, and the remainder of the book completely validates his choice of conscious awareness as the focus of study.

Having reviewed the challenges in exploring these mechanisms, he describes the lab setup that achieves the desired goal, including reproducibility, consistency, and controllability. The heart of the challenge is keep the experimental conditions the same and yet induce in a controlled manner either a conscious or an unconscious perception in the same viewer. He describes how a combination of masking and priming can achieve this goal, with a simple apparatus.

Before he can make use of this lab setup, he has to deal with one key point: the use of the viewer's report as the critical evidence of conscious perception. Noting that general use of subjective reports can be very misleading, he asserts that the subjective report of the viewer in the lab is exactly what we mean by conscious awareness. We do not read the viewer's subjective report as accurately determining what was presented to the viewer, but it is the required determinant of what the viewer thought they saw. Just to be sure that we get this point, he offers the Dehaene and Naccache Manifesto: "Subjective reports are the key phenomena that a cognitive neuroscience of consciousness purports to study. As such, they constitute primary data that needs to be measured and recorded along with other psychophysiological observations" (p. 42).

He is now ready to pose specific questions, and to show how they can be answered in the lab.

Chapter 2, *Fathoming Unconscious Depths,* is the heart of the book, showing us not only the experimental results but also the drama of scientific discovery and debate: how an experiment can seem to support a hypothesis, only to be knocked down by a more careful experiment. No summary can do it justice: in his hands the story is a spell-binding detective story, full of characters, highs and lows, and eventually the right ideas emerging triumphant. Along the way he has critical observations on various historical contributions to consciousness, including those of Freud.

In his use of the experimental setup, we see the mark of an outstanding scientist – the ability to frame the problem in a way that contributes useful insights and yields questions that are answerable in the lab. He first studies unconscious awareness to understand what mental activity humans are capable of processing unconsciously. He is then in a good position to study the types of awareness that require consciousness for full processing.

To capture the results and the methodology I will present some of the questions and answers without the intermediate steps. We can break down the main question – what can the brain process unconsciously – into a series of sub-questions: can visual binding of a stimulus occur

unconsciously? Yes. Can visual and auditory signals be fused unconsciously? Yes. Can words be recognized unconsciously? Yes. Can a word's meaning be unconsciously activated? Yes. Does unconscious processing reach into the cortex? Yes (into the fusiform gyrus, the premotor cortex, and the "number sense" area among others).

The latter part of Chapter 2 expands the discussion to include attention, value, and creativity, specifically clarifying the relationship between attention and awareness. Again he asks the pertinent question: it seems clear that conscious awareness needs attention, but does attention need consciousness? The answer may surprise some readers: "experiment after experiment has revealed the operation of selective attention without consciousness" (p. 75). And he goes on to assert that subliminal priming is not a passive, bottom-up process, operating independent of attention and instructions; instead, the degree of attention determines the strength of unconscious processing, just as for conscious processing.

One good question leads to another: how does our attention decide whether a stimulus is relevant? The answer is the assignment of value, and he describes some interesting experiments to show that we can assign this value unconsciously. Given the importance of value in psychoanalytic thought (being a key aspect in drive, motivation, and emotion), these passages will be of particular interest to psychoanalytic clinicians.

Finally, he ventures into the issue of unconscious creativity, setting up the key question to be addressed in Chapter 3, *What Is Consciousness Good For?*

His first answer comes in the form of a metaphor of the mind: consciousness is like the staff of the top executive of an organization, sifting and selecting from the output of "an unconscious army of workers." Then without burdening the reader with the mathematics, he replaces this metaphor with the idea of Bayesian inference, which gives precision to the idea that animals are faced with an overwhelming influx of sensory stimuli, which they need to resolve rapidly to one most likely interpretation. In this sense, the "army of unconscious workers" are processing the details of these stimuli, making some low-level sense of them, while consciousness selects from these possible interpretations the one that makes the most sense. His account is an easily understandable version of a mainstream neuroscientific theory of sensory consciousness.

One specific point is worth noting. Unconscious processes can handle incredibly sophisticated visual and motor coordination requirements, provided they have occurred very frequently in the past. The role of conscious processing is to deal with novel situations, and for that the brain needs to form potentially novel combinations of brain areas and to hold information available to those areas while they select the optimum response.

This leads to his second proposition about the evolutionary benefit of consciousness, described in the section entitled *Lasting Thoughts.* Quoting Daniel Dennett (1996), he suggests that a key role of consciousness is to condense sensory messages into a synthetic code, devoid of gaps and ambiguities, and compact enough to be stored for as long as conscious attention requires it for decision-making.

Using this distinction between conscious and unconscious processing, the subsection entitled *The Human Turing Machine* offers interesting insights into multi-digit arithmetic, the role of consciousness in supporting algorithmic processing and strategic thinking, and the relationship of the brain to a computer.

Chapter 3 ends with a discussion connecting consciousness to language and social engagement, and he offers many engaging examples. He discusses the value of sharing information to make better decisions, and introduces the default mode network, using it to consider the nature of introspection and ending with the provocative phrase that the Self is "a statistical deduction from observation" (p. 113).

Had the book ended here it would have been a very satisfying introduction to the scientific study of consciousness, and well worth its price. However, the next two chapters offer even more: clinically useful tools and a unifying theory.

The two main results

In Chapter 4, *The Signatures of a Conscious Thought,* he describes his first main result, four "signatures" of consciousness:

(1) The expansion of activity throughout the brain;
(2) A late EEG slow wave called P300 (or P3);
(3) A late and sudden burst of high-frequency oscillations;
(4) A synchronization of information exchanges across distant brain regions.

Dehaene's findings indicate that these are objectively identifiable physical events that are present when the brain registers a stimulus consciously, and absent when it registers a stimulus unconsciously. As he later describes, these signatures can be used as clinical tools in the diagnosis and treatment of minimally conscious patients.

The basic idea follows naturally from the earlier chapters: repeat the experiments that present a controlled mixture of stimuli presented above and below the threshold of consciousness, and simultaneously record the brain activity using electroencephalography (EEG), magnetoencephalography (MEG), or functional magnetic resonance imaging (fMRI).

The first signature marks the fundamental distinction between conscious and unconscious processing. As the brain processes a visual stimulus, the initial 200 milliseconds of activity is identical in conscious and unconscious awareness, but between 200 and 300 milliseconds the two types of processing diverge. The activity caused by a stimulus perceived without awareness dies out, but the activity from a stimulus reaching consciousness is amplified and flows out of the visual and temporal cortex to the associative and prefrontal cortex and back again. This difference is directly measurable and constitutes his first signature of consciousness.

The second signature follows directly from it: as each part of the cortex activates, it causes a strong positive electrical voltage on the scalp nearby, and the local EEG sensor registers a significant peak in its signal. As the neural activity spreads from the back of the brain to the prefrontal cortex, this local peak voltage tracks across the scalp and looks like a wave flowing forward. This wave is easily recognized by EEG systems and is called the P300 wave, because it often starts about 300 milliseconds after the onset of the stimulus.

Measuring the third and fourth signatures of conscious thought requires tracking activity occurring deep in the brain, and this is only available for humans under special circumstances. Fortunately for science, there are legitimate reasons for implanting electrodes in live human patients, particularly as a preoperative measure for surgery on epileptic patients. With this "deep brain" technology it became possible to measure more precisely where and when specific patches of neurons became active as the viewer perceived images, and again the distinction between conscious and unconscious processing was very clear. In one experiment (see p. 135), when consciously registering a word nearly 70% of the implanted electrodes registered activity, while only 25% were activated during unconscious perception. Figure 20 on p. 136 gives a clear picture of this third signature – a long burst of high-frequency (gamma band) activity accompanying conscious perception.

To arrive at the fourth signature requires a fascinating and wide-ranging discussion of experiments and their interpretation, and a very educational tour of hallucinations, "temporary lobotomies," stabilizing moving images and much more. The result is a point that is crucial to his theory, that consciousness is not *any* ignition of global brain activity, but a specific "brain web" that carries the signature of that particular conscious perception. He cites the famous "Bill Clinton" cell, which in a binocular rivalry experiment activates only when the image of Bill Clinton wins the binocular competition and falls silent when it loses. The brain web is a manifestation of synchronized activity across specific regions of the brain.

There is much more in this chapter, but he makes his main point very clear – we can find objective measurable signals that are present only when a stimulus causes conscious processing. And with this finding come two other important points: we can now experimentally probe for conscious activity in the human (and animal) brain, and we have experimental evidence that conscious activity is always associated with the temporary formation of a specific network of neuronal ensembles.

By the end of Chapter 4, the reader will have a good idea of the elements of Dehaene's theory of the brain's mechanisms to support consciousness, but in Chapter 5, *Theorizing Consciousness,* he brings it all together, summarizing his "global neuronal workspace" theory, which proposes that consciousness is brain-wide information sharing. He compares it to other theories, filling in detailed anatomical data to support his ideas. The chapter provides an excellent introduction to different theories of consciousness (e.g. those of Damasio or Tononi), and to the role of various parts of the brain.

It is important here to grasp his definition of "explaining" consciousness. It goes beyond the idea of neuronal correlates and introduces a sense of causality and potential quantification. For example, when in the kinetic theory of gases we connect the increase in motion of the atoms with the increase in temperature, we go beyond simple correlation: we aim to calculate how the temperature changes with specific changes in motion, and to explain how the transition from liquid to gas occurs. Dehaene has no less an ambitious goal for consciousness and subjective experience, though he accepts how difficult and time consuming it will be.

One idea is particularly stimulating – constructing a conscious state is as much about what stays out of consciousness as what is brought into consciousness. He elaborates the mechanism underlying the P300 signature wave, pointing out that it is a positive voltage that results from the firing of *inhibitory* neurons that dominates the EEG voltage on the scalp, thus representing those neurons excluded from the global excitation. Other experiments are required to identify the negative voltages that are the result of the ensembles that encode the resulting conscious thought.

This ties nicely to another thought reversal that might not be familiar to all readers: "Too many neuroscientists still adhere to the obsolete idea of the reflex arc as a fundamental model for the human brain" (p. 188). Instead, he points out, that *autonomy* is the primary property of the nervous system. Namely, the "restless" brain is highly active throughout its waking cycle, to the extent that the activation by an external stimulus is barely detectable in an EEG, and probably consumes less than 5% of the brain's energy.

Consciousness in the clinic

Having worked through and hopefully enjoyed the previous five chapters, some readers may well be asking whether these beautiful experiments and elegant theories are useful, and Dehaene obligingly provides some initial answers in Chapter 6, *The Ultimate Test,* by which he means clinical applications.

His main interest is in minimally conscious patients and he offers a useful if discouraging review of the current status of treatment, and of some existing diagnostic tools, together with a diagram of the various types of coma. His frustration at the haphazard approaches and frequent misdiagnoses is evident and subsequently leads to him taking action, using his "P300 signature" as a tool for discovering whether a minimally conscious patient is capable of conscious thought.

They have been using this approach at the Salpetriere Hospital in Paris with much success, though with many false negatives, failing to recognize potential consciousness in many patients who later partially recover. Similarly, Massimini in Milan directly stimulated various parts of the brain using transcranial magnetic stimulation (TMS), and measured with EEG whether widespread, extended duration activation of the patient's brain occurred. It appears to produce very useful results, but it is not yet ready to become a standard clinical tool. However, this chapter usefully draws attention to a real problem and the opportunity for new approaches that hold great promise for relatively cheap and effective diagnosis of different states of consciousness in brain trauma and stroke patients.

The chapter ends with a similarly cautiously optimistic account of how direct stimulation of the brain may reverse some patients' loss of consciousness. It also brings to an end the three main parts of the book: the scientific study of consciousness; the two main results – his proposal for "signatures" of a conscious thought and the Global Workspace Theory; and the initial attempts to use the signatures and the theory in clinical practice.

The future of consciousness research

The final chapter, *The Future of Consciousness,* has a purpose beyond mere speculation: he wants to show how we might begin to address harder questions of consciousness, building on his pioneering work and keeping his high scientific standards.

What are some of these questions?

- Can we determine the precise moment when consciousness first emerges in babies?
- Can we figure out whether a monkey, or a dog, or a dolphin is conscious of its surroundings?
- Can we solve the riddles of self-consciousness?
- Is the human brain uniquely capable of self-consciousness?
- Does the human brain have distinct circuits not found in other animals, and is their dysfunction the cause of uniquely human diseases such as schizophrenia?
- Will we ever be able to duplicate enough of the human brain to achieve artificial consciousness?

The great thing about having a theory is the ability to reframe or make more precise previously hard questions, and this applies directly to the first two questions above. Together with his wife and co-worker, Ghislaine Dehaene-Lambertz, Dehaene applied his auditory P300 signature test to two-month-old infants and found huge brain activation that extended beyond the auditory cortex, and he carefully eliminates other explanations to come to the conclusion that the infant has a level of consciousness.

As for animals, Dehaene observes that "I would not be surprised if we discovered that all mammals, and probably many species of birds and fish, show evidence of a convergent evolution to the same sort of conscious workspace" (p. 246). He also suggests that humans are not alone in having some form of self-knowledge.

His final thought on our human uniqueness is more controversial, but not out of the mainstream: human consciousness is the outcome of two nested mainstream evolutions – the development of a "global workplace," which occurred in all primates, and a later boost from language that occurred only in humans. This reflects a form of convergence between the old Chomskian view of language as a special capability of the human brain and the evolutionary biologist view that humans are much more similar to other animals than this view might imply. The convergence reduces the major language difference to the human ability to have recursive thoughts, rather than a genetically encoded Universal Grammar capability.

He ends by stating that David Chalmers, a philosopher at the University of Arizona, has wrongly swapped the labels in defining the "easy" and "hard" problems of consciousness. In his words, "it is the 'easy' problem that is hard, while the hard problem just seems hard because it engages ill-defined intuitions" (p. 262).

What a fitting end to a brilliant, engaging book.

Reference

Dennett, D. (1996). *Kinds of minds*. New York, NY: Basic Books.

Terence Rogers, Ph.D.
Icahn School of Medicine at Mount Sinai,
New York, NY, USA
rogersterry7@gmail.com