

Can neuroscience reveal the true nature of consciousness?

Victor A.F. Lamme
Cognitive Neuroscience Group, Room A626, Department of Psychology
University of Amsterdam, Roeterstraat 15, 1018 WB Amsterdam, The Netherlands

tel 31 20 5256675, fax 31 20 6391656, e-mail V.A.F.Lamme@UvA.nl

Consciousness is considered one of the ‘final frontiers’ in modern science. The phenomenon seems to escape all attempts to scientific reduction, and some philosophers argue that we may never be able to reveal its true nature. During the last decades, the subject has been taken up by neuroscientists, trying to find the ‘neural correlates of consciousness’ (the NCC). It seems, however, that this is not solving the riddle in any real sense. What would we learn about consciousness if we knew what neurons or brain structures are involved? I think the answer lies in taking a different approach than finding the NCC. Our starting point should be neuroscience itself, not consciousness (which is rather ill-defined anyhow). I have devised a hypothesis about how phenomena like visually guided behavior, visual attention, visual memory *and* conscious visual experience might emerge from different neural mechanisms. This hypothesis can be tested (and further refined) with experiments using awake behaving monkeys and human subjects, doing electrophysiological recordings and electrical and pharmacological manipulations, as well as brain imaging (fMRI / EEG) experiments. If supported by these experiments, this hypothesis will reveal that consciousness is not what we think it is now. It is different from attention, working memory, reportability, or ‘thinking’. In some cases, we may even be visually conscious without knowing it ourselves. Thus, this new approach in studying consciousness, inspired on neuroscience rather than psychology or philosophy, may reveal the true nature of consciousness. At least we will learn new things about consciousness, not accessible via our introspective intuition of it, or via experimental observations which take this intuition as a starting point. The answer will be: Yes, neuroscience can reveal the true nature of consciousness!

The problem: finding the neural correlate of consciousness isn’t going to solve anything

Consciousness: pondered by philosophers for millenia, studied by psychologists for a century, and now all eyes are on the neuroscientists, with their brain imaging tools, intriguing patient studies, trained monkeys etc., to reveal the answer to that age-old question – what is consciousness, and how does it emerge from our brains? But will they come up with a final solution? Research aimed at finding the ‘Neural Correlate of Consciousness’ (the NCC)^{1,2} is booming. Its goal is to find the difference between neural activity that produces consciousness and neural activity that does not. But where will that take us?

Consciousness is a rather ill defined phenomenon, and in fact one of the motivations for finding the NCC is to give a better definition of consciousness. Instead of first trying to exactly define consciousness, and then figure out how this works in the brain, the idea is to have (neuro)-psychological and neuroscientific findings converge towards a deeper understanding of consciousness. To start with, consciousness is loosely defined, as ‘the awake state in which we have experiences about which we can report at free will or request’, and the neural correlate of such a state and those experiences is sought. Then, we can refine the initial

definition, accordingly refine its neural correlate¹, and thus spiral down towards a full understanding of consciousness. There are some major problems with this approach:

1. If consciousness will in part be defined by its neural correlate, there is a circularity in proving that particular neural structures or events are, or are not, part of the NCC. This renders finding the NCC problematic. The endpoint of our spiralling journey towards understanding consciousness may not be the unique solution to the problem, as it will depend on the starting point.
2. By relying on heterophenomenological observations, such as the subjects' report about conscious experiences, we conflate consciousness with reportability. But a report, whether verbal or via a button press, is a motor output. We are thus studying a filtered version of conscious experience. It is not even sure that our introspective notion of conscious experience can be fully trusted, as a variety of experiments show.
3. By relying on reportability we also conflate conscious experience with attention. From the 'outside' consciousness looks very related to attention. Both seem to have a selective nature, and aspects of executive control and free will are common to both³. If indeed consciousness is the same as attention, as some assert⁴, then the whole search for the NCC seems futile, as the neural correlate of attention is already studied for decades.
4. Research thus far is primarily aimed at *localizing* the NCC in a particular structure or cortical area. But knowing where consciousness resides in the brain is not going to reveal something essentially new about it. Neither does it seem very logical that neurons in a particular part of the brain produce conscious experience while others do not.

It may very well be, therefore, that we are heading in the wrong direction. Moreover, some philosophers argue that finding the NCC will never solve the 'hard problem'. Neural processes may explain functions, such as how a particular sensory input is mapped onto a motor output. Complex neural processes may even explain complex (i.e. 'cognitive') sensorimotor mappings, including delays (as in working memory, or in planning the future), even explain abstract notions like motivation, intelligence or love. But it will never explain why these functions are accompanied by conscious *experiences*, instead of simply 'carried out in the dark'?⁵ Brain processes will never explain the essence of consciousness, being the private experience of the outside world, of emotions, of actions, of free will, and of self. It will not solve the 'qualia problem' nor bridge the 'explanatory gap'. And therefore, it will not tell us anything new about the true nature of consciousness⁶.

Neuroscience puts these philosophical issues aside for the moment. Let us first try to figure out what the NCC is, and then worry whether this solves the 'hard problem' or gives a better understanding of 'qualia'. As Crick and Koch⁷ put it 'No longer need one spend time attempting to understand the far-fetched speculations of physicists, nor endure the tedium of philosophers perpetually disagreeing with each other. Consciousness is now largely a scientific problem'. Maybe a deeper insight into the neural basis of consciousness will automatically answer whether the hard problem is as hard as it seems¹, or whether qualia really 'exist' or are a figment of philosophical imagination. But is neuroscience right?

So what is the outlook? If we ever find the NCC, it might be not the 'real' one because we took a wrong starting point. And even if we do find it, what will this tell us about the essence of consciousness, which are conscious experiences? Neuroscience may just end up as the third discipline that is unable to provide new insights into the problem of consciousness.

The solution: Taking neuroscience as a starting point

The goal of this project is to prove the latter point wrong. I agree with philosophy that merely linking neural activity to reported consciousness is leading us in the wrong direction. Then,

consciousness is confounded with cognitive functions enabling the report, such as decision processing, attention, and executive control. The ‘hard problem’ remains untouched, and nothing new is learned about the nature of consciousness.

I therefore propose a radically different approach. In this *forward* approach, conscious reports, states or actions are not taken as the starting point of which the NCC is sought. Instead, neural functions form the foundation, and with those a new ‘psychology’ is defined. Here, I will limit this to visual processes, to prove that visual consciousness can be understood and distinguished from functions like visually guided behavior, attention, working memory, and control⁸. It will reveal that conscious visual experience is not always what we think it is, and may even defy our introspective notion of consciousness⁹.

A central tenet of my approach is the distinction between feedforward and recurrent cortical processing. My hypothesis is that a single feedforward pass through the cortical network may mediate automatic, reflex-like, yet intelligent and complex (cognitive) behavior. However, feedforward processing is unconscious. Consciousness arises only after recurrent cortico-cortical interactions, mediated by feedback and horizontal connections^{9,10}. These interactions invoke processes related to memory formation, which in turn are central to the emergence of conscious experience.

The hypothesis has various counterintuitive but testable predictions. For example, conscious experience is not related to attention or executive control, and not necessarily associated with high level cortical processes^{9,10}. Second, the hypothesis argues for the existence of an unreportable form of consciousness, i.e. the (sensory) experience itself, stripped from cognitive functions, not even ‘known’ to the self, yet distinct from unconscious processes. It thus corroborates a distinction between Access and Phenomenal consciousness^{9,11}. Third, the hypothesis guides the way towards understanding consciousness in relation to memory formation and synaptic plasticity.

This hypothesis and its predictions will be tested at the macroscopic level of cortico-cortical communication up to the microscopic level of neuronal interactions. I propose experiments in human and awake monkey subjects, using brain imaging tools like EEG, fMRI and single or multi-neuron recordings, and manipulations like TMS and pharmacological inactivation. Their goal will be to either falsify or verify the hypothesis, and to further refine it: what is the exact relation between recurrent processing, memory and conscious experience, both at the level of information processing and at the level of synaptic processes, as both may be of importance in bridging the explanatory gap.

When successful, this will show that neuroscience *can* give fundamental new insights into the true nature of consciousness. The gap between brain and mind will be closed, not by bridging it, but by moving our notion of mind towards that of brain. In my mind, that is the way to start solving this problem, which is central to a variety of disciplines, such as philosophy, psychology, and biology.

Theoretical background: Identifying the main features of cortical processing

I first propose *neural definitions* of consciousness and attention, and thus build a hypothesis about what consciousness is. From that hypothesis several new insights into the nature of consciousness emerge, that will be tested experimentally. The proposed neural definitions emerge from two important features of cortical sensory and sensori-motor processing, which are 1. depth of processing, and 2. the distinction between feedforward and recurrent processing.

Features of cortical processing 1: Depth of processing

The cortical sheet contains tens of cortical areas, each with their own anatomical and functional characteristics. Cortico-cortical fibers connect these areas, such that information can flow from sensory to other sensory areas, or to motor areas, and vice-versa. These connections define a hierarchy among these areas. Primary visual cortex (V1, also striate cortex or area 17), where visual information enters the cortex, is the lowest area in the visual sensori-motor hierarchy. From V1, information is distributed to the extra-striate areas (like V2, V3, V4, MT), and from there to areas in the parietal and temporal cortex, constituting the ‘dorsal’ (parietal), and ‘ventral’ (temporal) visual streams¹². The dorsal stream is mainly translating sensory input into motor behavior, while the ventral stream plays a central role in object recognition¹³.

Tuning properties in low level areas are ‘simple’ (such as orientation selectivity in V1), while higher areas express more complex tuning, related to object recognition (such as face-selectivity in temporal cortex) and sensori-motor transformations (such as tuning to position and motion in parietal cortex). This culminates in motor and other ‘executive’ areas, where the sensory input is finally translated into motor programs or plans^{14,15}.

Response properties of neurons along this hierarchy have mainly been studied using isolated stimuli. But natural scenes typically contain many objects. In that case, competition between these stimuli arises^{16,17}, such that not all stimuli reach into the highest levels of this hierarchy; only a few will reach the motor cortex. This constitutes a very important first feature of cortical processing, which I call *depth of processing*. Depth of processing varies from stimulus to stimulus, depending on a variety of things. Stimuli may have more impact because they are stronger along some feature dimension, such as brightness, contrast, or colour, and hence activate neurons more strongly at the outset. Foveal stimuli have an advantage over peripheral ones, simply because more neurons are ‘looking’ at them. But also stimuli that are ‘recognized’ by dedicated neurons (such as faces) have a much better chance than stimuli that are not. Depth of processing is thus also determined by how a particular stimulus or object matches the cortical object detection machinery¹⁸. To put it in artificial neural network lingo: depth of processing depends on how the features of an object match the stored synaptic weights of the cortical network. In the animal brain, these weights are shaped by genetic factors and by learning.

But depth of processing is also influenced by processes of shorter time scale. For example, if a stimulus is preceded by a stimulus that shares its features, the second stimulus can benefit from the remaining trace of activity. Also, such ‘paving of the way’ may be achieved ‘top-down’, via feedback connections, such that higher level areas support the processing of particular stimulus properties (such as location, or colour, or handyness, etc) in lower areas^{17,19}. In this way, depth of processing is determined by short-term (memory) effects of remaining or actively sustained neural activity.

We can thus, at least in principle, understand how, and why, a particular stimulus is processed more or less deeply. Of course many details of this selection process still need to be worked out, but it is an essential part of sensori-motor processing, that can readily be studied with a variety of techniques, such as neurophysiology in primates¹⁷, fMRI in human subjects^{19,20}, and psychophysics. Therefore, depth of processing is a useful concept in our understanding of cortical functioning.

Features of cortical processing 2: Feedforward versus recurrent processing

The fast feedforward sweep

Humans and other primates move their eyes about three times per second when exploring a visual scene. Each time the eyes come to a standstill, a new image is projected on the retina, that very rapidly informs the brain. Within 40 milliseconds, neurons in area V1 start to fire action potentials. Only 10 milliseconds later, cells in extrastriate areas (V2, V3) respond. This implies that their responses must be based on the very first spikes that are emitted from V1. In equally short time steps of about 10 ms per area, successively higher areas of the cortical hierarchy start to respond. Within 100 to 120 ms, the whole brain, including areas in motor and frontal cortex, is updated about the new image on the retina⁸.

We can thus identify a *fast feedforward sweep* (FFS) of information transfer, from primary visual cortex to the extrastriate and dorsal and ventral stream areas, all the way up to motor cortex, and other regions involved in controlling and executing movement. In some respect, we could call this a *cortical reflex arch*, mediated by the feedforward connections that go from low level visual areas up to motor cortex, and potentially transforming visual input into a motor response (in fact multiple feedforward sweeps may be identified, for example a fast one and a slower one, according to the magno- and parvo cellular inputs from the retina, or a dorsal and a ventral one, etc.²¹)

This cortical reflex arch has remarkable properties. Given its speed, it can only express the most elementary and directly hardwired computations, using a minimal amount of synapses going from one area to the next. Nevertheless, these early responses already fully express the receptive field tuning properties that characterize cortical processing. Orientation tuning in V1 cells is present from the very first spikes that are fired. The same goes for tuning to colour, direction of motion, stereo depth, etc, that characterizes the extrastriate areas. Even the highest levels of selectivity appear to be mediated by feedforward connections: cells in inferotemporal cortex distinguish between face and non-face stimuli with their first spikes^{8,22,23}.

The FFS thus enables a very rapid categorization of visual stimuli into all sorts of (probably) behaviorally relevant classes, such as face / non-face, animal / non-animal, etc²⁴. The FFS can not only detect elementary features like orientation, motion and colour, but also discriminate between complex feature constellations, such as objects and faces. In other words, a great deal of object recognition is established during the FFS, and potentially related motor responses are initiated.

Recurrent processing

As soon as the FFS has reached a particular area, horizontal connections start to connect distant cells within that area, and feedback connections start sending information from higher level areas back to lower levels, even all the way down to V1 and the thalamus²⁵. Together, these connections provide what is called *recurrent processing* (RP).

Much less is known about the roles of RP than is known about the FFS. RP may induce dynamic changes of tuning properties. For example, a V1 neuron tuned to vertical contrast may change that preference to, say, horizontal during the course of its activity²⁶. Something similar occurs in IT neurons: these neurons initially only discriminate between face and non-face stimuli, but at longer latencies also discriminate between the faces of different individuals²⁷. It is however not firmly established that these are expressions of RP.

We have studied the role of RP by recording from neurons in V1 of awake monkeys. An example is shown in figure 1. Monkeys were shown scenes with texture defined figures (fig. 1a). The figures were at different positions relative to the receptive field of the neurons that we recorded from. Obviously, the orientation of the texture patch that fell within the

receptive field influenced the early phase of the V1 responses (~50ms). However, if the texture patch in the receptive field was held constant, we observed that at longer latency (>100ms) the cells responded more strongly when that patch was part of a figure than when it was part of the background. Apparently, at the longer latency, some information from *beyond* the V1 receptive field, information about its visual context, influenced the response. Hence the name of the phenomenon, *contextual modulation*^{28,29,30}.

That contextual modulation (CM) is mediated by RP was confirmed by a study where V1 was isolated from feedback from higher areas (because these were lesioned), yet still received input from the thalamus (and the eyes). Orientation tuning (a FFS property) was not changed, but CM was abolished; the cells no longer discriminated between patches of textures that belonged to figure or background³¹. CM therefore can be viewed as an electrophysiological marker of RP, just like the early tuning properties of neurons are a marker of the processing established during the FFS.

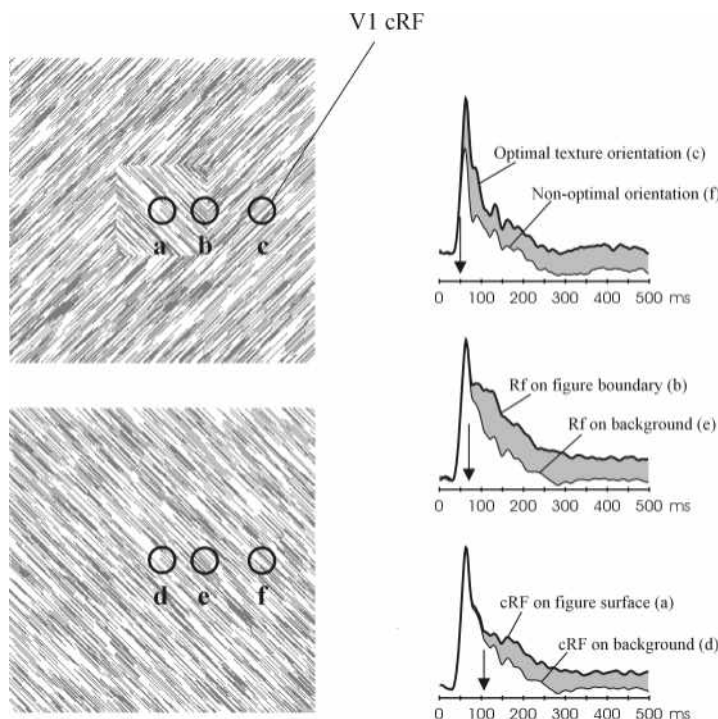


Figure 1: Contextual modulation

Neural responses (spike frequency on vertical axis, time after stimulus onset on horizontal), to stimuli as shown on the left, containing a textured figure (top), or only background (bottom). Responses are shown for the cases when the receptive field of the neuron is at the position indicated by the circle (V1cRF). Gray shading indicates the difference between two responses that are compared. C vs f indicates orientation selectivity, a vs d indicates contextual modulation. Note that the stimulus within the RF is identical for a and d.

We, and many others, have shown that CM occurs in many different conditions. CM typically expresses the *perceptual* interpretation of the scene. A stimulus that is perceived as more bright or salient due to its context, evokes a stronger V1 response. When it is grouped with other stimuli into a discernable object, its evoked response will likewise be stronger. In general, CM seems to be expressing fundamental visual processes like perceptual grouping and figure-ground segregation^{29,30}.

We are however only beginning to understand the full potential of RP. As soon as the FFS has finished, RP can start, but may then take a variable amount of time. As time goes by, the recurrent interactions may grow more and more widespread, integrating information from distant regions of the brain, and thus enabling evermore complex and recursive computations. Theoretically, this could be the basis of many of our higher-level cognitive capabilities⁸. We can unravel these capabilities by studying EEG and fMRI equivalents of CM in human subjects^{52,53}.

The Hypothesis: From two neural definitions to four types of consciousness

From the characteristics of cortical processing identified above, I draw *neural definitions* that form the basis of a hypothetical theoretical construct in which attention and consciousness are redefined⁹. Here it is: I simply equate *depth of processing* to attention, and the *FFS / RP dichotomy* to the distinction between unconscious and conscious processing. Let's start with the first one:

The depth of processing that a stimulus reaches is the neural definition of the amount of attention that is allocated to that stimulus:

Depth of processing = Attention

In many respects this fits quite well with the established notion of attention as derived from psychological and neural experiments. Indeed attended stimuli are processed more strongly and more efficiently and have a better chance of influencing behavior. The neural mechanisms that govern depth of processing (the genetic, long term, and short term 'memory' effects described above) are known in the psychological literature as 'saliency', 'capture', 'bottom-up attention', 'priming', 'cueing', 'top-down attention' and other attentional phenomena^{17,32}. But in some respects this definition of attention may not fit established knowledge. For example, the pre-attentive / attentive distinction that emerges from the visual search literature³³ does not fit into it. But let me stress once more that in my *forward* approach, neural mechanisms form the basis. Phenomena that are traditionally viewed as attention-related are here better viewed as belonging to another aspect of neural processing. That brings me to the second neural definition:

The distinction between feedforward and recurrent processing is the neural definition of the unconscious / conscious dichotomy:

The Feedforward Sweep = Unconscious

Recurrent processing = Conscious

What I mean with conscious in this regard is processing that is accompanied by some sort of *experience*. This implies that feedforward processing, no matter what area is activated by it (or, no matter how *deep*), is never accompanied by a conscious experience. Feedforward activation of V1 is not accompanied by conscious experience, but neither is feedforward activation of face selective cells in area IT or even cells in prefrontal cortex.

There is in fact already strong support for the claim that feedforward activation of the cortex is not sufficient for conscious experience. Stimuli that are made invisible by masking (presenting a second stimulus shortly afterwards)⁷⁰, nevertheless evoke brief, but widespread activation in areas such as V1³⁴, IT²³, frontal cortex³⁵ and motor cortex³⁶. Also, stimuli that are made invisible by presenting them with opposite colors to the two eyes (dichoptic masking, see also below), evoke equally widespread activation of cortical regions as visible versions of the stimulus³⁷. Feedforward activation of neurons can also be found in anesthetized, and therefore unconscious, animals. Tuning properties often hardly differ between awake and anesthetized conditions³⁸. Finally, feedforward activation of V1 neurons is equally strong for stimuli that are seen or not seen by animals engaged in a figure-ground detection task³⁹.

From those findings we could tentatively conclude that RP is at least *necessary* for conscious experience. This is supported by several findings: Masking suppresses contextual modulation⁴⁰, which we have identified as a marker of RP. Contextual modulation is also absent when stimuli are spontaneously not seen by animals engaged in a figure-ground detection task³⁹. Transcranial Magnetic Stimulation (TMS) to the occipital cortex at the time

that feedback signals arrive, suppresses conscious experience of stimuli presented 100 ms earlier^{41,42}.

It thus seems that the neural definitions proposed here already make some sense. Moreover, in this way, attention, motor outputs and reportability are clearly distinguished from the conscious / unconscious dichotomy, as we will see below.

From neural definitions to hypothetical constructs: four types of processing

Depth of processing (attention) and feedforward or recurrent processing are orthogonal properties of cortical processing. Feedforward processing may be deep or shallow, and likewise, recurrent processing may be shallow and limited to a small group of visual modules, or widespread and engaging a large part of the brain, including 'executive' areas. Combining the neural definitions we thus get (at least) four types of visual processing that may be discerned:

- 1. Shallow feedforward processing (unattended, unconscious).** In this case, a stimulus only activates a limited number of visual areas, and therefore cannot directly or indirectly influence or produce behavior, and cannot be reported. Given the absence of recurrent processing there is no conscious experience of the stimulus.
- 2. Deep feedforward processing (attended, unconscious).** Here, the stimulus activates not only low- and high-level visual areas, but also the associated motor regions and / or prefrontal cortex. It thus has the potential of directly producing (or inhibiting) a motor response, or otherwise influence behavior (f.i. priming)³⁶. Potential motor programs are activated by it, not necessarily executed. Because it is only feedforward, it is unconscious. Think of it as a cortical reflex, as when you grasp a bottle of wine that accidentally drops from the table, or when a tennis-pro has to return a serve that has been hit at 140 mph. These are complex actions, requiring complex (and probably cortical) computations. Yet, at the moment of initiation of those actions, there is no conscious experience yet. The conscious experience only sets in afterwards, when the third type of processing emerges:
- 3. Shallow recurrent processing (unattended, conscious).** Typically, at the moment the FFS has reached frontal and motor areas (~100ms), RP between visual areas starts (figure 2). At first this is shallow, and limited to several visual areas. Hypothetically, this type of processing can be evoked by many objects in the scene⁴³, because at these mostly retinotopically organized areas there is little competition between objects at different locations. Therefore, this kind of processing evokes a conscious experience of many of the parts and objects in the scene. Moreover, the recurrent interactions resolve ambiguities, group features, detect figure surfaces from background etc., resulting in the process of perceptual organization³⁰. However, because the recurrent interactions do not yet incorporate motor, prefrontal or language areas, this kind of conscious experience cannot influence behavior, and cannot be reported about, maybe not even to the self⁴⁴. I speculate that we cannot incorporate the results of shallow RP into many executive cognitive processes, such as 'mental rotation' or others forms of 'thinking' about visual input. Consequently, this is a very fragile representation, that is quickly overwritten with every new glance or stimulus presentation, resembling iconic memory⁴⁵ (when the stimulus is removed). Nevertheless (and by definition), it is a conscious representation.

4. Deep recurrent processing (attended, conscious). This occurs when a stimulus evokes widespread recurrent interactions, ranging all the way from visual cortex to prefrontal and motor structures. Now the stimulus is full-blown accessible for cognitive operations, and under attentive conscious control. It can be reported. However, given the widespread nature of the interactions involved, this can only occur for a limited set of stimuli. In my opinion, this is where the classical attentional bottleneck comes into play, where only one to four items can be attended, held in working memory, and reported about⁴⁶.

The Feedforward sweep & Attentional selection

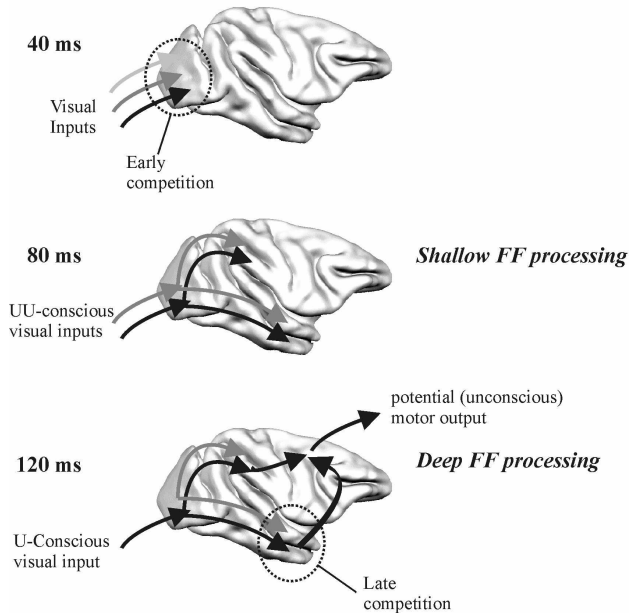


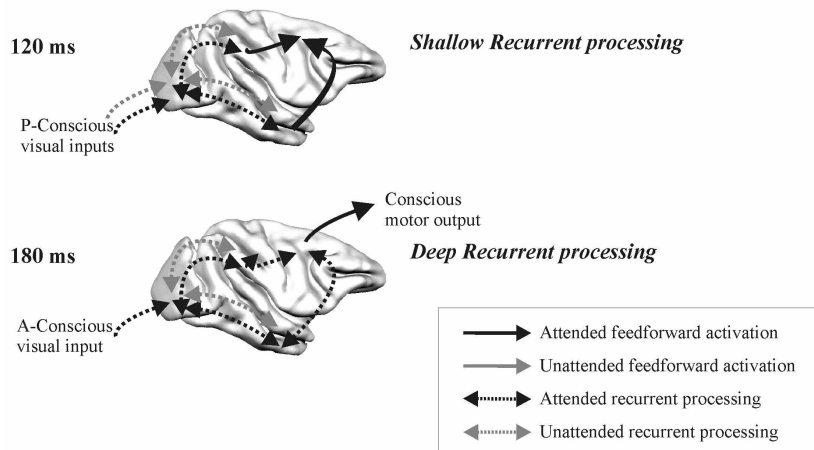
Fig 2: Characteristics of cortical processing

The top three panels show the progression of the feedforward sweep over time (40-120ms), and the selection of inputs that becomes more and more competitive from lower to higher levels. Feedforward processing is unconscious, yet may mediate, initiate or modify ongoing behavior.

The lower two panels show the progression of recurrent processing (RP), from shallow RP, resulting in the phenomenally conscious experience of visual inputs (P-consciousness), to deep RP, resulting in full-blown attentive consciousness (A-consciousness)

Four fundamentally different types of processing can be discerned:
 UU-conscious = Shallow FF processing
 U-conscious = Deep FF processing
 P-conscious = Shallow Recurrent Processing
 A-conscious = Deep Recurrent Processing

Recurrent processing & Attentional selection



This is only a brief description of the four types of cortical processing a visual stimulus might undergo. Somewhat more detail is given below, and can be read elsewhere^{8,9,10,47}. Strongly related is a distinction made before¹¹, between Phenomenal (P-) consciousness, and Access (A-) Consciousness. In P-consciousness (just like in stage 3), there is phenomenal experience

of stimuli, yet without cognitive access to that experience, while in A-consciousness (just like in stage 4), there is full access accompanying the conscious experience. In the following, I will refer to the four stages of processing as UU-conscious (unattended, unconscious), U-conscious (unconscious, yet attended), P-conscious (conscious, yet unattended), and A-conscious (conscious and attended).

To some extent, these associations seem logical. That shallow feedforward processing corresponds to unattended and unconscious is unsurprising. That stimuli evoking widespread recurrent interactions are consciously attended will neither raise many eyebrows. The critical categories are 2 and 3. That feedforward processing is always unconscious, even when it reaches high level areas, is debatable. Even more controversial is the third category. Does it make sense to call RP that is limited to the visual areas conscious? Even when you cannot report about the stimuli? Should we call that conscious? Similarly, the distinction between P- and A-consciousness is rather controversial and highly debated among philosophers and neuroscientist. However, these categories follow directly from the neural definitions that I have adopted. Therefore, they form the focus of the research testing this hypothesis.

Why is recurrent processing conscious?

I have already reviewed some evidence that RP is *necessary* for conscious experience, and that feedforward processing alone is unconscious. A central tenet of the hypothesis, however, is that RP is also *sufficient* for conscious experience. More strongly even, RP *is* conscious experience. This tenet is obviously in need of experimental support. Moreover, it begs the question how that relates to the more or less unitary nature of conscious experience. It could very well be, for example, that many islands of RP exist at the same time. Would we then have various, maybe even competing, conscious experiences at the same time? Related to that; what exactly is a sufficient amount of RP for conscious experience to arise? The recurrent interaction between two neurons? Between two layers? Two areas? The whole visual brain?

I hypothesize that once a core of RP exists, it rapidly grows as widespread as possible. Competing cores are quickly absorbed or overwritten by the most potent one. The ‘winning’ core acts somewhat like an attractor in a chaotic system. Multiple islands of RP may exist very briefly, but very quickly, large groups of neurons coherently interact, to represent a unitary solution to the many possible interpretations of the visual scene. Multiple recurrent cores may coexist (f.i. representing multiple objects, thoughts etc), as long as they are not competing. This means, for example, that neurons that represent the contours of an ambiguous object (that can be seen as possible object A or B) can never engage in recurrent processing with neurons encoding A, *and* with neurons encoding B. Thus, RP contributes to establishing the unitary nature and final solution of perceptual organization.

Finally then, the question remains why these recurrent cores would yield conscious *experience*. Of course this is a question within the realm of the ‘hard problem’, touching the problem of linking the physical (the brain) with the mental (experience). I think that is still a step to far right now. But maybe I am only scratching the surface by linking conscious experience to RP. It is very likely that the prolonged recurrent activation of cells trigger the activation of NMDA-type receptors, following Hebbian rules⁴⁹, and thus mediate synaptic plasticity⁴⁸. It could be that the final neural definition of consciousness processing is processing that evokes cortical synaptic changes. If that idea can be proven, the mind-brain problem could be attacked from a more fundamental level, or put in a different perspective altogether. A comparison could be made to an older question: While the theory of evolution gave a new perspective on ‘the meaning of life’ the discovery of DNA has totally abolished the question of ‘what life is’, and eradicated once popular notions like *elan vital*.

References

1. Crick, F. and Koch, C. (1998) Consciousness and neuroscience. *Cerebral Cortex* 8, 97-107
2. Rees, G., Kreiman, G. & Koch, C. (2002) Neural correlates of consciousness in humans. *Nature Reviews in Neuroscience* 3: 261-70.
3. Posner, M.I. (1994) Attention: the mechanisms of consciousness. *Proc Natl Acad Sci U S A.* 91, 7398-7403.
4. O'Regan, J. K. and Noe, A. (2001) A sensorimotor account of vision and visual consciousness. *Behav Brain Sci.* 24, 939-1031
5. Chalmers D.J. (1995) The puzzle of conscious experience. *Scientific American*, 273, 80-86
6. Bennet, M.R. & Hacker, P.M.S. (2003), Philosophical foundations of neuroscience. Oxford: Blackwell.
7. Crick F., Koch C. (1995) Are we aware of neural activity in primary visual cortex? *Nature*, 375, 121-123.
8. Lamme, V.A.F., and Roelfsema, P.R. (2000) The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences* 23, 571-579
9. Lamme, V.A.F. (2003) Why visual attention and awareness are different. *Trends in Cognitive Sciences*, 7, 12-18
10. Lamme, V.A.F. (2000) Neural mechanisms of visual awareness: A linking proposition. *Brain and Mind* 1, 385-406
11. Block, N. (1996) How can we find the neural correlate of consciousness. *Trends in Neurosci.* 19, 456-459
12. Felleman DJ, Van Essen DC. (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, 1, 1-47
13. Goodale MA, Milner AD. (1992) Separate visual pathways for perception and action. *Trends in Neurosci.* 15, 20-25
14. Maunsell JH. (1995) The brain's visual world: representation of visual targets in cerebral cortex. *Science*, 270, 764-769
15. Kurata K. (1994) Information processing for motor control in primate premotor cortex. *Behav Brain Res.* 61, 135-42
16. Desimone, R. (1998) Visual attention mediated by biased competition in extrastriate visual cortex. *Philosophical Transactions of the Royal Society of London, series B*, 353, 1245-1255
17. Desimone, R. and Duncan, J. (1995) Neural mechanisms of selective visual attention. *Ann. Rev. Neurosci.* 18, 193-222
18. Desimone, R. (1996) Neural mechanisms for visual memory and their role in attention. *Proceedings of the National Academy of Sciences USA*, 93, 13494-13499.
19. Driver, J. and Frackowiak, R.S. (2001) Neurobiological measures of human selective attention. *Neuropsychologia* 39, 1257-1262
20. Rowe J, Friston K, Frackowiak R, Passingham R. (2002) Attention to action: specific modulation of corticocortical interactions in humans. *Neuroimage*.17, 988-998.
21. Bullier, J. (2001) Integrated model of visual processing. *Brain Research Reviews*, 36, 96-107
22. Oram, M.W., and Perrett, D.I. (1992) The time course of neural responses discriminating between different views of the head and face. *J. Neurophysiol* 68, 70-84
23. Rolls ET & Tovee MJ (1994) Processing speed in the cerebral cortex and the neurophysiology of visual masking. *Proceedings of the Royal Society of London, series B, Biological Science*, 353, 1935-1942
24. VanRullen, R. and Thorpe, S.J. (2001) Is it a bird? Is it a plane? Ultra-rapid visual categorisation of natural and artificial objects. *Perception* 30, 655-68.
25. Salin PA, Bullier J. (1995) Corticocortical connections in the visual system: structure and function. *Physiological Reviews*, 75, 107-154
26. Shapley R, Hawken M, Ringach DL. (2003) Dynamics of orientation selectivity in the primary visual cortex and the importance of cortical inhibition. *Neuron*, 38, 689-699
27. Sugase Y, Yamane S, Ueno S, Kawano K. (1999) Global and fine information coded by single neurons in the temporal visual cortex. *Nature*, 400, 869-873
28. Lamme VAF (1995) The neurophysiology of figure-ground segregation in primary visual cortex. *Journal of Neuroscience*, 15, 1605-1615
29. Zipser K, Lamme VA, Schiller PH. (1996) Contextual modulation in primary visual cortex. *Journal of Neuroscience*, 16, 7376-7389
30. Lamme, V.A.F. and Spekreijse, H. (2000) Modulations of primary visual cortex activity representing attentive and conscious scene perception. *Frontiers in Bioscience*, 5, D232-243
31. Lamme, V.A.F. et al. (1998a) Feedforward, horizontal, and feedback processing in the visual cortex. *Curr. Opin. Neurobiol.* 8, 529-535
32. Egeth, H.E. and Yantis, S. (1997) Visual attention: control, representation, and time course. *Annu. Rev. Psychol.* 48, 269-297
33. Treisman, A. (1996) The binding problem. *Curr Opin Neurobiol.* 6, 171-178
34. Macknik SL, Livingstone MS. (1998) Neuronal correlates of visibility and invisibility in the primate visual system. *Nature Neuroscience*, 1, 144-149.
35. Thompson KG & Schall JD (1999) The detection of visual signals by macaque frontal eye field during masking. *Nature Neuroscience*, 2, 283-288
36. Dehaene, S. et al. (1998) Imaging unconscious semantic priming. *Nature* 395, 597-600
37. Moutoussis, K and Zeki, S. (2002) The relationship between cortical activation and perception investigated with invisible stimuli. *Proc Natl Acad Sci USA.* 99, 9527-32.
38. Lamme, V.A.F. et al. (1998b) Figure-ground activity in primary visual cortex is suppressed by anaesthesia. *Proc. Natl. Acad. Sci. USA.* 95, 3263-3268
39. Super, H. et al.. (2001) Two distinct modes of sensory processing observed in monkey primary visual cortex (V1). *Nat Neurosci.* 4, 304-310
40. Lamme, V.A.F. et al., (2002) Masking interrupts figure-ground signals in V1. *J. Cognitive Neurosci.* 14, 1-10

41. Corthout, E., Uttl, B., Juan, C.H., Hallett, M., Cowey, A. (2000) Suppression of vision by transcranial magnetic stimulation: a third mechanism. *Neuroreport*, 11, 2345-2349
42. Pascual-Leone, A. and Walsh, V. (2001) Fast backprojections from the motion to the primary visual area necessary for visual awareness. *Science* 292, 510-512.
43. Landman R, Spekreijse H, Lamme VA. (2003) Set size effects in the macaque striate cortex. *Journal of Cognitive Neuroscience*, 15, 873-882.
44. Gallagher, S. (2000). Philosophical conceptions of the self: implications for cognitive science. *Trends in Cognitive science*, 4, 14-21.
45. Coltheart, M. (1980) Iconic memory and visible persistence. *Perception & Psychophysics*, 27, 183-228.
46. Cowan, N. (2001) The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav Brain Sci.* 24, 87-185
47. Lamme, V.A.F. (2004) Separate neural definitions of visual consciousness and visual attention; a case for phenomenal awareness. *Neural Networks*, 17, 861-872
48. Dudai, Y. (2002) Molecular bases of long-term memories: a question of persistence. *Curr Opin Neurobiol.* 12, 211-216
49. Renart A, Parga N, Rolls ET. (1999) Associative memory properties of multiple cortical modules. *Network*, 10, 237-255
50. Ridderinkhof KR, Band GPH, & Logan GD (1999). A study of adaptive behavior: Effects of age and irrelevant information on the ability to inhibit one's actions. *Acta Psychologica*, 101, 315-337
51. Botvinick MM et al. (1999) Conflict monitoring versus selection-for-action in anterior cingulate cortex.. *Nature*, 179-181
52. Scholte HS (2003) Scene segmentation. PhD thesis, University of Amsterdam
53. Lamme VA, Van Dijk BW, Spekreijse H (1992) Texture segregation is processed by primary visual cortex in man and monkey. Evidence from VEP experiments. *Vision Research*, 32, 797-807.
54. Leopold, D.A. and Logothetis, N.K. (1999) Multistable phenomena: changing views in perception. *Trends Cogn Sci* 3, 254-264
55. Rensink, R.A. (2002) Change detection, *Annu Rev Psychol* 53, 245-277
56. Landman, R. et al. (2003) Large capacity storage of integrated objects before change blindness. *Vision Research*, 43, 149-164
57. Mack, A. and Rock, I. (1998) *Inattention blindness*, MIT Press, Cambridge, Massachusetts
58. Wolfe, J. M. (1999) Inattentional amnesia. In V. Coltheart, *Fleeting Memories*. Cambridge: MIT Press.
59. Breitmeyer BG, Ro T, Singhal NS (2004) Unconscious color priming occurs at stimulus- not percept-dependent levels of processing. *Psychological Science*, 15, 198-202
60. Nakayama, K., He, Z.J. and Shimojo, S. (1995) Visual surface representation: a critical link between lower-level and higher level vision. In Kosslyn, S.M. and Osherson, D.N. *Vision*. In Invitation to Cognitive Science. M.I.T. Press, p. 1-70
61. Schacter, D.L. et al (1993) Implicit memory: a selective review. *Annu. Rev. Neurosci.* 16, 159-182
62. Henson RN, Shallice T, Gorno-Tempini ML, Dolan RJ (2002) Face repetition effects in implicit and explicit memory tests as measured by fMRI. *Cerebral Cortex*, 12, 178-86.
63. Murray MM et al. (2004) Rapid discrimination of visual and multisensory memories revealed by electrical Neuroimaging. *Neuroimage*. 21, 125-135.
64. Stoerig P, Cowey A. (1997) Blindsight in man and monkey. *Brain*. 120, 535-559
65. Dudkin, K.N., Kruchinin, V.K., Chueva, I.V. (2001) Neurophysiological correlates of delayed differentiation tasks in monkeys: the effects of the site of intracortical blockade of NMDA receptors. *Neuroscience and Behavioral Physiology*, 31, 207-218
66. Chalmers, DJ (1995) Facing up to the problem of consciousness. *Journal of Consciousness Studies*, 2, 200-219
67. Penrose R (2001) Consciousness, the brain, and spacetime geometry: an addendum. Some new developments on the Orch OR model for consciousness. *Annals of the New York Academy of Science*, 929, 105-110
68. Woolf NJ, Hameroff SR. (2001) A quantum approach to visual consciousness. *Trends in Cognitive Sciences*, 5, 472-478
69. Denno DW. (2003) A mind to blame: new views on involuntary acts. *Behav Sci Law*. 21, 601-618.
70. Enns, J.T. and Di Lollo, V. (2000) What's new in visual masking? *Trends Cogn Sci*. 4, 345-352
71. Kolb FC, Braun J. (1995) Blindsight in normal observers. *Nature*. 377, 336-338.
72. Friston KJ, Harrison L, Penny W. (2003) Dynamic causal modelling *Neuroimage*, 19, 1273-1302.