The neurochemical correlate of consciousness: exploring neurotransmitter systems underlying conscious vision

van Loon, A.M.

Citation for published version (APA):
Chapter 1. Introduction

What makes us become aware of the things we see? How does our brain transform the information that we get into our eyes to a full experience of a nice piece of art or an intriguing person? When we see a person, we do not only experience the color of their sweater, someone’s face or the direction of the person’s gaze, instead we experience the person as a whole. How and where in the brain does this information become integrated into a unifying percept and into a conscious experience?

In search for the neural correlate of consciousness (NCC) different neural correlates have been proposed, depending on the kind of neural signals recorded, the type of manipulation used, and the interpretation of behavioral results (for reviews see Crick & Koch, 2003; Dehaene, Changeux, Naccache, Sackur, & Sergent, 2006; Lamme, 2006; Seth, 2007; Tononi & Koch, 2008). But the role of neurotransmitters in the NCC has not received much attention, while neurotransmitters are essential in neural communication. Intervening with neurotransmitters systems to induce loss of consciousness occurs every day, for example with anesthetics in surgery rooms. Neurotransmitter systems could therefore be very informative to study the NCC. But which neurotransmitter systems are involved?

We know that anesthetics operate on many different neurotransmitters systems: N-methyl-D-aspartate (NMDA), gamma-aminobutyric acid (GABA), glycine and acetylcholine (for a review see Alkire, Hudetz, & Tononi, 2008). In general, anesthetics specifically manipulate neurotransmitters by antagonizing or agonizing their receptors, causing a depression of neural activity by reducing excitatory glutamatergic neurotransmission (agonizing) and potentiating inhibitory GABAergic neurotransmission (agonizing) (Alkire, et al., 2008; Hemmings, et al., 2005). How does manipulation of these neurotransmitter systems affect consciousness? It is suggested that disrupting the balance between excitation and inhibition selectively affects information integration by means of reducing recurrent processing (Ferrarelli, et al., 2010; Roelfsema, Lamme, Spekreijse, & Bosch, 2002; Wyatte, Curran, & O'Reilly, 2012). Interestingly, recurrent processing has also been proposed to be the NCC on very different grounds (Ahissar & Hochstein, 2004; Dehaene, et al., 2006; Dehaene & Naccache, 2001; Lamme, 2006; Tononi & Koch, 2008). This will be summarized below, after a brief introduction into how visual information is processed in the brain.

Visual perception and feedforward processing

Our brain has to perform several complex steps to integrate the incoming visual information. First, after light hits our retina, information is sent via the optic nerve fibers through the lateral geniculate nucleus (LGN) of the thalamus to the primary visual cortex (V1) in the occipital lobe. From V1, information is sent forward to the other visual areas and finally distributed throughout the brain. It is thought that each area performs a cortical algorithm specific to that area, building on the computations performed by the preceding areas. Therefore, the visual areas form a hierarchy from low-level to higher-level areas. Neurons in early visual areas such as V1 have small receptive fields that respond to small image elements, such as oriented contrasts (Hubel & Wiesel, 1962, 1968). As information moves up the
hierarchy, the size of the receptive fields increases and eventually neurons respond to complex shapes and even to specific object categories such as faces in the inferior temporal lobe (Gross, Rocha-Miranda, & Bender, 1972). This initial sweep of activation is often referred to as the fast feedforward sweep (FFS) (see Figure 1.1A). During the FFS many brain areas are activated by the visual input. This could suggest that one becomes aware of an object in a scene when the information has reached these higher-level object-related areas. For example, activity in these areas is associated with object recognition performance (Grill-Spector, Kourtzi, & Kanwisher, 2001) and lesions in this region impair recognition of objects (James, Culham, Humphrey, Milner, & Goodale, 2003). However, processing by the feedforward sweep is not accompanied by conscious experience of that input (see Figure 1.1.A).

**Figure 1.1. Schematic representation of the feedforward sweep and recurrent processing**

(A) The feedforward sweep, where each area extracts visual information about e.g. orientation, shape, motion, or objects but the activation does not result in a conscious experience of the visual input. (B) Recurrent processing, where information is exchanged between higher and lower areas and within areas, by means of horizontal and feedback connections. Once RP extends towards areas in executive space such as in the frontoparietal network (Fr-Par) access to and report about the conscious experience is possible.

For example, higher level areas also get activated with invisible stimuli (Fahrenfort, et al., 2012) and even in anesthetized monkeys (Ku, Tolas, Logothetis, & Goense, 2011). Therefore, the mere activation of object selective neurons does not seem sufficient to generate visual awareness of the detected objects. Furthermore, it was shown that the early visual cortex is not only active in the first stage of vision: object recognition alters object representations also in early visual cortex (Hsieh, Vul, & Kanwisher, 2010) and even the contents of visual working memory and mental imagery can be successfully decoded from the activity patterns in early visual areas (Albers, Kok, Toni, Dijkerman, & de Lange, 2013; Harrison & Tong, 2009). These studies thus suggest that activity in early visual areas not simply reflects the sensory input but is also involved in higher-level vision and receives input from higher-level regions. Therefore, visual processing is not strictly hierarchical but follows an inverse hierarchical path (Ahissar & Hochstein, 2004; Hochstein & Ahissar, 2002; Lamme, 2006). The anatomical substrate of this is that the majority of the feedforward connections enabling the FFS are reciprocal: as soon as the FFS has reached a certain
area, feedback connections between higher and lower level areas and horizontal connections within that area get activated. This process is referred to as recurrent processing (RP) (Lamme & Roelfsema, 2000) (see Figure 1.1B).

Recurrent Processing & Anesthesia
Recurrent processing has also been directly linked to anesthesia. Recordings in macaque monkeys anesthetized with isoflurane (binding to GABA, NMDA and glycine receptors) demonstrated suppressed recurrent processing, whereas feedforward activity remained unaffected (Lamme, Zipser, & Spekreijse, 1998). Here, a figure-ground stimulus was used, where a figure square is defined by line segments with orientations that are orthogonal to those of the background (see Figure 1.2). Activity of V1 neurons was recorded with their receptive field either on the figure or on the background. The circle illustrates the receptive field of the neuron, which is too small to ‘know’ whether it is part of the figure or part of the background. Therefore, it needs input from surrounding neurons and neurons that are in areas higher up the visual hierarchy – which have larger receptive field sizes - to ‘tell’ the neuron whether the line segments it responds to belong to figure or background (Lamme, 1995 and Roelfsema et al., 2002 for a formal model). This is where RP comes into play which is reflected in an increased response for the figure compared to the background after ~100 ms, enabling figure-ground segregation (see Figure 1.2A). However, when the monkey is anesthetized, the figure-ground modulations disappear, whereas the initial FFS is still present (see Figure 1.2B).

Figure 1.2. Figure-ground modulation in awake and anesthetized macaque monkeys
(A) A square is visible because its line elements are orthogonal to that of the background. When the receptive field of a neuron located in primary visual cortex (V1) falls on the figure (depicted by the orange circle) instead of the background (depicted by the blue circle), the response becomes stronger after ~100 ms. Even though its receptive field is too small to encapsulate its context. Therefore it needs input from higher visual areas with larger receptive fields by means of recurrent processing. (B) Recordings during anesthesia. The FFS was still intact but RP was abolished.
Furthermore, contour integration responses in V1 disappeared under anesthesia using Pentobarbital (acting at the GABA<sub>A</sub> receptor) (Li, Piëch, & Gilbert, 2008). Moreover, when loss of consciousness was induced by Midazolam (GABA<sub>A</sub> receptor agonist) in humans, cortical effective connectivity was disrupted (Ferrarelli, et al., 2010). But not only under anesthesia RP seems reduced. Recently it was shown that integration of information is also disrupted during sleep (Tononi & Massimini, 2008) and in patients in vegetative state (Boly, et al., 2011). Therefore, the presence of RP seemed to be linked to someone’s state of consciousness.

**Recurrent Processing & Visual awareness**

On the other hand, theories on the *content* of consciousness have also stressed the importance of RP and the integration of information (Ahissar & Hochstein, 2004; Dehaene, et al., 2006; Dehaene & Naccache, 2001; Lamme, 2006; Tononi, 2005; Tononi & Koch, 2008; Tononi & Massimini, 2008). They propose that with the FFS visual features can be rapidly extracted, but that for a conscious visual experience feedback interactions that lead to a reverberating “ignition” are required. For example with backward masking a stimulus becomes less perceptible or even invisible through the presentation of a second stimulus - the “mask” - shortly after the first. Accumulating evidence suggests that when stimuli are effectively masked, RP in visual cortex is reduced, whereas feedforward activity remains relatively unaffected (Boehler, Schoenfeld, Heinze, & Hopf, 2008; Di Lollo, Enns, & Rensink, 2000; Fahrenfort, Scholte, & Lamme, 2007; Lamme, Zipser, & Spekreijse, 2002). Moreover, transcranial magnetic stimulation (TMS), when applied on early visual cortex during the time window of RP (around 100 ms after stimulus presentation), has also been shown to impair perception (Boyer, Harrison, & Ro, 2005; Juan & Walsh, 2003; Wokke, Sligte, Scholte, & Lamme, 2012), even though at that time the information has already been passed forward towards higher areas.

**Recurrent Processing & Neurotransmitters**

Modeling studies have demonstrated that RP consists of both excitatory glutamatergic feedback connections and inhibitory horizontal connections (Dehaene, Sergent, & Changeux, 2003; Lumer, Edelman, & Tononi, 1997). This is in line with the molecular profile of anesthetics, where also both excitatory glutamatergic neurotransmission and inhibitory GABAergic neurotransmission are modulated (Alkire, et al., 2008; Hemnings, et al., 2005). More specifically, the feedback connections in RP are thought to be mediated by the NMDA receptor (Dehaene, et al., 2003; Lumer, et al., 1997). This involvement of the NMDA receptor is further supported by the fact that it has some distinctive properties. The NMDA channel opening is voltage-dependent; it only opens when the magnesium blockade is removed by sufficient prior depolarization caused by AMPA receptors (thought to carry the feedforward signal) (Collingridge & Bliss, 1987; Salin & Bullier, 1995). This unique property might explain the modulatory function of NMDA receptors, as it ensures amplification of the firing rate of neurons that are driven by feedforward connections (Roelfsema & van Ooyen, 2005). Evidence comes from a monkey study that showed that blocking the NMDA receptor with 2-amino-5-phosphonovalerate (APV) and Ifenprodil in primary visual cortex specifically reduced feedback activity,
whereas blocking AMPA receptors specifically reduced the FFS (Self, Kooijmans, Supér, Lamme, & Roelfsema, 2012).

Furthermore, inhibitory horizontal connections have also been associated with consciousness perception for example in perceptual bistability (Blake, 1989; Blake & Logothetis, 2002; Hering, 1964; Lehky, 1988; Noest, Van Ee, Nijs, & Van Wezel, 2007). In bistable perception, perception fluctuates spontaneously between two distinct interpretations of a constant sensory input. A simple and classic example is shown in Figure 1.3. Here, we may see either a young woman looking to the rear, or an old woman looking forward, and perception spontaneously (or to some extent willfully) switches from the one interpretation (young woman) to the other (old woman). That the dynamics of such bistable illusions depend on inhibitory horizontal connections further strengthens the role of specific neurotransmitter systems in building the contents of consciousness.

Figure 1.3. Young woman-old woman bistable visual illusion

This image can be perceived as either a young woman looking to the rear or an old woman looking forward. Perception between these two interpretations can spontaneously or to some extent willfully switch.

Taken together, both excitatory glutamatergic neurotransmission and inhibitory GABAergic neurotransmission seem important neurotransmitter systems in consciousness. Moreover, both the state of consciousness as well as the content of consciousness seems to depend on recurrent processing. What is missing, however, are studies that manipulate both neurotransmitter systems (via e.g. pharmacological interventions) - and hence the state of consciousness - as well as the content of consciousness (via e.g. masking, bistable illusions etc.) while recording neural activity to assess RP. Such a combined effort may give more direct insights into the neural and pharmacological underpinnings of consciousness, and may open the way towards a molecular biological understanding of consciousness. Therefore, in this thesis, I have conducted several experiments combining measurements and manipulations of consciousness with manipulations of neurotransmitter systems.

Overview of this thesis

In the first experimental chapter (Chapter 2), I combined three different pharmacological manipulations: Dextromethorphan (a NMDA receptor antagonist), Lorazepam (a GABA<sub>A</sub> receptor agonist), Scopolamine (a muscarinic receptor antagonist), or a placebo with a visual backward masking paradigm. I recorded electroencephalographic (EEG) signals while participants discriminated between trials containing a figure or no figure, under various masking strengths (ranging from fully “seen” figures to fully “unseen” figures). This experimental set-up allowed us to explore the existence of a common neural pathway for two entirely different ways of manipulating visual consciousness: pharmacological intervention and backward masking.
As a follow-up on Chapter 2, I wanted to further investigate the role of GABA in visual awareness. Therefore, in Chapter 3, I linked visual GABA concentrations to the dynamics of bistable visual illusions. Models of bistable perception posit that mutual inhibition between stimulus-selective neural populations in visual cortex plays a key role in the dynamics of perceptual bistability (Blake, 1989; Blake & Logothetis, 2002; Hering, 1964; Lehky, 1988; Noest, et al., 2007). If this is true, one may expect a central involvement of the inhibitory neurotransmitter GABA in bistable perception. We used a computational model of these competitive neural interactions in visual cortex (Noest, et al., 2007) to derive specific predictions about the effect of GABAergic inhibition on the perceptual dynamics. As a proxy of cortical inhibition, we measured GABA concentrations in human participants by means of magnetic resonance spectroscopy (MRS). Previous research has demonstrated that there are consistent relationships between GABA concentration in visual cortical areas as measured with MRS and performance on visual tasks involving inhibition (Eden, Muthukumaraswamy, Freeman, & Singh, 2009; Yoon, et al., 2010). We measured from a region encompassing retinotopic visual cortex which has been shown to modulate during perceptual alternations (Donner, Sagli, Bonneh, & Heeger, 2008, 2013; Haynes & Rees, 2005; Lee, Blake, & Heeger, 2004; Tong & Engel, 2001). We also measured from a high-level area in the right dorsolateral prefrontal cortex (rDLPFC) which is thought to be involved in bistable perception (Knapen, Brascamp, Pearson, van Ee, & Blake, 2011; Sterzer & Kleinschmidt, 2007). To assess whether GABA is involved in different visual illusions, we used three bistable illusions namely binocular rivalry (BR), motion-induced blindness (MIB) and structure from motion (SFM). If we find a correlation between individual GABA concentrations and perceptual dynamics of the three visual illusions, GABA could be the common underlying mechanism.

To test the relationship between GABA and bistable perception more directly we have to manipulate GABA pharmacologically. In Chapter 4, I therefore manipulated the GABA\_A receptor with Lorazepam, while participants perceived the same bistable illusions as in Chapter 3. If we observe that Lorazepam administration has effects on the perceptual dynamics of these illusions that are congruent with the results of Chapter 3, we will have converging evidence for the reciprocal inhibition account of bistable perception.

Thus, the first experimental chapters focus on the role of GABA and inhibition in visual awareness, but as previously mentioned feedback connections which are important for visual awareness are thought to operate via the NMDA receptor (Dehaene, et al., 2003; Lumer, et al., 1997). A study using neurophysiological recordings in monkeys showed that blocking the NMDA receptors abolished neural signals related to figure-ground segregation and feature integration (Self, et al., 2012). However, it is unknown whether blocking the NMDA receptor also affects perceptual integration itself (Chapter 5). Here, we tested with a texture discrimination task whether ketamine (a NMDA receptor antagonist) reduces feature integration in humans. If ketamine disrupts feature integration, this would further support the evidence that recurrent processing is mediated via NMDA receptors.

Next, I wanted to assess the neural effect of partially blocking feedback activity with ketamine. Therefore, in Chapter 6, I conducted a pharmacological functional magnetic resonance imaging (fMRI) study to investigate how ketamine
influenced neural object representations in higher- and lower-level visual areas. A previous study demonstrated that, both in the lateral occipital complex (LOC) and in early visual cortex, recognition changed the neural representation of objects (Hsieh, et al., 2010). The authors suggested that the neural representations in early visual cortex were shaped by the top-down activation from LOC. In this chapter, we could now test this hypothesis directly. We presented ambiguous two-tone Mooney images of objects (Mooney, 1957; Moore & Cavanagh, 1998) and compared the neural representations before and after recognition of the Mooney image using multi-voxel pattern analysis (MVPA), and studied the effects of the ketamine on the recognition induced changes.

In Chapter 7, we wanted to assess the effect of Mooney recognition on neural object categorization responses. Previous research using fMRI demonstrated that category information; especially animate / inanimate categorization could be decoded using representational dissimilarity matrices (RDM) (Kriegeskorte et al. 2008). Is recognition of the image required for these category-selective responses? Or is it possible to observe category information in the brain based on the bottom-up feedforward input? To answer these questions, we studied the category-selective responses to animate and inanimate Mooney images in LOC before and after recognition. Our results seemed to suggest that animate / inanimate category selective responses in LOC can arise based on the bottom-up input or physical similarities of the images, independent of recognition. Whereas, perceptual or object specific responses in LOC might require recognition.

In the final Chapter 8, I will summarize and interpret the obtained results and outline possible future directions for the study of neurotransmission in conscious vision.