The Crack of Dawn

Perceptual Functions and Neural Mechanisms that Mark the Transition from Unconscious Processing to Conscious Vision

Victor Lamme

There is conscious vision, and there is unconscious visual processing. So far so good. But where lies the boundary between the two? What are the visual functions that shape the transition from "processing in the dark" to having a conscious visual percept? And what are the neural mechanisms that carry that transition? I review the findings on feature detection, object categorization, interference, inference, Gestalt grouping, and perceptual organization, and examine to what extent these functions correlate with the presence or absence of conscious vision. It turns out that a surprisingly large set of visual functions is executed unconsciously, indicating that unconscious vision is much "smarter" than we might intuitively think. Only when these unconscious mechanisms fail, and more elaborate and incremental processing steps are required, is consciousness necessary. The function of conscious vision may be to add a final layer to our interpretation of the world, to solve relatively "new" visual problems, and to enable visual learning.

Author

Victor Lamme
Victorlamme@gmail.com
Universiteit van Amsterdam
Amsterdam, Netherlands

Commentator

Lucia Melloni
lucia.melloni@brain.mpg.de
Max Planck Institute for Brain Research
Frankfurt a. M., Germany

Editors

Thomas Metzinger
metzinger@uni-mainz.de
Johannes Gutenberg-Universität
Mainz, Germany

Jennifer M. Windt
jennifer.windt@monash.edu
Monash University
Melbourne, Australia

1 Qualia 2.0

What do we need to know about consciousness? Which aspect of it is most mysterious? What do we want philosophy, psychology, neuroscience, computer science, or even physics to tell us about consciousness that we do not already know? The answer to that question may vary from person to person. To me it is this very simple thing: why do I see? Why do I have conscious experiences whenever I open my eyes? What makes the 1.5 kilograms of protein and fat in my head give me the wonderful sensations I experience every day, from the second I wake up until the moment I fall asleep?

The point is probably best illustrated by the difference between a digital photo camera and the human mind. A camera nowadays can do wonderful things. It can record an image at extreme resolutions, with the right focus and exposure, all by itself. It can identify a face, putting it in a frame on the screen, and writing the name below it of the person it recognizes. You can leave it to push the button at the moment everybody smiles. Connect it to a com-

puter, and it will detect emotions, recognize objects, or read handwriting on a letter. Surveillance cameras can detect suspicious movements or strange behaviours in crowds, outperforming human night-guards or intelligence agents. There is one big difference between the camera and the human mind, though. The camera does not see.[1] I do. And so does the night guard, most of the time. It is this aspect of visual processing that is in need of an explanation. Not the fact that I recognize the person in front of me, can read his emotions, talk to him, or pick up the cup of coffee he gives me. I can vaguely understand how my brain enables me to do that. What I do not understand is how it is that I see all those things.

Is that the "hard problem" all over? Am I talking Qualia? Not in the strict sense. In its original formulation, the hard problem would argue that there is no function, no neural process whatsoever that could ever explain conscious sensations (Chalmers 1995). Functions explain functions, but not the fact that I see. Qualia are defined as ineffable aspects of information: the redness of red, stripped of every possible functional property or reactive disposition. And with that comes the whole charade of inverted spectra, colour scientists called Mary, and explanatory gaps. Which didn't get us all that far—so let's not chase that unicorn again.

It's not that I don't want to address the hard problem, or bridge the explanatory gap. That is in fact exactly what I am after (Lamme 2010a, 2010b). But I would like to leave that for later. What we need to recognize, first, is that there must be some functions and some neural processes that are more closely connected to seeing than others (Crick & Koch 1998, 2003). For example, it is fairly reasonable to assume that an understanding of the neural basis of a reflexive motor response—like the pulling away of your hand when it touches fire— does very little towards explaining consciousness (Lamme & Roelfsema 2000; Lamme 2006). Other func-

tions may offer a better gateway. For example because they explain some fundamental aspect of seeing (Seth 2010), such as its unity, or because they coincide with the difference between conscious visual processing and visual processing that occurs "in the dark" (Lamme 2010a, 2010b). In trying to bridge the explanatory gap, I think it is important to first find the right tree up which to bark. We must first identify the exact boundaries between conscious and unconscious processing. The hard problem can then be attacked afterwards. Or maybe that whole explanatory gap will vanish right before our eyes once we are there.

This paper is about exactly that. Let's find the visual functions and neural processes that take us as close as possible to the hard problem, as close as possible towards explaining why we humans see, while photo cameras do not. And let's avoid barking up the wrong tree.

2 Why dolphins are not fish

To find the cognitive functions and neural processes that take us towards understanding the phenomenality of consciousness it is important to establish a boundary—a boundary with what we should call unconscious processing. This will by no means be an easy job (Lamme 2006). In fact, the whole issue of understanding consciousness and solving the explanatory gap is about positioning that boundary. There are situations where it is in fact unclear whether we should talk about a conscious sensation or not. Take the situation of a split-brain patient: here, a stimulus presented to the left visual field will be processed by the right half brain, typically devoid of communication via language. Hence, the subject will *tell* you that she did not see that stimulus. He may draw the stimulus, however, using his left hand. Or the left hand may point at the stimulus, or match it to a related subject (Gazzaniga 2005). Who are we to believe, then? The hand or the mouth? What types of behaviour may count as evidence for conscious sensations? Just speech? What about aphasic subjects, then? Similarly, there are conditions like neglect, or manipulations of attention (change blindness, inattentional blindness),

---

1 Or at least we assume it does not. This is the basic intuition we start from in trying to explain consciousness. If not, one easily slides into pan-psychism. That is a viable option of course: it could be that the camera does see, yet cannot "tell" us. However, the arguments put forward in the remainder of this paper seem to suggest that the camera does not see.

where it is difficult to be entirely sure that what appears to be not seen is in fact maybe just not attended to, and hence forgotten or not cognitively accessible and hence not reportable (Lamme 2003, 2006, 2010). This uncertainty has sparked a lively debate on the nature of consciousness, its potential independence of cognitive functions like attention, working memory, or access (Lamme 2004, 2010a, 2010b), and whether consciousness can ever be separated from a *report* about consciousness (Block 2005, 2007; Dehaene et al. 2006; Cohen & Dennett 2011; Fahrenfort & Lamme 2012). This debate is all about the difference between seeing and knowing, between phenomenality and access (P-consciousness and A-consciousness), between qualia and higher-order thoughts. In this debate, the issue that seems unsolvable is where exactly the boundary between conscious and unconscious processing should be laid.

In such attempts to establish boundaries, it is perhaps good to start from the extremes, as an example from zoology will illustrate. Superficially, one could argue about whether a dolphin is a fish or a mammal. Science has resolved that argument by looking at animals that we all agree are either mammals (such as dogs, cows, or monkeys) or fish (such as sole, tuna, or piranha). From that perspective, the key differences between these species lies in the way they breathe and reproduce.[2] Why are these the key differences? Well, differences in breathing do all the explaining for why fish are generally more adept at living in water instead of on land. Similarly, evolution towards the land has called for eggs with protective layers (amnios), as anamniotic eggs (that fish lay) cannot survive on land. The most extreme version of that is the intrauterine development of the egg. Mammals and fish are thus at the two extreme ends of evolutionary adaptation towards breathing and reproducing on land.[3] We understand why a mammal behaves differently to a fish from these key properties. From these key properties we understand why mammals roam the surface of the earth, why they look the way they look, and

why they behave the way they do.[4] In classifying animals, we use these features for a discrete taxonomy. This means that there are other features that do not qualify as defining characteristics, which are disregarded in animal taxonomy. Among these are behaviours like swimming in water, or living in groups. The key differences, obtained from looking at the extreme ends of the spectra, lead us to conclude that dolphins are mammals and not fish, even though appearances may suggest otherwise. We can draw a sharp boundary, and do not have to resort to saying that dolphins are "fishy mammals", because we recognize that the swimming behaviour of dolphins is irrelevant to their taxonomy.[5] *Defining features* and *irrelevant features* enables a proper and discrete taxonomy, making most sense of all the available data. Moreover, a taxonomy based on such features allows for an understanding that goes towards a deeper level, in this case the evolutionary pressure that came from the transition from sea to land dwelling.

I propose to undertake something similar with consciousness. What is the proper taxonomy of conscious versus unconscious vision? What are the defining features of this difference, and what features are irrelevant? And do the defining features take us towards a somewhat more fundamental level of understanding consciousness (Lamme 2010a, 2010b)? To find those features, we start from the extreme ends: the mammals and fish of consciousness research, the things most people will agree on as representing either conscious or unconscious processing.

3 The mammals and fish of consciousness

When I am awake and say I see a face, am able to report its identity; I can identify the colour of its eyes and hair, and judge its emotional ex-

---

2 Among other things, like whether they maintain body temperature or have hairy skin.
3 With reptiles and birds in between, laying amniotic eggs on the land.

4 Of course there is the occasional mammal that lays eggs (e.g., the platypus) or fish that give birth to live young (e.g., the hammerhead shark). Still, calling these mammals or fish depends on the relative weight of other defining features, such as their way of breathing, feeding, body temperature maintenance, etc.
5 In a somewhat more mathematical analogy one could take all properties of all animals in the world, and perform a cluster or factor analysis. A good taxonomy has clusters that are aligned along the primary factors. Traditional taxonomy seems to have operated in this way implicitly.

pression. There is little reason to doubt that I have a conscious sensation of that face.[6] If we study the properties of visual processing in this condition, we can be pretty sure we are studying the properties of conscious visual perception. This is our "mammal" of consciousness. We can study the properties of this species fairly easily. We can resort to introspection, verbal reports, or more strictly formalized approaches like detection or discrimination tasks. In favour of using introspection is that our introspective idea of consciousness is the very thing we are trying to explain. I would like to understand why the world looks the way it looks in my mind's eye. This is the explanandum. Even so, we should be cautious in fully "trusting" introspection,[7] and that is where more formal approaches may come in handy.

What would be the proper "fish" of consciousness? Are there conditions where everyone agrees that consciousness is absent? Dreamless sleep (Tononi & Massimini 2008) and anaesthesia (Alkire et al. 2008) seem to be good candidates, although not very useful ones, given that visual stimuli are difficult to deliver, and that one can only resort to objective measures (brain signals) to assess what is still processed or not. Awake subjects are easier to assess in that respect, but there it is hard to find truly unequivocal manipulations of consciousness. "Unequivocal" in this context means that the manipulation can truly be regarded as a manipulation of consciousness, i.e., in the case of vision is a manipulation of visibility (Kim & Blake 2005; Lamme 2006). An example of the latter

would be visual masking (Breitmeyer & Ogmen 2000; Enns & Di Lollo 2000). Here, a target stimulus is presented very briefly, and immediately followed by another stimulus, known as the mask. When properly done, this will render the target completely invisible. People will be at chance detecting presence or absence, or in judging another property of the target stimulus. It is safe to assume invisibility in masking, because there is no conceivable reason that could prevent the subject from reporting his visual percept, had he had one: the subject is sitting there, focussing his full attention to the target location, ready to push the button as soon as he sees the target. The not-seeing can therefore not be attributed to the absence of attention, to a lapse of memory, or to any other cognitive function sitting between a potentially conscious sensation and its report (Lamme 2003, 2010a, 2010b). As we are ready to believe the presence of consciousness in the case of someone verbally describing the face he sees, we should be equally ready to believe its absence in the case of masking (or dreamless sleep and anaesthesia).[8]

Another popular paradigm to render stimuli invisible is continuous flash suppression (CFS; Tsuchiya & Koch 2005. Here, the target stimulus is shown to one eye, while the other eye receives a rapid stream of brightly coloured patches, serving as a mask. This typically results in the target stimulus being rendered invisible, although stimuli may "break through" after a while.[9] A third paradigm is dichoptic masking, where two oppositely coloured stimuli are shown to the two eyes, that when properly fused combine into an invisible stimulus (Mout-

---

6 One could do so, of course, which would lead to the denouncement of consciousness as a scientific phenomenon altogether, much along the lines of eliminative materialism (e.g., Churchland 1981). Daniel Dennett, in his categorical denouncement of anything coming close to qualia or even the phenomenology of consciousness, seems to follow a similar agenda (1993). It is entirely possible indeed that consciousness is a figment of our imagination, one that will evaporate upon close scientific scrutiny. Something like that happened to 'elan vital'—the unique property of living matter—once we learned about chemistry, biology, DNA, and natural selection. For now, let's assume that consciousness exists, and is in need of an explanation. If not, I would rather not be spending my years in neuroscience.

7 One important caveat is that for introspection we have to resort to cognitive functions like attention, memory, and "internal report". This may result in both a potential underestimation of what we actually see (see for example the iconic/fragile/working memory discussion), and to an overestimation of what we actually see (as in the illusion of peripheral colour vision). This has been dealt with extensively elsewhere (Lamme 2010a).

8 Note that a proper treatment of response bias is important in this case. "Shy" subjects may feel inclined to respond "not seen" on most trials, more liberal subjects may feel inclined to respond "seen" on most trials. Only treating the responses in terms of signal detection theory (Swets et al. 1978) can truly establish the absence of any sensation (because the number of false alarms—subjects saying "seen" on trials without a target—is taken into account). From that perspective, using only partially-effective masks is not a proper method, not even when only those trials are used in which subjects reported not seeing the target.

9 A potential problem with the CFS manipulation is that "time to breakthrough" is often used as a measure of relative awareness of stimuli. Time to breakthrough is more or less analogous to a "yes" response (or hit) in a masking paradigm, and hence can suffer from response bias. CFS studies where responses are more rigorously treated in terms of signal detection theory are scarce. See Stein et al. (2011) for a more elaborate discussion on this problem with the CFS paradigm.

---

oussis & Zeki 2002; Fahrenfort et al. 2012). From all the available neuropsychological patients, patients suffering from hemianopia due to a V1 lesion (often accompanied by blindsight) are probably the clearest cases of impaired visual consciousness (Weiskrantz 1996).

I select these consciousness manipulations because they seem to be the safest bets for highlighting situations where conscious vision is really absent. The absence of conscious vision in these cases has purely visual origins. There is no other function precluding the report of a potentially present visual sensation, as may be the case in split-brain patients or neglect, or in manipulations like inattentional blindness, change blindness, or the attentional blink (Lamme 2003). The two extreme ends—the mammals and fish of consciousness—may serve as a guideline towards establishing the properties of conscious versus unconscious processing. What are the differences between awake conscious vision and vision in sleep, anaesthesia, blindsight, and the various forms of masking?

4 Categorization: From low to high level features

Above, I used the example of seeing a face. What does seeing a face mean, in terms of the visual functions being executed? Recognizing a face first of all entails that one identifies the stimulus as belonging to the class "faces", as opposed to any other class of stimuli, such as "animals", "teapots", "houses", or "letters". This is a process of categorization. Intuitively, categorization seems a key property of consciously seeing and recognizing a face. It is not, however. Since the first findings of blindsight it has been recognized that categorization can occur fully independently of conscious sensations (Weiskrantz 1996; Boyer et al. 2005). Patients without awareness of stimuli in the blind part of the visual field can nevertheless categorize these stimuli, as long as the categorization is framed in a two-alternative forced choice: is it a square or a circle, is it moving upwards or downwards, is it red or green, vertical or horizontal? In such cases, patients' responses fall well above mere chance, indicating that the categorization of

stimuli in two distinct classes is still functioning, and hence does not necessarily require consciousness.[10]

Categorization is the main function of cortical visual neurons, in that each neuron is *feature-selective*: it only responds to a stimulus when that stimulus possesses certain visual features. A Nobel prize was awarded for this finding, as it is fundamental to the operation of the visual cortex (Hubel 1982). It ranges from low level features such as spatial frequency, orientation, direction of motion, or colour to higher-level features such as the geometry of a shape or the class of an object. Each feature-selective neuron can be seen as doing a simple, often one-dimensional categorization: it signals "vertical orientation", "moving upwards", or "rectangular shape" (Lamme & Roelfsema 2000). Face-selective neurons shout "face!" (Oram & Perrett 1992). The categorization responses of visual neurons are so fundamental to their operation that they are fully independent of consciousness: most neurons are equally feature selective in anaesthesia as they are in the awake condition (Dow et al. 1981; Snodderly & Gur 1995; Lamme et al. 1998a). Feature-selective responses of neurons are mediated via feedforward connections, and visual categorization proceeds along these feedforward connections in an unconscious way (Lamme et al. 1998b; Lamme & Roelfsema 2000).

Additional evidence dissociating categorization from consciousness comes from a multitude of sources. Unseen stimuli in backward-masking are also categorized, as can be judged from the specific priming effects they may evoke. For example, a fully masked digit 7 may speed up (or slow down) responses to categorizing a second digit (or number word) as either being above or below 5, showing that the masked and unseen number (the 7) is categorized according to its numeric value (Dehaene et al. 1998).[11] Many similar examples exist. Moreover, it has been shown that masked and hence unseen stimuli evoke category-specific responses from the brain, either in the form of se-

---

10 Note, however, that categorization is typically far better for stimuli than patients—or normal subjects—are aware of.
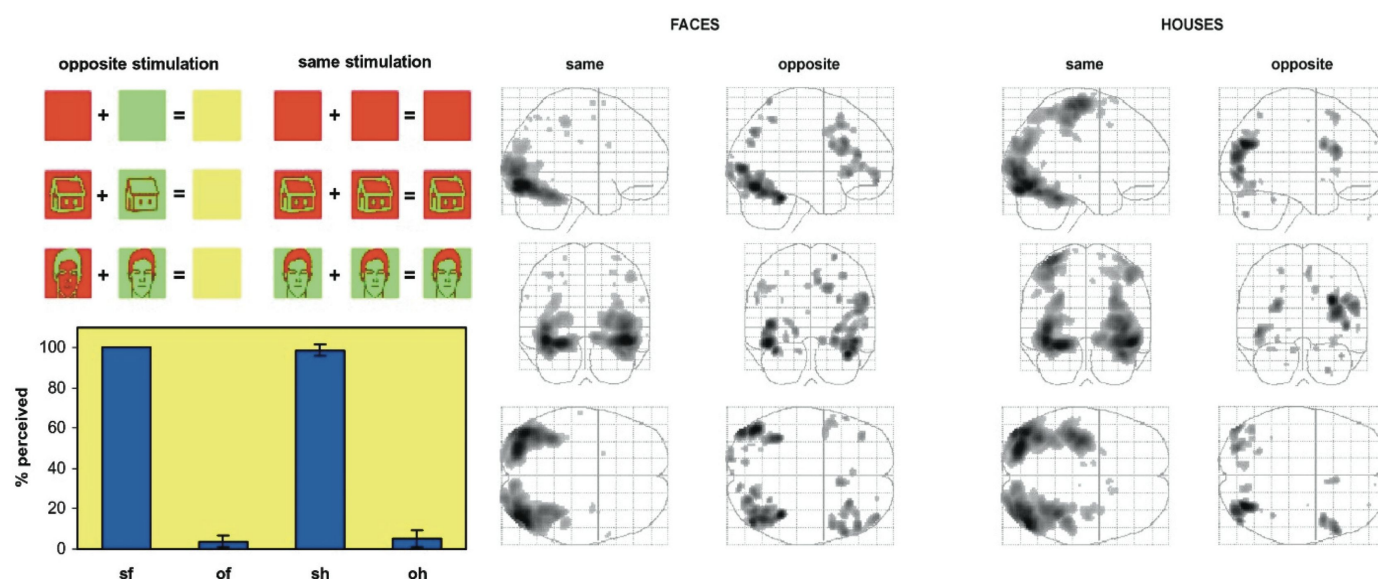11 Or more precisely: as being either below or above 5, in this experiment.

Figure 1: Faces and houses were made invisible using dichoptic masking—i.e., presenting oppositely coloured versions to each eye. Regardless of (in-)visibility, these faces and houses evoked selective activations of category specific regions of the brain (from: Moutoussis & Zeki 2002).

lective single unit responses (Rolls & Tovee 1994; Macknik & Livingstone 1998), or in the form of selective activation or category-selective regions such as the Fusiform Face Area (FFA) (figure 1) (Moutoussis & Zeki 2002; Kouider et al. 2009),[12] or in the Visual Word Form Area (Dehaene & Naccache 2001)—indicating that they are categorized up to the level of face vs non-face or word vs non-word (Dehaene et al. 2004). There is a large body of literature covering the unconscious processing of emotional valence in either faces or words (Straube et al. 2011).

Particularly far-reaching levels of unconscious categorization have been reported for behaviourally or socially relevant stimuli. Tools evoke selective activation of the dorsal stream areas—and selective priming effects—when made invisible with CFS (Fang & He 2005; Almeida et al. 2008). Faces that have their eyes turned towards the viewer break from CFS sooner than faces that are turned away—a finding that is probably explained by the fact that faces turned towards the viewer pose a very relevant or even threatening social signal (Gobbini et al. 2013). Similarly, the gender of naked bodies is processed during CFS (Jiang et al. 2006). Also, the mismatch between object categories is identified for stimuli made invisible using CFS: scenes with mismatching objects (e.g., a cook taking a chess-board out of the oven instead of a dish) break from CFS sooner than matching scenes (Mudrik et al. 2011).

The latter finding is related to various non-visual "categorization" processes that occur for invisible stimuli: it has been shown that masked stimuli travel throughout the brain, even reaching high-level areas involved in inhibitory cognitive control, response error selection, or evidence accumulation, exerting high-level cognitive effects (Van Gaal & Lamme 2012). So invisible stimuli not only activate visual categorization processes, but also activate extremely high-level and very abstract categories such as the stimulus being a "stop signal", an "error", or "evidence for a right button press".

---

12 It is unclear to what level invisible faces are processed exactly. Clearly, face/non-face categorization takes place for masked stimuli (see below), but whether face identity is also preserved depends on the exact experiment. Some find face-*identity*-specific priming and suppression of activation of the FFA and related face-selective-regions for backward masked stimuli (Kouider et al. 2009). However, this effect was only present for famous faces, not for unknown faces, showing that it may not be identity itself that is processed but "level of fame" or something similar. Others have made faces invisible using CFS, and found that face-specific adaptation only occurred for visible, and not for invisible faces (Moradi et al. 2005). The two studies are hard to compare, partly because of the different techniques used to make faces invisible (masking vs. CFS), but mostly because the latter used an adaptation effect as independent variable. It may be that unconscious categorization still occurred, yet did not result in learning (e.g., Meuwese et al. 2013; Meuwese et al. 2014).

From a neural perspective, categorization is feature selectivity, which may range from very simple to highly complex features and categories. This kind of categorization proceeds entirely independent of consciousness.[13] So how does conscious recognition differ from categorization? To answer this question, we have to take a closer look at categorization responses. What a face-selective cells does, is to categorize a face as belonging to the class of faces versus non-faces. That's all. When we consciously see a face, however, we do much more than this: we classify the stimulus as a face, but at the same we identify its shape, colour, identity, and emotional expression. So we distinguish between "that brown face of my sad-looking friend Peter" and very many other faces—and also between that face and millions of other potential visual stimuli.

Gulio Tononi uses the metaphor of a photo-diode to illustrate the point (2008, 2012). For a photo-diode a black screen is different from a white screen. That's a distinction it can make. The photo-diode carries information about the brightness of the screen, so its signal carries one bit of information (or a few bits, depending on its sensitivity). For us, however, consciously seeing a black screen is very different. Seeing the black screen implies that we distinguish it from a grey screen, a red screen, a black table, a green house, a pink face, a dog, a sound, a feeling, or any other sensory event that would have been possible. Consciously seeing the black screen thus carries a huge amount of information, as it excludes an almost endless set of alternative sensations. And that makes seeing "that brown face of my sad-looking friend Peter" very different from what a face-selective neuron does when it signals "face". The neuron behaves much like the photodiode, in that it signals presence or absence of a feature along a single dimension. That is because neurons tend to combine feature-selectivity with invariance

for other features: a face-selective cell signals faces regardless of colour, size, identity or expression (Rolls 1992).[14]

Tononi proceeds from a photo-diode to the photo camera as a metaphor for explaining another central feature of conscious sensations (2004, 2008, 2012). He argues that the critical difference between a conscious representation in the human mind and what happens in a camera is that in the camera information is distributed and not integrated. Each and every pixel signals a particular level of luminosity, but it does so entirely on its own. It does not "know" what other pixels are doing. To the camera it would not matter if all the pixels were cut apart and became separate cameras. Conscious sensations, on the other hand, are integrated.

It thus seems that to find for visual operations that are more closely linked to consciousness, we must look for something beyond basic categorization. We must look at processes where the individual pixels in our camera—the billions of neurons each signalling particular features—are interacting, and are integrating their information.

5 Interference: A loss of independence

The pixels in the "camera of the human mind" do not work independently. A strong case in point are illusory brightness or colour shifts. Patches of the exact same brightness may be perceived as entirely different, depending on their surroundings, and depending on the global configuration of brightness and contrast. A striking example is the cylinder with checkerboard illusion shown in the right half of figure 2. Similar illusions exist in the domain of colour (figure 2, left). Relatedly, everyone who has ever tried to paint a picture has experienced that it takes an astonishingly rich palette of reds, purples, browns, yellows, and even greens or blues to construe a veridical depiction of a simple red apple. The unitary experience of see-

---

13 Another illustration of the separation between feature selectivity and conscious experience is the observation that many neurons signal features of which we are not aware: V1 neurons signal the orientation of gratings that are of too finely spaced for us to perceive (He et al. 1996; Foster et al. 1985), respond to 3D disparity where we do not see depth (Cumming & Parker 1997), or signal invisible temporal frequencies (such as the flickering of light beyond the flicker-fusion frequency of about 15–25Hz, Maier et al. 1987).

14 Responses are modulated by such features, but typically this happens only after some delay (Sugase et al. 1999). The initial feedforward response is typically fully governed by a basic feature, like face vs non-face. Later on, responses are modulated by face identity or expression, and this is mediated by horizontal or recurrent interactions between neurons. We then enter the domain of feature integration, which *is* a hallmark of conscious recognition; see below.
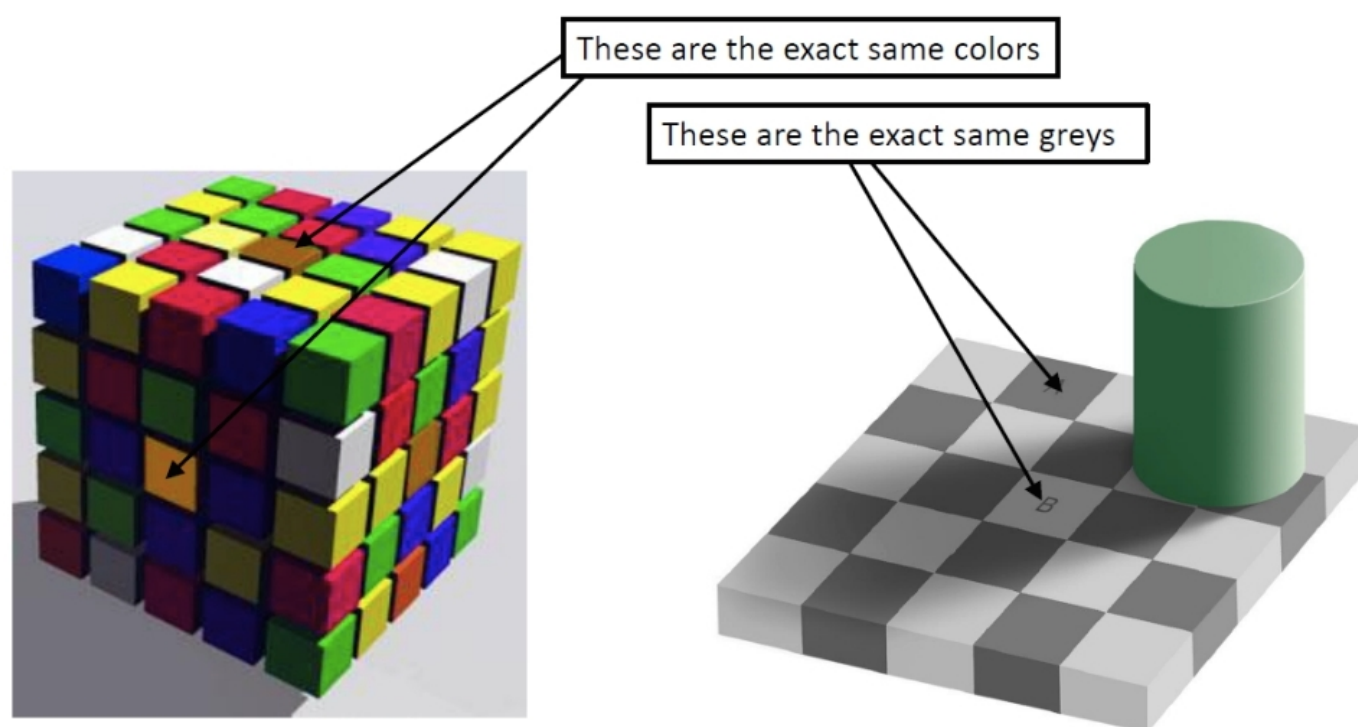

Figure 2: Two strong shifts in the perception of colour and brightness. Although the indicated patches are identical, they are perceived as having quite different colour and brightness. Visit Michael Bach's website (http://michael-bach.de/ot/) for these and many other examples.

ing a red apple is in fact composed of the detection of a multitude of wavelengths, all interacting to compose that one colour. Only with extreme focused scrutiny (or by covering surrounding elements) are we able to isolate the elements that make up our unitary conscious experiences.

Another illustration is the phenomenon of colour constancy. When we look at a bowl of fruit in the blue morning light the spectral composition of wavelengths reflected from the fruits is very different from the wavelengths coming from the fruits at sunset (figure 3). Nevertheless, we see the banana or the apple as having the same colour whether it is dusk or dawn. Our visual system is not interested in the wavelength coming from fruits; it is interested in their potential taste or edibility. Therefore, it discounts the illumination, and computes "colour", which is a property of the object, rather than of the light coming from it.[15] Colour is not wavelength; colour is a meaningful property of

objects that is based on wavelengths, yet transcends it.

To what extent do these phenomena depend on consciousness? Harris et al. (2011) studied a brightness illusion much like that in figure 2. Two circles were shown, of either the same or different brightness. By placing these circles in a dark and bright surround respectively, their brightness suffered from an illusory shift. In the critical condition, the surrounds were made invisible by presenting them to one eye, and filling the other eye with a continuously flashing Mondrian stimulus. This resulted in CFS of the surrounds. Cleverly, the two circles were shown in both eyes, so remained visible throughout. Regardless of the CFS-induced invisibility of the surrounds, the circles still showed illusory brightness shifts.[16]

The neural mechanisms of illusory brightness perception were studied extensively

---

15 It probably discounts the illumination by very much the same mechanisms as the illusory brightness shifts discussed above (via inhibitory lateral interactions). However, precise neural mechanisms may be different, as might be the cortical level at which neural responses reflect colour rather than wavelength.

16 It must be noted that in this experiment, the surrounds were not always fully invisible. In 86% of the trials, subjects reported not seeing the surrounds. Only these trials were used for the analysis. Within these trials, discrimination of the background (is the darker half left or right?) was at chance level, leading to the argument that indeed there was a full absence of awareness of the surround.

Figure 3: These images show a bowl of fruit photographed in three lighting conditions—artificial light (left), hazy daylight (middle), and clear blue sky (right). Notice the marked variation in colour balance caused by the spectral properties of the illuminant. We are not normally aware of this variation because colour constancy mechanisms discount illumination effects (image and legend from http://www.psypress.co.uk/mather/resources/topic.asp?topic=ch12-tp-04).

in the macaque monkey and cat visual cortex. It was found that perceived brightness (modulated by flanking regions) influenced neural responses in area V1 of the cat, but not at earlier stages such as the LGN or the optic tract, thereby showing a gradual progression from physical brightness to perceptual brightness in the visual pathways (Rossi & Paradiso 1999; Rossi et al. 1996). Using the Cornsweet brightness illusion,[17] it was found that in the monkey's visual cortex, V2 cells represents surface brightness whereas V1 cells do not, pushing the level at which perceived brightness is calculated somewhat higher (Roe et al. 2005). Either way, these results were recorded in anaesthetized animals, showing their independence from consciousness.

How the visual system goes from the detection of wavelength towards the representation of colour is still a topic of controversy. Initially, there was thought to be a modular progression from V1 cells encoding wavelength towards V4 cells encoding colour. That view was challenged by various findings showing that the responses of V1 cells are influenced by surrounding hues. The view that V4 is the "colour module" has also been challenged, in part by strong disagreement on the homology between monkey V4 and alleged human counterparts.[18] Moreover, the coding of colour is intricately linked to the coding of object shape, and hence can no longer be viewed as a simple "add-on" to our visual percept.[19] It is now thought that the perception of colour depends on the interaction between neuronal groups, or is best understood as a population code (Shapley & Hawken 2011).

Given this controversy, it is difficult to know to what extent colour perception depends on consciousness. Many of the recordings in monkey visual cortex were performed in awake animals, some in anaesthetized animals (Shapley & Hawken 2011). A clear-cut difference in results between the two conditions is hard to establish. A remarkable finding is that blindsight patients report no conscious sensation of colour, yet may have spectral sensitivity curves that have a similar shape in the lesioned and intact hemi-fields (Stoerig & Cowey 1989). Spectral sensitivity is, however, mostly carried by wavelength. Similarly, patients with cortical colour blindness (achromatopsia) do not consciously perceive colour, yet can detect objects or patterns based on wavelength contrasts (Cowey & Heywood 1997). Colour constancy mechanisms, on the other hand, are absent in the lesioned hemi-field of blindsight patients (Barbur et al. 2004; Barbur & Spang 2008), and

---

17 In this illusion, two surfaces of identical brightness are perceived as having different brightness, because there is a contrast edge between them.

18 I am not even going to dare mentioning their names here.

19 The fact that black-and-white photography works so well, has led us to believe that colour is a feature that is "painted" onto objects, as a sort of extra, independent of any other feature. We are now coming around from this view. For example, to compute the colour of an object, the object's shape has to be taken into account, otherwise shadings would be misinterpreted. Object identity also influences colour perception: a brownish colour on a banana will be seen as more yellow than it would on a tomato.
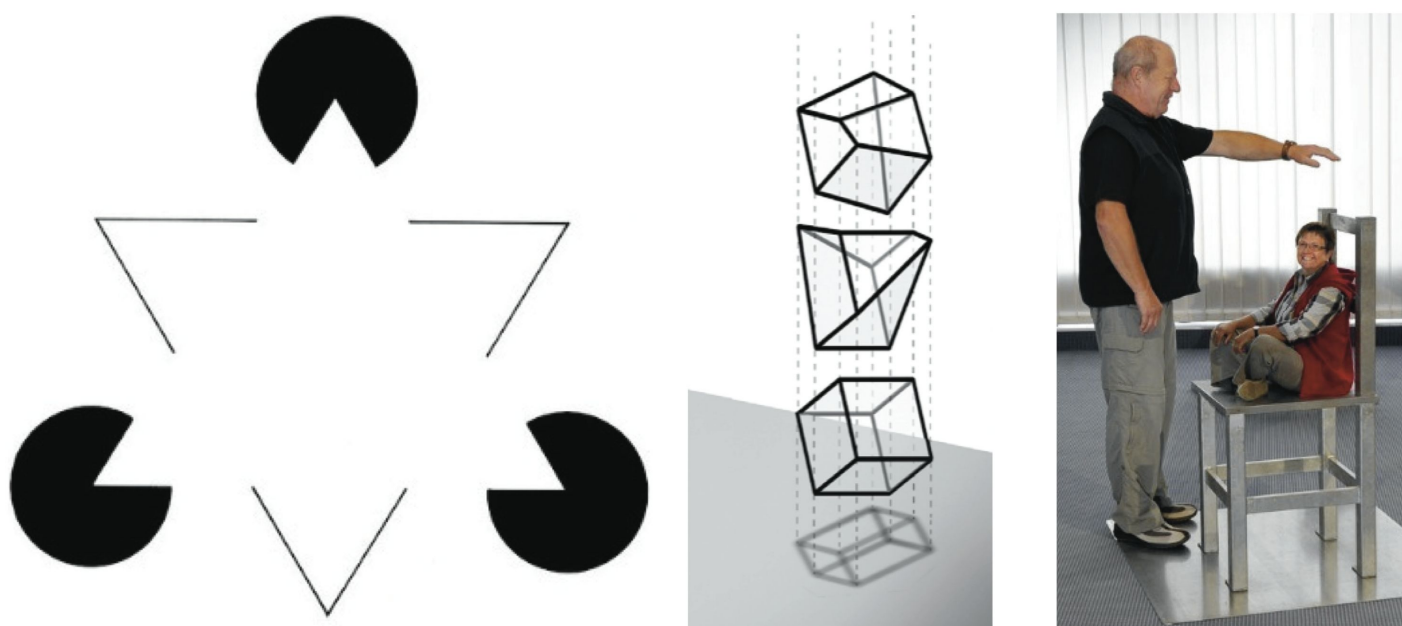

Figure 4: Left: the Kanizsa triangle. Note the illusory brightness increase inside the region of the illusory triangle. Middle: the 2D projection of a cube can in fact originate from a multitude of 3D objects. We regularly interpret it as a cube, however. Right: we see the woman as small, despite our cognitive ability to realize that "this cannot be true". Our 3D "priors" force us to see her as small (from https://richardwiseman.wordpress.com/2009/09/09/great-table-illusion/).

hence seem more closely linked to conscious perception.[20]

The difference between perceived colour and wavelength, and its relation to conscious vision, has been directly addressed in a masked priming experiment. In this experiment, subjects were shown desaturated blue, green, or white coloured disks. Perceptually, the white was closer to the blue disk. From the point of view of the phosphor activations on the monitor screen, on which the disks were shown (i.e., their "wavelength composition"), the white disk was, however, closer to the green disk. What was studied was the effects of these disks when they acted as primes for a subsequent colour discrimination. It was found that masked, and hence invisible white disks, acted more like green primes than like blue ones. Visible white disks, on the other hand, acted more like blue primes than like green ones (Breitmeyer et al. 2004; Breitmeyer et al. 2007). Apparently, unconscious priming acts on wavelength similarity,

whereas conscious priming acts on perceived colour similarity.

All in all, it remains difficult to assess the relation between consciousness and phenomena like brightness or perceived colour illusions, or mechanisms related to colour constancy. Perceived brightness seems to depend on largely unconscious mechanisms, and on fairly low level and short range mechanisms. The transition from wavelength analysis to the perception of colour is more likely to accompany the transition from unconscious processing to conscious vision. A firm conclusion, however, relies upon settling the debate about mechanisms of colour perception and their neural substrates in humans and animals, and more direct experimentation on how these mechanisms are affected by manipulations of consciousness.[21]

6 Inference: Beyond the input

In the phenomenon of colour constancy we have already seen a hint of another visual function.

---

20 This argument is, however, weakened by the fact that other long-range colour interactions remain in the blind hemi-field (Barbur et al. 2004), and by the finding that colour constancy mechanisms may depend on fairly early, monocular mechanisms (Barbur & Spang 2008). Moreover, it is reckoned that several colour constancy mechanisms exist, some of which are based on retinal adaptation mechanisms (Kamermans et al. 1998).

21 Obviously, these empirical issues about colour perception and consciousness have very direct consequences for many philosophical debates as well, given the many thought experiments that rely on colour perception and the whole notion of qualia.

Colour is not about the wavelength coming from objects. It is a property of objects that we infer from wavelengths. At some point, conscious perception starts to diverge from the mere physical properties of the input, in a process we call *inference*. There are many more examples of inference, and many visual illusions capitalize on the fact that our visual mechanisms are constantly trying to make sense of the world. Figure 4 shows the famous Kanizsa triangle. The minimal, strictly physical interpretation of the image is that of three Pac men pointed at each other and three arrowheads pointing outwards. But our perceptual interpretation goes beyond this, in that we *see* a white triangle hovering over three black circles, occluding another outlined triangle. The illusory triangle is seen as slightly brighter than its surround, and illusory contours mark its "borders".

This process of inference seems to strongly fit the intuitive difference between a camera and conscious vision. It requires the integration of multiple "pixels", their interaction, and their interpretation beyond what is strictly given by the image itself. And it is in this last aspect in particular that prior knowledge about the world comes into play, and starts to interfere with the stimulus-driven feature-selective categorization of the input.

The Kanizsa triangle can be seen as a specific example of the more general propensity of the visual system to arrive at a representation of surfaces in 3D space (also called the 2.5D sketch). In that representation we seek the most natural interpretation, consistent with our existing experience of how things are in the world. It is simply much more likely that there is a triangle covering circles than that there are three Pac men that happen to be facing each other at exactly 60° angles. The triangle interpretation is generic, whereas the Pac men one would be accidental (Albert & Hoffman 2000). Nakayama & Shimojo (1992) studied various configurations of 3D stimuli, and found that our visual system always strives towards the interpretation that is most generic, i.e., that would least depend on an accidental viewing position of the observer. Interpretations that would not change when the observer happened to shift position are fa-

voured, given that we are constantly moving relative to objects. For example, the 2D image of a cube can in fact arise from an infinite number of shapes (figure 4, middle), yet we tend to favour the "cube" interpretation because it is the most generic one.

Another way of putting it would be to say that the cube interpretation fits our common experience, in that most of the time, these kinds of 2D projections arise from regular 3D cubes: it is the most ecologically valid interpretation. In a modern guise, this aspect of inference is formalized as a Bayesian approach, where vision uses a set of prior probabilities to arrive at the most likely 3D interpretation of a 2D image. The cube has the highest prior, compared to the more irregular shapes. Illusions like the Ames room (where someone changes size when he walks from one corner to the other), or the size illusion shown in figure 4 (right) capitalize on these assumptions: we assume that rooms have rectangular floors and walls, we assume the woman is sitting on a chair. These assumptions are so strongly embedded in our visual hardware that even in the face of strange consequences, such as people growing in size within a few steps or a man holding his hand over a mini-woman, this inference is maintained.

Many more illusions display non-veridical inferences. In the Ebbinghaus illusion, the perceived size of a disk depends on the size of surrounding disks. In the Ponzo and Müller-Leyer illusions we see line segments as having different lengths, while in fact they are the same. These illusions show that the size of an object is an inference that we draw from its context, rather than from the space it occupies on the retina.

To what extent does inference depend on conscious vision? When we have to pick up the disks in the Ebbinghaus illusion, it appears that our hands open at a pre-grip aperture that is in accordance with the disk's actual size, not its perceptual size. Apparently, size context effects influence perception, and not automatic action—which has led to the idea that we have two largely separate neural pathways, one transforming visual input into conscious perception, and the other translating visual input into automatically guided action (Goodale & Milner 1992).
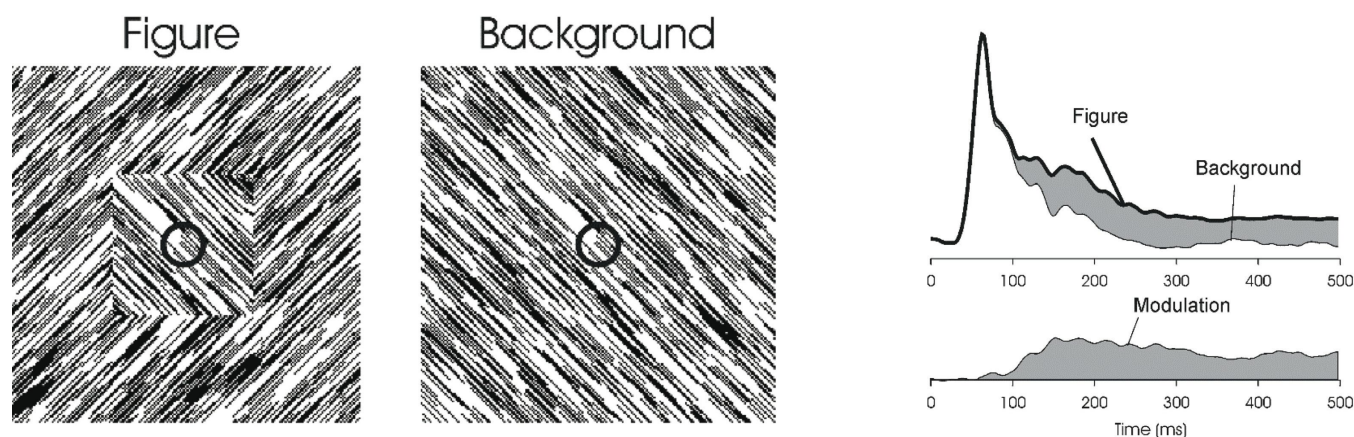
Figure 5: On the left, we see a textured square overlying a textured background. This is because we automatically group all line segments with the same orientation into one object, and segregate it from the line segments with the orthogonal orientation. The small circle represents the receptive field of a V1 neuron, that would not be able to differentiate between the "figure" and the "background" stimulus, because the line segments within that receptive field are identical. Indeed, V1 responses are identical up to about 100ms after stimulus onset. Beyond that, the two responses start to diverge, however, indicating that the response of the V1 neuron is modulated by the perceptual context of what is within its receptive field (Lamme 1995; Lamme et al. 2000).

There is more evidence linking perceptual inference to conscious vision. Harris (Harris et al. 2011) studied whether Kanizsa triangles were still inferred when the inducers were made invisible using CFS. The same setup was used that showed the presence of brightness illusions under CFS (see above). Subjects had to indicate whether the triangle in the suppressed eye was pointing left or right. They were at chance level, indicating that the Kanizsa-type inference depends on consciousness. Another study, however, found that Kanizsa triangles broke through CFS earlier than control stimuli with inducers pointed outwards (Wang et al. 2012), suggesting that Kanizsa-type inference does occur pre-consciously.

At the single neuron level, the detection of illusory contours has been studied quite extensively. Initially, it was found that V2 cells respond in an orientation-selective manner to Kanizsa-type illusory contours (Von der Heydt et al. 1984). More recently, other areas have been shown to be involved as well (Sáry et al. 2008)—area V4 in particular (Cox et al. 2013). And in human neuroimaging studies it was found that Kanizsa-type illusory contours activate many early visual areas (Seghier & Vuilleumier 2006). All these studies used awake animals or humans, so it is difficult to infer whether these responses depend on the conscious state.

Marcel studied the processing of illusory triangles in two blindsight patients. Two inducers were presented in the sighted hemi-field, while one critical inducer was presented in the blind field, either completing the triangle or not. Completed triangles were detected far above chance (~80%), while the detection of the inducer shape was at chance. Moreover, one of the subjects described the illusory triangles as "brighter", "out there on the screen" and "on top of something" (Marcel 1998).

All in all, the relation between inference and consciousness is unclear, mostly because fairly little work has been done as yet to study the relation directly (i.e., to study the effect of consciousness manipulations on inference and its neural correlate), but also because much of the work that has been done focuses on a single (though very important) phenomenon: the Kanizsa triangle.

7 Integration: Feature grouping and segregation

Both in interference phenomena such as brightness or colour shifts and in inference phenomena like the Kanizsa triangle we see some

aspects of the integration of information. Visual responses go beyond the encoding of individual pixels, and start to influence each other, either on the basis of more or less hardwired lateral interactions, or on the basis of the incorporation of prior knowledge. In the end, conscious vision seems to be about reaching *full* integration.[22] We have one visual percept, where all information is combined.[23] This is a property of conscious vision that has interested scientists for a long time. Gestalt psychologists formulated a multitude of laws, along which image elements may be combined into larger wholes (Rock & Palmer 1990; Wagemans et al. 2012). In this grouping process, all features, together with their interactions, inferences, and meanings are combined into a final percept: the thing we see, the whole scene containing shapes, objects, and backgrounds. This is a highly dynamic process in which various Gestalt laws may compete for one interpretation or another, and where subtle changes may influence the meaning of pixels at long distances. We enter the domain of feature integration, grouping, binding and segregation. In short; the domain of perceptual organization.

Two levels of integration may be distinguished, where a subdivision between "base groupings" and "incremental groupings" may be useful (Roelfsema 2006). Base groupings are those that depend on the fact that some feature combinations automatically ride together. An orientation-selective cell in the primary visual cortex, for example, is often at the same time

also direction-selective. It may be tuned to particular binocular disparities as well. And it will have a limited receptive field. So the firing of that neuron already goes beyond a one-dimensional feature-detector, beyond the photo-diode. It signals an orientation, moving in a particular direction, at a particular 3D depth, and located in some part of the visual field. Such base groupings exist for many feature combinations, such as colour and shape (e.g., V4 cells), or motion and disparity (e.g., middle temporal, MT, cells).

Another type of base grouping is visible in the feature selectivity of a particular cell, where we may recognize the combination of feature-selectivity of cells at earlier levels. From the start, Hubel and Wiesel recognized that orientation selectivity could be viewed as a convergence of information from retinal ganglion cells lying in a row. The feedforward convergence of information from orientation selective simple cells leads to the receptive field structure of complex cells, which are orientation and direction selective (Hubel & Wiesel 1968). Many higher-level feature-selective cells can be seen as converging information from lower level cells (Tanaka 1996).

Base grouping does not depend on consciousness. The combined feature selectivity of neurons, as well as high-level feature selectivity based on the feedforward convergence of lower-level feature selectivity are still present in anaesthesia or masking (Lamme & Roelfsema 2000; Roelfsema 2006).

Of a very different nature are "incremental groupings". Here, the information from separate neurons has to be combined to obtain a higher level categorization. A good example is texture based figure-ground segregation, shown in figure 5 (Lamme 1995; Zipser et al. 1996). Here, we automatically perceive a textured square overlying a textured background. This is entirely due to the fact that the centre square is made up of line segments of a particular orientation, different from the line segments that make up the background. There is no luminance difference or any other cue that gives the square "away". Line segments of one orientation are automatically grouped into a coherent surface—the square—that is segregated from the surface that is

[22] Tononi similarly argues that consciousness always strives for "maxima of integrated information", for which he uses the metaphor of the internet (2012). Like the brain, the internet is a highly interconnected structure where information travels from one part to other parts. In contrast to the brain, however, the internet is designed to transfer information from one specific part of the net (computer A) to a specified other part of the net (computer B), and it would in fact be rather counterproductive if this information were influenced by other information flowing from computers C to D or E to F. At another moment information may flow from A to C or D or F. The internet therefore does not strive for "maxima of integrated information", whereas the brain typically does. Focussed attention, in such a view, would then be in fact a mechanism that counters this propensity towards maximally-integrated information, and which enables the brain to operate more strongly along the principles of the internet.

[23] This is in fact such a strong intuition that it has led us to believe for a long time that consciousness must be some place in the brain "where it all comes together". Descartes envisaged the pineal gland as such a place, and hence theories that lean towards such an explanation of consciousness are often said to suffer from the fallacy of the "Cartesian theatre".
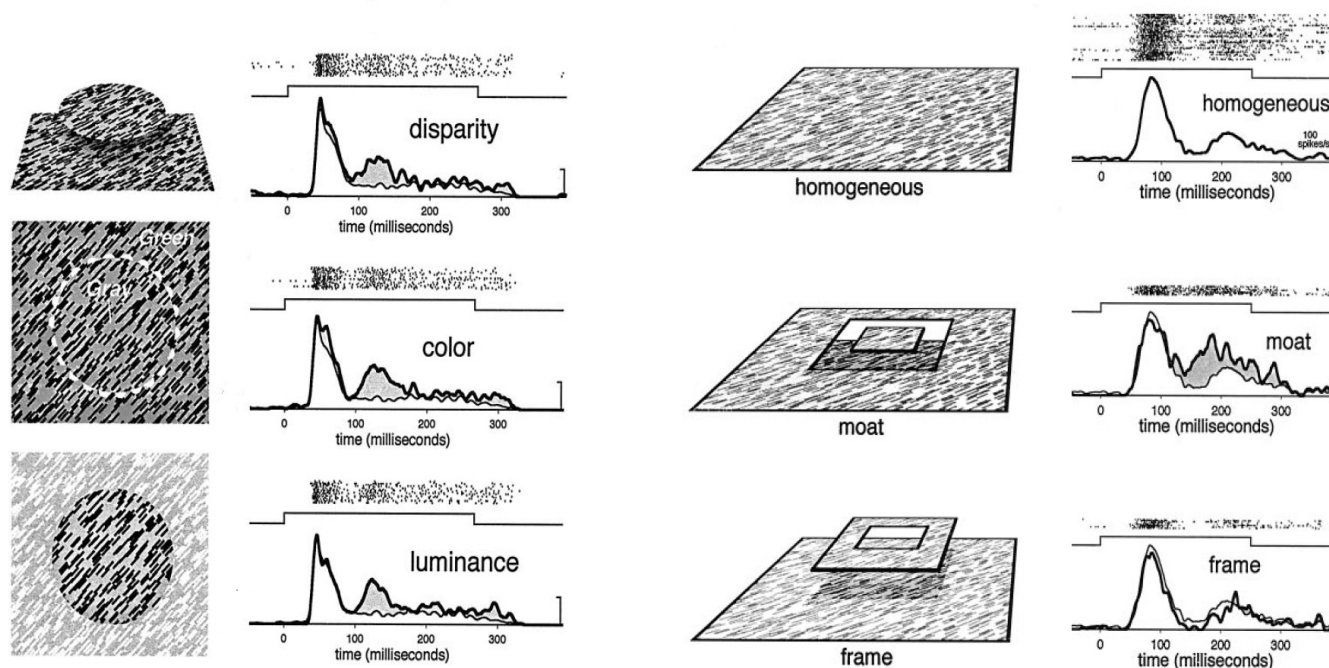
Figure 6: Contextual modulation of V1 responses follows the global perceptual interpretation of images. In all cases, the V1 receptive field is stimulated with the exact same line segments. When these line segments belong to a homogenous background texture, a response indicated by the thin line is given. Left: when the line segments belong to a figure that is defined by differences in disparity, colour, or luminance, the responses are larger. Right: differences in 3D disparity were used so that the patch of texture was either part of a figure square "floating in a moat behind it" or in the background with a "frame" hovering in front of it. The contextual modulation always followed these perceptual interpretations, in that "figure" interpretations always evoked larger responses (Zipser et al. 1996).

formed by line segments of the other orientation —the background. Orientation-selective neurons in V1 typically have small receptive fields, which would only cover a small part of either the figure or background. The grouping of line segments into coherent surfaces, segregating from each other, requires the integration of information from a large set of separate V1 cells. This constitutes "incremental grouping" (Lamme & Roelfsema 2000; Roelfsema 2006).

The neural basis of the integration of image elements into larger units, and the subsequent segregation of such units into figure and ground has been studied extensively at the single unit level, both in anesthetized and awake monkeys. The key finding is that of "contextual modulation", where the response of a neuron to a particular feature within its receptive field is modulated by the larger perceptual context of that feature (Lamme 1995; Zipser et al. 1996; Lamme et al. 1999). In the example of figure 5, the small circle represents a V1 receptive field,

which typically has a size of ~1 degree of visual angle. From the "point of view" of that receptive field, there is no difference between the "figure" or the "background" stimulus: in both cases, identical line segments cover the receptive field, and if the neuron were just signalling the presence of this feature ("left diagonal orientation present"), the responses of this neuron should be identical for the two stimuli. Indeed they are, as shown in the panel on the right, showing fully overlapping responses, until ~100 ms after stimulus onset. At that point, however, the responses for figure and background start to diverge. Apparently, information on the context of the line segments starts to influence the response, so that the response is larger for the "figure" than for the "background" context (Lamme 1995).

These kinds of figure-ground modulations follow the perceptual interpretation of scenes to a large extent. For example, when figure-ground relationships are ambiguous, or reversed, the

Figure 7: Contour grouping. In all cases shown here, oriented image elements are grouped together to form either a line (left), a circle (center), or an animal (right). They group according to the Gestalt principles of proximity, similarity, and colinearity. These stimuli were also used in neurophysiological experiments, typically showing that elements that group and segregate evoke larger neural responses than isolated or background elements.

modulation follows the globally-organized percept, rather than local orientation differences or gradients (figure 6, right panel) (Zipser et al. 1996; Lamme & Spekreijse 2000).

The perceptual grouping of image elements into larger units follows certain rules and principles, the formulation of which was the largest contribution of the Gestalt psychologists to modern vision theory (Wagemans et al. 2012). Among these Gestalt laws of perceptual organization are "similarity" (elements that look alike will be grouped), "common fate" (elements that go together in time, e.g., move together, will be grouped), "proximity" (elements that are close together will be grouped), and "good continuation" (elements that lie along a smooth line will be grouped). Contextual modulation of V1 neurons behaves according to these rules, in that elements that share luminance, colour, disparity, orientation, direction of motion, or co-linearity induce facilitatory interactions (figure 6 & 7) (Lamme et al. 1993; Lamme 1995; Kapadia et al. 1995; Zipser et al. 1996; Lamme et al. 2000).

How does Gestalt grouping and segregation depend on consciousness? To some extent, contextual modulation seems to survive during anaesthesia. This is, however, largely limited to fairly short range interactions between neurons, barely beyond or entirely within the receptive field (Allman et al. 1985; Gilbert & Wiesel 1992; Nothdurft et al. 1999). More long-range interactions, and interactions that express more global scene interpretations can only be recorded in awake monkeys (Knierim & Van Essen 1992; Lamme 1995; Kapadia et al. 1995; Zipser et al. 1996). For example, the figure-ground specific modulation of V1 responses shown in figures 5 and 6 (and structure from motion defined figure–ground modulation) is fully absent when monkeys are anaesthetized. At the same time, the orientation and motion selectivity of these neurons (i.e., their ability to categorize certain features) is not affected at all (Lamme et al. 1998a).

Similarly, backward masking disrupts figure–ground modulation. In monkeys, the visibility of texture orientation defined figure-ground targets was manipulated by masking with a stimulus consisting of randomly-positioned texture-defined figures (figure 8). The animals were at chance in detecting the location of the target figure for stimulus-onset asynchronie (SOA) of up to 50 ms (i.e., 50 ms between the onset of the target figure and the mask). At larger SOA's, behaviour quickly rose to ceiling. Figure-ground contextual modulation followed the same pattern: absent up to and including SOA's of 50 ms, and increasingly present at longer latencies. At the shorter latencies, however, V1 neurons still responded vigorously to the texture patterns in an orientation-selective manner,
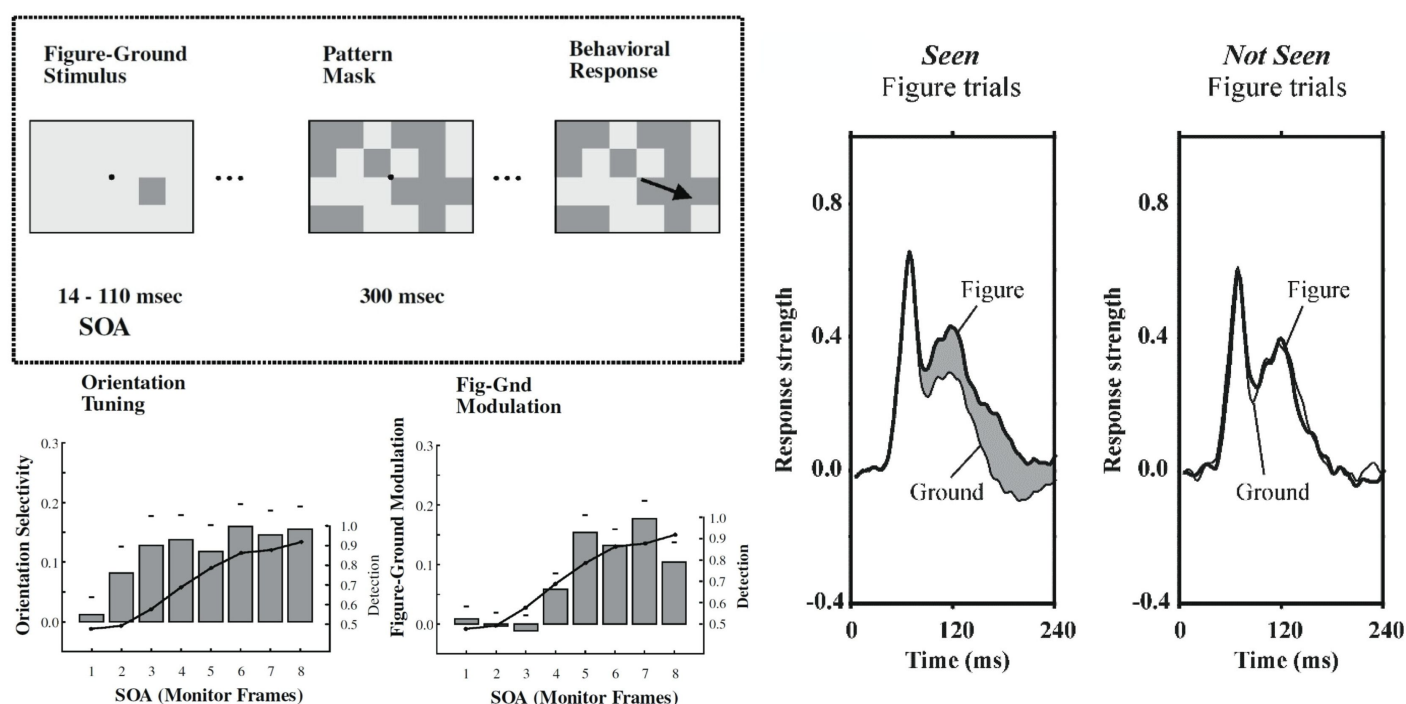
Figure 8: Left, above: textured figure-ground squares (like shown in figure 5) were presented either left or right of the fixation spot, and monkeys had to indicate their position with an eye movement. The figure targets were masked with a pattern of randomly-positioned texture squares. Left, below: the graphs show—for different SOA's—the ability of monkeys to correctly identify the position of the squares (line graph) versus the strength of either orientation-selective responses or figure-ground modulation (bars). Monkeys do not see the figures at SOA's of up to 3 frames (~50ms), and likewise, contextual modulation is absent in those cases, whereas orientation selectivity is not (Lamme et al. 2002). Right: monkeys had to indicate the presence or absence of textured figure targets by making an eye movement or deliberately maintaining fixation. When figures were not seen, contextual modulation was absent (Supèr et al. 2001).

showing that lower level classification was still present for unseen orientations (Lamme et al. 2002). Similar results were obtained in human subjects using EEG responses (Fahrenfort et al. 2007).

Contour grouping, as displayed in figure 7, is particularly susceptible to masking. When these displays are temporally alternated, so that each element rotates 90^0 in successive displays, a strong masking effect is observed.[24] Depending on the angle between elements forming a contour, visibility drops to chance at alternation frequencies between 12 and 1Hz. This implies that the integration of these contours takes between 80 to 1000 ms (Hess et al. 2001).

Zipser used dichoptic masking to render orientation-defined figures invisible. Figure-ground stimuli like those of figure 5 were shown to the two eyes of awake and fixating monkeys, yet with opposite orientations in either eye. As a result, the dichoptically-fused image consisted of cross-like elements, in which a figure was no longer visible.[25] Figure-ground modulation was absent in this case (Zipser et al. 1996). In a similar experiment in human subjects, Fahrenfort used face stimuli that were defined by oriented texture differences. A face was present in each image presented to the two eyes. Yet when binocularly combined, the face disappeared in the fused percept. He compared the neural signals obtained for such stimuli to responses to similar stimuli where binocular fusion resulted in a vis-

---

24 This manipulation is a combination of backward and forward masking, and also somewhat reminiscent of dichoptic masking, in that in subsequent displays images with the opposite orientation contrast are shown. See the two images of figure 9, but then not presented to the two eyes but in rapid alternation.

25 A similar setup was used in the curious case of alleged "blindsight in normal observers". In one of the experiments in that paper, target figures were made invisible using the same manipulation of dichoptic presentation of orthogonally-oriented elements. It was claimed that despite their subjective invisibility, subjects were able to localize the targets above chance, just as blindsight patients do for unseen stimuli (Kolb & Braun 1995). The findings were not replicated, however (Robichaud & Stelmach 2003).

ible face (figure 9) (Fahrenfort et al. 2012). A striking finding was that visibility (although rigorously checked behaviourally) had no effect on the ability of the Fusiform face area to distinguish between face and non-face stimuli, once more corroborating the independence of categorization responses and consciousness. In addition, invisible face stimuli could be classified from neural responses when training the classifier on visible stimuli and vice versa. The difference between visible and invisible binocular faces was found in the fact that visible faces evoked strong recurrent interactions between the FFA and earlier visual areas, both expressed in the fMRI signal (assessed using psychophysiological interaction analysis with the FFA as a seed), as well as in the EEG signal (showing a larger amount of theta, beta and gamma synchronization, and the presence of figure-ground modulation only in the visible condition).

The most direct relation between contextual modulation and consciousness was perhaps demonstrated by Supèr et al. (2001). Monkeys were shown oriented texture figure-ground targets at different locations, and had to signal their presence by making an eye movement towards their positions. Importantly, however, in 20% of the trials, no figure was presented at all, and the monkeys had to maintain fixation on those catch trials for the duration of the stimulus.[26] Indeed the monkeys refrained from making eye movements on catch trials (as they were trained to do). But also on some 8% of trials in which a figure *was* presented they maintained fixation, as if to say "I did not see a stimulus here". There was a striking difference in the level of contextual modulation for seen versus not-seen figure targets: modulation was fully absent for not-seen figures (figure 8). Seemingly, on some trials contextual interactions spontan-

eously fail to develop, and the result is that figure targets were invisible.[27]

That brings us to the question of neural mechanisms. Seemingly, the visual functions of perceptual organization, grouping according to Gestalt laws, and figure-ground segregation all depend strongly on the conscious state, and on the objective (or subjective) visibility and perceptual interpretation of the stimulus. Do these functions have similar neural mechanisms? There has been much debate on the neural connections underlying contextual modulation effects. Given the latency of the effects (typically several milliseconds after the initial categorization or feature response) it was originally hypothesized that they depended on feedback signals from higher-level visual areas (e.g., V4, IT, MT, etc.) toward lower levels (e.g., V1, Zipser et al. 1996). Experiments using cooling or lesioning of higher-level areas gave mixed results. Local inactivation of V2 using GABA injections had no effect on short- to medium-range contextual effects in V1 (Hupé et al. 2001). Cooling area V5/MT, on the other hand, had effects on figure-ground signals in V1, V2, and V3 (Hupé et al. 1998). These effects, however, worked on the early part of the response, and were evoked using stimuli where segregation depended more on contrast differences than on the long-range integration of information (Bullier et al. 2001). Others also found figure-ground effects that were faster than those discussed here (Sugihara et al. 2011). There is thus a whole range of contextual effects, some of which are faster than others, and some of which may depend on feedback while others do not.

There is one counterintuitive aspect of interpreting these results in this way: in fact, feedback connections are not slow, but just as fast as feedforward connections, where both are at about 3.5 m/s (Girard et al. 2001). Horizontal connections that run via unmyelinated fibres in layers 2 and 3 of the cortex are about 10 times slower (Sugihara et al. 2011). Many of the Gestalt principles of perceptual organization

---

[26] This paradigm has been shown to distinguish between seen and not-seen stimuli in monkeys with a V1 lesion in one hemi-field, and was used to differentiate between "conscious" visual responses and unconscious blindsight behaviour: without catch trials (i.e., when in forced choice mode), monkeys react to both stimuli in the intact and in the lesioned field, expressing blindsight capabilities. In catch trials, however, monkeys only respond to stimuli in the intact and not in the lesioned hemi-field, as if expressing conscious sensation instead of a mere reflex (Moore et al. 1995). Supèr et al. used the same paradigm in intact monkeys to assess conscious percepts of figure–ground stimuli.

[27] A later investigation into neural activity preceding either seen or not-seen figure trials showed that not-seen trials are preceded by somewhat lower level of spontaneous activity, and also express less inter-neuronal synchrony (Supèr et al. 2003; Van der Togt et al. 2006).
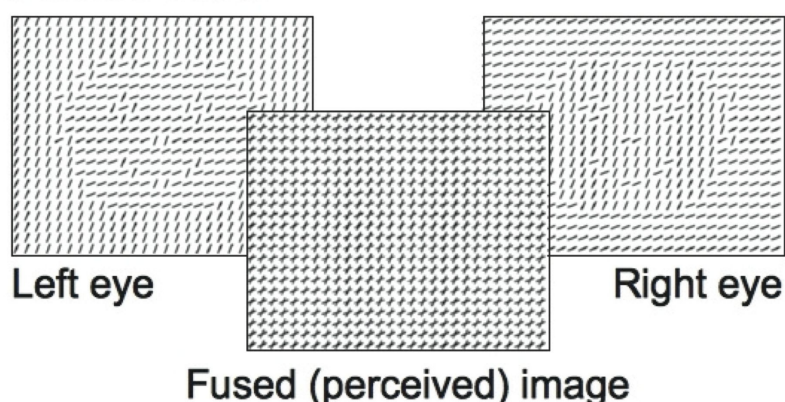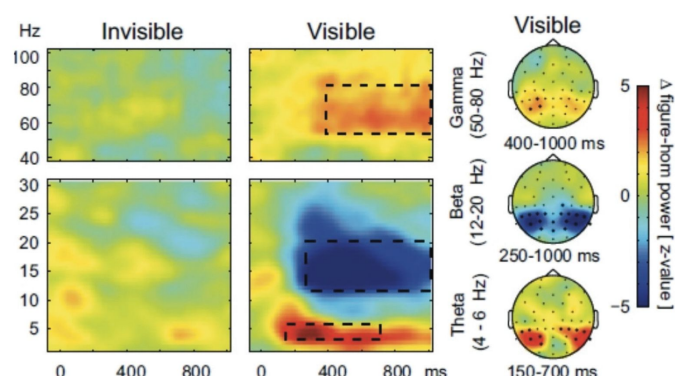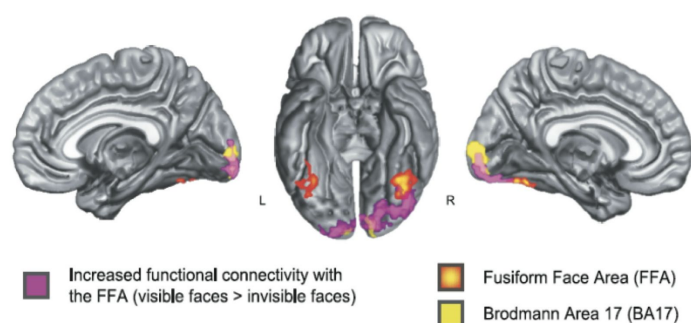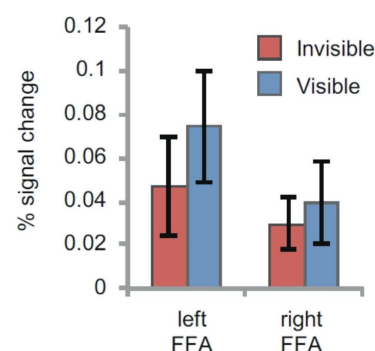

Figure 9: Top left: texture-defined faces were presented in either eye of subjects, yet with different orientations of line segments. As a result, the face was not visible in the fused percept (compare manipulation of figure 1). By using other orientation combinations, the same design could also result in a visible face (not shown). Top right: category-specific responses in the FFA did not differ for visible or invisible faces. Below: visible faces are characterized by strong recurrent interactions between FFA and earlier visual areas (left), and by strong synchronous activity in the theta, beta, and gamma bands (right). From: Fahrenfort et al. (2012).

are, however, embedded in these slow horizontal connections: V1 cells with a similar orientation preference are selectively interconnected via so-called patchy horizontal fibres. Moreover, these interconnections are strongest for oriented cells that have their receptive fields aligned along their orientation preference. Horizontal connections are also strongest between nearby cells (Gilbert & Wiesel 1989; Malach et al. 1993; Bosking et al. 1997). As such, these horizontal connections thus form the neural substrate of the well know Gestalt rules of "similarity", "co-linearity", and "proximity". A similar arrangement of preferred interconnectivity has been found for motion-direction selective cells in MT (Ahmed et al. 2012), potentially forming the substrate of the grouping principle of "common fate". Neurophysiological correlates of these

grouping principles are relatively fast, however (Knierim & Van Essen 1992; Kapadia et al. 1995).

The figure-ground segregation effects of figures 5, 6, 7, and 8 are among the longest latency contextual effects reported. That may be because they depend on both horizontal and feedback connections. Figure 10 shows the result of an experiment where the complete peristriate belt of visual cortex surrounding V1 and V2 was subjected to suction lesioning, removing (parts of) areas V3, V3A, V4, V4t, MT, MST, FST, PM, DP, and 7a (Lamme et al. 1998b; Supèr & Lamme 2007). Before the lesion, an oriented texture figure–ground stimulus evoked elevated activity in all neurons responding to the figure elements. Response modulation was even somewhat stronger, and occurred earlier at
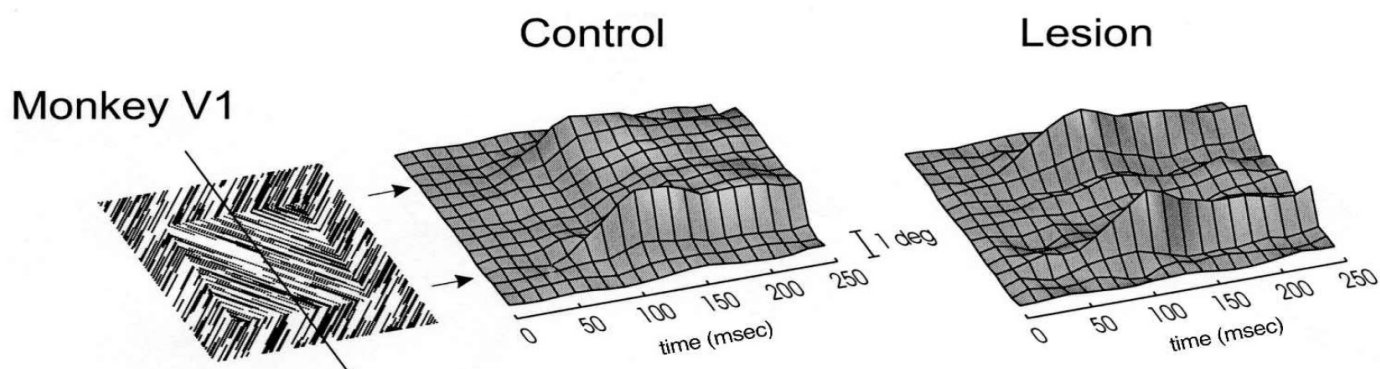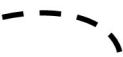
Figure 10: Contextual modulation (i.e., figure–ground responses, see figure 5) for various positions of the receptive field of V1 neurons (vertical axis), and extending over time (horizontal axis). In an intact monkey, modulation arises first at the figure-ground boundary, followed by a "filling-in" of the boundaries. After a lesion to the peri-striate belt of the visual cortex, only the boundary modulation remains, while filling-in has been abolished (Lamme et al. 1998a).

the boundary between figure and background. This was followed by a sort of "filling in" of enhanced activity between the boundary regions. We thus see an incremental process, starting with boundary segmentation and followed by surface segmentation. Similar findings have been reported in humans using combined EEG and TMS (Wokke et al. 2012).

After the lesion, the boundary enhancement remained, which may indicate that texture boundary detection mechanisms do not depend on feedback from higher visual areas and hence are mediated by horizontal connections within V1, or by recurrent interaction with V2. The centre modulation, where the figure elements are "neurally elevated" from the background elements, was completely abolished after the lesion, indicating that these figure-ground signals do depend on recurrent interactions between V1 and higher-tier areas. This finding was modelled on a realistic neural network of spiking neurons, indeed formalizing the idea that local orientation contrast—and hence the boundary between figure and ground—is mediated by inhibitory horizontal interactions between oriented receptive fields, whereas the figure–ground signal depended on excitatory feedback interactions trickling down from higher to lower areas (Roelfsema et al. 2002). Recently, laminar recording of figure-ground signals in V1 confirmed this idea (Self et al. 2013). These results show

that the long-latency figure-ground segregation effects depend on incremental interactions mediated by both horizontal and feedback connections. That may be the reason why they are most vulnerable to anaesthesia, masking, and other manipulations of consciousness.

Tononi modelled several neural architectures in order to find the connection parameters that fulfil the requirements for achieving maximally-integrated information. The optimal architecture consists of neurons that each have specific and different connections patterns, yet are sufficiently interconnected for each neuron to be able to connect to another via a few steps. Uniformly, or strictly modularly organized networks are less optimal. The thalamo-cortical system fits these requirements very well. On the one hand, neurons should be interconnected, otherwise information is not integrated. On the other hand, too much interconnection leads to a loss of specific information, as all neurons start doing the same thing, which happens in epilepsy or deep sleep—states that are indeed accompanied by a loss of consciousness (Tononi 2004, 2008, 2012). The contextual modulations that have been explored here seem to exactly express these properties: on the one hand, the neural responses are very specific, in that the major part of the response is driven by the features that are within the (small) receptive field. But on the other hand, the integration of these

VISUAL FUNCTION	Example	Conscious Vision		Anaesthesia	Hemianopia Blindsight	Backward Masking	Dichoptic Masking	Continuous Flash Suppression
Categorization Feature detection								
Higher level Categorization	word							
Interference						Brightness Colour		Brightness Colour
Inference					With intact hemifield			Breakthrough Discrimination
Base Grouping							Green	
Incremental Gestalt Grouping				Short range Long range		Absent		
Figure-Ground Organization				Absent		Absent	Absent	

■ Function is Present	■ Conflicting Results	■ Function is Absent	■ Unknown

Figure 11: Table summarizing the influence of consciousness manipulation on various visual functions. Colours indicate whether functions (rows) still operate under aa particular manipulation (columns). In the case of conflicting or uncertain evidence (yellow), the cases or conditions where the function still seems present is written in green; the cases where the function is absent are written in red. All functions are assumed to be present in conscious vision. For each visual function, an icon depicts its most prominent example. See text for explanation.

features rides on top of that response as a moderate modulation, expressing perceptual integration that may cover a large spatial extent, yet never even beginning to fully override the information carried by the neuron. In other words, visual neurons have categorization as their main priority, yet they also integrate these categories at some point in their response. That is the moment in time where the seed for conscious perception is laid (Lamme 2003, 2006, 2010a, 2010b).

8 Is there a functional boundary between unconscious and conscious vision?

I have taken the two extreme ends of consciousness manipulation: clear-cut visible and above-threshold items in awake subjects veridically reporting their visual experiences versus visual processing in anaesthesia, blindsight, or during profound masking or suppression (figure 11). If we don't accept conscious vision in the former, and the absence of it in the latter, there is no use arguing about the phenomenon of con-

sciousness. Even so, it has been surprisingly hard to find fundamental differences in the workings of many visual functions in the two conditions. Categorization of visual stimuli, even up to high levels, clearly stands independent of conscious visibility. It is unclear whether interference—i.e., the fact that features are no longer treated independently—depends on consciousness: shifts in brightness perception do not depend on consciousness, while it is uncertain whether the transition from wavelength to colour perception (and colour constancy) marks the conscious-unconscious divide.[28] Similarly, the status of inference phenomena like those observed in the Kanizsa triangle is uncertain.

This is all the more surprising given that many of these functions have traditionally been viewed as expressing the transition from merely physical features detected by sensor arrays towards the perceptual interpretation of this in-

---

28 Yet this transition may be the "Holy Grail" for those willing to understand qualia—or at least for those believing in "soft qualia", i.e., phenomenal properties that are not entirely detached from visual functioning, and having some sort of neural substrate (Block 1996, 2005, 2007).

formation. Moreover, they mark the integration of stimulus-driven input with our knowledge of the world, such that we arrive at visual "meaning". Recently, there has been quite some interest in so called predictive coding frameworks of vision (Rao & Ballard 1999; Panichello et al. 2012). In these frameworks, vision is seen as a type of Bayesian inference, where our prediction (prior) of the outside world is continuously matched with our sensory input, and where the difference is propagated through the network as an "error signal", which then results in an updating of the model (posterior). Indeed, expectations bias our perception of the world, most strongly in the face of ambiguous stimuli, but also in the case of unambiguous stimuli (Panichello et al. 2012). Although it has been suggested that either the matching process, the prior, or the posterior in this type of inference have some relation to consciousness,[29] this is questionable given the automaticity of many expectation effects. For example, the mere statistical dominance of a particular stimulus type is sufficient to bias perceptual interpretations (Chopin & Mamassian 2012). Also, expected words break from continuous flash suppression sooner than unexpected words (Costello et al. 2009).[30]

All in all, the relation between consciousness on the one hand and categorization, interference, and inference processes on the other hand ranges from non-existent to weak. A much stronger case seems possible for functions like the grouping of image elements according to Gestalt laws and figure-ground segregation. These operations seem to depend strongly on the conscious state, and on conscious perception of the stimuli involved (figure 11). This is surprising, given their relative "simplicity". For example, the grouping of similarly-oriented or collinear line segments may be achieved by horizontal connections in the primary visual cortex (Bosking et al. 1997, see above). Figure–ground

segregation—and its neurophysiological correlate—has been successfully modelled in a recurrent network architecture consisting of orientation-selective visual neurons in three hierarchically-organized visual areas, combined with some inhibitory horizontal interactions and excitatory feedback (Roelfsema et al. 2002). Regardless, the experimental data clearly show that if we want to identify visual functions that mark the transition from unconscious processing to conscious vision, grouping according to Gestalt laws (incremental grouping) and figure-ground segregation[31] (or perceptual organization in general) are our best bets.[32]

9 Is it all about distance, or time?

So why do Gestalt grouping and segregation bear such a close relation to consciousness? From a neural perspective, they differ from most other functions in that they depend on interactions between neurons at rather large distances. For example, for a neuron to "know" whether it sits on the figure or the background of the stimulus in figure 5, information has to travel over a distance of about 20 millimetres in the visual cortex.[33] Moreover, the modulations of neural activity that accompany this "knowing" depend on the incremental push-pull interactions between horizontal and feedback connections (Lamme & Roelfsema 2000; Roelfsema et al. 2002; Roelfsema 2006). These require quite extensive processing steps, given that the con-

---

[29] In my reading, the predictive coding models are sometimes rather vague about exactly which signal mediates conscious experience. It is often seen to be a combination of the matching process and the posterior, e.g., Seth et al. (2011).

[30] But note that this is in fact nothing more than a semantic priming experiment. The results primarily show that if a semantic category has been activated, this category will then break earlier from CFS.

[31] Of course one could argue that in the case of a face on a blank background there also is figure–ground segregation. This type of segregation clearly does not depend on consciousness. This touches on the debate on whether categorization is possible without segregation (Wagemans et al. 2012).

[32] A promising theory of consciousness holds that conscious representations and states are characterized by the integration of information, or more precisely, on the formation of complexes of integrated information (Tononi 2004, 2008, 2012). That integrated information characterizes consciousness is, however, mainly derived from a set of axioms and introspective or intuitive thought experiments, most of which have already been discussed in the previous text or footnotes (Tononi 2012). What this review of experimental findings however shows is that the "integration of information" comes in many guises, not all of which are equally strongly related to consciousness. A somewhat more precise definition of "integrated information" may be guided by these experimental findings.

[33] The figure is 4 degrees of visual angle wide. Neurons in human V1 with receptive fields at that distance are about 20 mm apart, given a cortical magnification factor of 0.2 degrees per millimetre at 2.0 degrees eccentricity (Duncan & Boynton 2003).

textual Gestalt effects typically manifest themselves at long latency.

Intuitively, seeing an illusion like the Kanizsa triangle, or the contextual shifts in brightness or colour perception discussed above, also seems to depend on "long range" interactions: information travels over large distances in the visual field. But distance travelled over the visual field does not always equal distance travelled in the brain. These phenomena may depend on fairly hardwired and feedforward mechanisms, and their neural correlates typically have relatively short temporal latencies (Von der Heydt et al. 1984). Seemingly, these phenomena tap into mechanisms that have high ecological relevance to the visual system, and are hence solved in a few processing steps, using dedicated feedforward mechanisms. The same holds for all categorization responses in the brain, regardless of their apparent complexity: the progression from low-level to high-level feature detection (including categorization of faces or other complex stimuli) proceeds in a feedforward "sweep" that lasts 100 ms or less (Lamme & Roelfsema 2000).

What emerges is the nagging feeling that consciousness has nothing to do with the seeming complexity or "high-levelness" of a visual function. Whether a visual function depends on consciousness may simply be related to the amount of space that has to be travelled in the brain, how many processing steps have to be taken in between, and hence how much time it takes to complete. This converges onto a thesis that we may call:

The STERP-property of phenomenal representations $=_{\text{Df}}$ conscious representations depend on the spatio-temporally extended neural processing mediated by recurrent interactions.

What that extent is remains to be specified, but has been studied directly by Faivre & Koch (2014), who measured the effects of stimuli made invisible using CFS on the perception of subsequent visible stimuli. Both for apparent motion and for biological motion walkers, it was found that unconscious motion integration only

occurred for relatively short (100 ms) and not for longer (400, 800, 1200ms) temporal intervals. Meng et al. (2007) observed that neural signals representing the spatial filling-in of a grating over a gap in the visual field depended on conscious experience of the grating.[34] This suggests that for visual information to literally "bridge a distance" across the visual field, consciousness is required.

The importance of the spatial and temporal extent of neural processing in consciousness also emerges from an entirely different field: that of disorders of consciousness. It is generally believed that there is a gradual decrease of consciousness from the healthy awake state towards minimally conscious, vegetative state and coma. These states also show a gradual decrease in the extent of neural interactions, in both space (Casali et al. 2013) and time (Bekinschtein et al. 2009). Particularly striking is the finding that the presence or absence of consciousness (in this case: the difference between minimally conscious and vegetative state patients) could be classified by simply looking at the amount of "shared symbolic information" in the EEG[35] at various distances in the brain. Shared symbolic information at distances of 10 cm and beyond signalled the presence of consciousness, and moreover was indicative of the prognosis of vegetative state patients (whether they would eventually awaken or not). Strikingly, this measure hardly depended on the location of the interactions (King et al. 2013). In other words, whenever and wherever neurons share information at distances of 10 cm or more, there is consciousness.[36]

Both distance and time are continuous. Arguing that consciousness is related to the temporal or spatial extent of neural processing therefore almost automatically seems to imply that the transition from unconscious to con-

---

34 It did not depend on attending the grating, however, which is of relevance to the discussion on the relation between attention and consciousness. See below (Lamme 2003, 2004, 2006, 2010a, 2010b).

35 At each electrode, EEG signals were first transformed into symbolic shapes (e.g., up-down-up) for various temporal intervals. Then it was determined to what extent these EEG "symbols" covaried between electrode pairs of various distances, after the exclusion of covariance that was caused by simple volume conduction.

36 Which made me wonder whether any piece of cortex of 10 cm or larger that is held on life support in a petri-dish might have consciousness.

scious processing is gradual rather than discrete. This is not necessarily so, however. Recurrent processing is mediated by highly non-linear interactions, and in such interactions, rather discrete phase transitions are possible (Steyn-Ross et al. 1999; Del Cul et al. 2007; Hwang et al. 2012). It could thus very well be that there is a discrete transition from a phase where information integration is rather limited to a phase that is characterized by extensive information integration, and that this transition depends on the temporal or spatial extent of recurrent interactions.[37]

Whether the transition from unconscious to conscious processing is discrete or continuous has been argued on different grounds, such as on the distribution of behavioral responses ("seen" versus "not seen") in relation to manipulations of stimulus variables (Sergent & Dehaene 2004; Overgaard et al. 2006). In signal detection theory, the strength of perceptual information is considered to be continuous, while the decision criterion imposes a discrete boundary between what is reported as "seen" or "not seen". In its classic form, however, signal detection theory is agnostic about whether consciousness is pre- or post-decisional. Recently, many attempts have been made to incorporate consciousness into the framework of signal detection theory, and in many of these models consciousness is considered post- rather than pre-decisional (Maniscalco & Lau 2012; King & Dehaene 2014)—thus the boundary between the conscious and unconscious is taken to be discrete. Based on neurophysiological findings in the monkey visual cortex, a signal-detection model was devised in which consciousness was considered pre-decisional. In this model, the distribution of sensory information was considered bi-modal, reflecting either a conscious or an unconscious state. The model could explain both the behavioral and neurophysiological findings in the monkey visual cortex, obtained using a variety of stimulus strengths and decision cri-

teria (Supèr et al. 2001). Note that also in this pre-decisional model the conscious–unconscious divide is discrete (or at least bi-modal), rather than gradual.

10 The function of conscious vision

Could it be that Gestalt grouping and figure-ground segregation (of textured images) only happen to go along with consciousness because they take more time; because they require more elaborate computations, not provided by the many dedicated feedforward pathways and modules of the brain? Normally, vision proceeds in a fast and feedforward fashion, where dedicated neurons detect features and categories. Using its hardwired connections, the visual system can swiftly detect the most relevant objects: food, mates, or dangerous animals. Some objects are more difficult to discern, and require prior knowledge or the computation of neighbourhood relations between image elements: food behind a leaf, a sweet versus a sour apple. That takes slightly more—but not too much more—time, because many of the required interactions are hardwired as well. They are hardwired because the visual system has been exposed to these "visual problems" very often, either during evolution or during visual experience. Then there are visual problems that are even more difficult: a camouflaged animal in a crowded forest (figure 7), only visible via subtle differences in overall texture or motion. In this case, all visual resources and mechanisms have to come to the rescue. Only by combining the input from many neurons in a versatile way can the visual "solution" be found. That may be the function of consciousness in the visual domain: to combine the otherwise unconscious modules and mechanisms in a flexible way so as to solve otherwise unresolvable visual problems leading to a second thesis that we may call:

> The SUPER-property of phenomenal representations $=_{\text{Df}}$ neural representations require consciousness and invoke phenomenality as soon as what needs to be represented can no longer be represented by a single dedicated module or mechanism, yet

---

[37] It could even be that the mere fact that information exchange extends over a particular time and space is critical for that exchange to be accompanied by a conscious sensation. When the same amount of information would be exchanged much slower — as in plants — or much faster — as in a supercomputer — or over a smaller or larger space (as in a microchip or over the internet) no conscious sensation ensues.

requires the interaction of these modules so that a super-positioned representation emerges.

From the point of view of consciousness, a hierarchy of visual functions can then be made. This starts with largely unconscious feature detection and object categorization. These features start to influence each other, and are no longer treated independently, so that categories form that are about the relations between image items (base groupings, short range incremental grouping). With this, there is a transition from the physical properties of the visual input as they are presented to the sensor array to the meaning[38] of these properties (e.g., wavelength to colour). During these operations, features and categories are matched with our knowledge and expectations of the world, embedded in the anatomical organization of the visual cortex, aiding in the transformation from visual input towards meaning (inference). Finally, all this information is combined into an organized percept. The longer these operations take, the more distance has to be travelled in the brain, and the more conscious these operations become.[39]

If nothing interferes, the visual system will always strive towards optimally integrating the available information, so that the richest interpretation of all available information is achieved, and all features have been detected, all inferences have been made, all image elements are combined and all potential ambiguities have been resolved. If this process is cut short, for example by masking or a TMS pulse (Pascual-Leone & Walsh 2001; Silvanto et al. 2005), there is no integrated end-result. And seemingly there is no conscious sensation either. Regardless of this, many features have still been detected, many inferences have been made, and

the brain can use this information to achieve its goals. Behaviour may be influenced, or set into motion (Dehaene et al. 1998). Priming will occur, as well as all sorts of unconscious cognition (Van Gaal & Lamme 2012). Without consciousness, and without maximal integration, the visual system is far from helpless. It can do less, but it can still do a lot.

From this perspective, the function of consciousness in vision is just to enable that last push. That is, to resolve the visual issues that cannot be dealt with otherwise.[40] And with that, visual functions grow more complex, and evolve from their basic form into more sophisticated versions. A good example comes—once again—from the processing of faces. The core property of face-selective neurons is to respond in a category-selective manner: they distinguish between faces and other objects. They do so from the very first action-potentials that are fired. At that moment, however, category specificity is still very basic, in the sense that all types of faces evoke a similar response (Rolls 1992). At a later moment in time, however, responses typically become more and more specific. In the monkey visual cortex, face cells distinguish between different viewpoints and different emotional expressions of faces with a delay of about 50 milliseconds relative to the categorical face/non-face response (Sugase et al. 1999). View invariant identity representations arise even later, with a delay of about 200 ms (Freiwald & Tsao 2010). At these delays, the face-selective neurons will have established recurrent interactions with lower (and higher) level neurons across the brain, allowing for these more sophisticated classifications to be expressed in the response.

We may thus conclude that face recognition "as we know it"—i.e., not just categorizing face versus non-face, but seeing that face, knowing what it looks like, who it is, and what emo-

---

38 Note that "meaning" in this context refers to the meaning information has to the organism, shaped by and in accordance with its evolutionary history and ontogenesis (like colour has the "meaning" of the edibility of fruit). It does not refer to "meaning" in any linguistic sense.

39 That may explain why two seemingly similar phenomena like the brightness and colour shifts of figure 2, and the arrival at colour constancy in figure 3 are depend on consciousness in different ways. Colour constancy requires the computation of the full distribution of wavelengths over the entire image, which takes more time than the computations required to compute brightness of adjacent patches.

40 Maybe that is the reason why the transition from unconscious to conscious processing also marks the transition between veridical and inferred representations (e.g., from wavelength to colour). Dedicated modules can do their thing in isolation, and therefore have no need to compromise towards a non-veridical representation of the outside world. When modules interact, the necessity may arise to compromise veridical representations to achieve global coherence into the combined super-positioned representation that cannot be represented otherwise.

---

tion it carries—is a visual function tightly linked to conscious rather than unconscious vision. The main reason for this lies in the fact that in conscious recognition we go beyond simple categorization, and move towards a function where the integration of all possible information about that face (its viewpoint, colour, identity, emotional expression, etc.) is required.

This may raise the question of how we then become conscious of an extremely simple stimulus, such as an oriented black line on a completely white background. With such a simple stimulus, there seems to be no need for any elaborate binding, incremental grouping, or inference. Neurons in the primary visual cortex can detect the line and its orientation within a few action potentials. There seems to be no need to call in the functions that are enabled by conscious processing. So why is it, then, that we still see the black line on the white background?

First, it should be noted that the notion of "simple" stimuli is more complex than one would expect. For example, it was shown that subjects can rapidly detect animals or vehicles in complex natural scenes, even when their attention is simultaneously focused on another task. Discriminating large T's from L's, or bisected colour disks from their mirror images was impossible under the same dual task paradigm. Apparently, seemingly simple letter or disk stimuli require more attentive processing than seemingly complex natural scenes (Li et al. 2002), suggesting that they take longer and more elaborate processing. In blindsight, subjects can discriminate lines of different orientations, suggesting that conscious processing is not required for these simple stimuli. However, discrimination performance—although above chance—is typically worse than for consciously-seen line segments, suggesting that something is "missing" from the neural representations formed in blindsight compared to those in conscious vision.

So what might the more elaborate processing steps that lift the unconscious representation of a black line towards a conscious representation of that line be? First, it is known that neurons in many visual areas beyond V1 respond to orientated line segments. At each level,

receptive fields, and hence spatial frequency preferences, differ. This means that (the orientation of) the line segment is represented at many different spatial scales across the visual cortex. Only the integration of these differently-scaled representations, via recurrent interactions, yields a precise and conscious representation. The same holds for other properties of the "simple" line segment, such as its colour, its depth, and its relation to the background.[41] Indeed, oriented lines are fairly easy to mask (in fact easier than faces), indicating that their conscious percept depends on more elaborate processing steps than expected for such a simple stimulus.

11 The impact of conscious vision on the brain

If a particular visual problem has to be dealt with often, the brain will start to build connections so that the problem can be resolved more rapidly. Visual problems that require long and elaborate processing will eventually be resolved in milliseconds. By building new and dedicated connections, elaborate processing steps may be simplified into a fast and short set of interactions. Conscious processing will turn into unconscious processing, because conscious processing has triggered perceptual learning that in turn evokes synaptic changes that create new "dedicated modules" that can do the job unconsciously. This leads to a third thesis:[42]

> The LEARN-property of phenomenal representations $=_{Df}$ neural representations that require consciousness and invoke phenomenality, at the same time evoke synaptic plasticity mechanisms and learning,

---

[41] Even something as simple as a white background will give the black line another visual "meaning" than a yellow background, a green background, or a textured background. The same point has been formulated by Giulio Tononi (2004, 2008, 2012): a conscious representation is conscious because it differentiates from the endless other potential representations that could have been. In this case: the oriented black line on the white background is one of the endless possible configurations of lines on backgrounds, and only by integrating the information of line and background is it known which of these configurations is actually present.

[42] Similar ideas exist in the context of motor learning: a task that first requires extensive conscious practice will gradually become more and more automatic, up to the point where it can be executed fully unconsciously.

---

in an attempt to make these representations less dependent on consciousness and invoking less phenomenality.

Indeed, there are several arguments for linking consciousness to perceptual learning. Plasticity in the visual cortex comes in many temporal and spatial scales. There are fast- and short-range adaptations or recalibrations, expressed in altered stimulus-response dependencies (e.g., contrast normalization). But receptive fields may also change in size or feature selectivity when exposed to repeated stimulation. Receptive fields literally grow or shift position when their surrounds are stimulated but the receptive field is not (Gilbert & Wiesel 1992). Prolonged depletion of input leads to the induction of new connectivity via fast axonal sprouting of horizontal connections (Yamahachi et al. 2009). Horizontal connections in particular play an important role in both immediate and longer term plasticity of the visual cortex (Gilbert et al. 1996). The repeated execution of Gestalt grouping via the same connections may therefore induce learning (Gilbert et al. 2001), as, for example, is observed in the learning of texture segregation (Karni & Sagi 1991) or in the gradual improvement of contour integration during childhood development (Kovács et al. 1999). In addition, perceptual learning induces a reorganization of the areas involved in encoding the learned object—a process that is mediated by feedback connections (Sigman & Gilbert 2000; Sigman et al. 2005). It seems that the neural machinery that mediates Gestalt grouping and segregation is also the machinery that mediates perceptual learning.

Furthermore, feedback and horizontal connections have been linked to the molecular mechanisms of neural plasticity. A key component in neural plasticity is the NMDA receptor pathway, and in the monkey, NMDA receptor blocking using APV reduces contextual figure–ground modulation (Self et al. 2012). Similarly, in humans, figure–ground segregation is impaired using Ketamine, an anaesthetic which selectively blocks the NMDA receptor at low doses (Meuwese et al. 2013). Also, it was found that Ketamine at sub-anaesthetic doses interferes with the leaning of Mooney figures. Mooney figures are high-contrast versions of images that are hard to recognize when you don't know what the image is about. Once you have seen its original natural contrast version, however, the Mooney image is readily recognizable. It was found that the neural representation of Mooney images starts to resemble that of their natural versions once they are learned. Ketamine disrupts this rapid learning process, but only in V1, and not in higher visual areas, indicating that feedback from higher areas to V1 is selectively disrupted by Ketamine (Van Loon et al. submitted).

In sum, there are strong indications that link conscious visual processing and its neural machinery—horizontal and feedback connection—are linked to perceptual learning and the molecular mechanisms involved. This may open up a path to a more molecular understanding of consciousness. In addition, it provides us with a clear idea about the function of consciousness: that of building a new repertoire of visual functions, so that eventually conscious processing is no longer necessary.

It must be noted however, that the link between consciousness and learning is controversial. Many instances of "unconscious" perceptual learning exist (e.g., Gutnisky et al. 2009; Seitz et al. 2009; Seitz & Watanabe 2003; Schwiedrzik et al. 2011). An important issue here, however, is whether these are cases of learning without conscious experience of the stimuli that induce the learning, or whether they are instances of learning without cognitive access or attention to these stimuli (see Meuwese et al. 2013). A further clarification of the role of consciousness in learning is required.

12 The dolphins of consciousness research

I have examined the defining characteristics of conscious versus unconscious vision. Incremental grouping and segregation according to Gestalt laws seems to be a defining characteristic of conscious vision. Other visual phenomena and functions, like interference or inference, are less

strongly linked. Feature detection and higher-level categorization clearly do not mark the transition from unconscious to conscious vision. From a neural perspective, it can be argued that conscious processing is linked to those operations that require spatially and temporally extended processing, where neurons engage in incremental interactions involving many steps. These processes are selectively dependent on horizontal and feedback connections. Moreover, these interactions induce learning, as they operate along highly plastic neural pathways, and use the molecular machinery that is directly involved in neural plasticity.

We can now start using these defining characteristics to answer more difficult questions. Is there consciousness in the right half-brain of a split brain patient (Sperry 1984)? Is there consciousness without attention (Koch & Tsuchiya 2012)? Is there consciousness in neglect or extinction (Lamme 2003)? Is it appropriate to talk about inattentional "blindness", where people do not remember having seen something while their attention was engaged elsewhere? What exactly happens during change blindness (Simons & Rensink 2005)? Is there consciousness in animals (Edelman & Seth 2009), or in a vegetative state (Owen et al. 2006)? These are the "dolphins" of consciousness research, situations that are hard to position in the current taxonomy of conscious versus unconscious, because much controversy exists about the presence or absence of conscious experience in those conditions. With this, I hope to have given some usable arguments that can settle such controversies. My claim would simply be that whenever we see the defining properties of conscious vision that have been laid out here (i.e., incremental Gestalt grouping and segregation), there is conscious vision, regardless of whether there is conscious access or report (e.g., Scholte et al. 2006). More in general, the more fruitful stance towards consciousness would be to let all the available evidence converge into general theses, such as those derived here, and then take these as the defining characteristics of conscious processing and consciousness, regardless of whether they fit our introspective intuition of what consciousness is or should be. Defining consciousness as the process that builds on spatio-temporally extended neural processing (STERP property), that enables the building of super-positioned representations that individual modules cannot provide (SUPER property), and that evokes synaptic plasticity and learning (LEARN property) yields clear defining characteristics. These characteristics go a great length towards elucidating important features of phenomenality (its integrated nature, Gestalt properties), towards explaining the nature of conscious experience (perceptual organization, interference, inference), and are hinting towards a potential function of consciousness (learning) and its molecular basis. What I consider irrelevant characteristics (such the ability to report about an experience, see Lamme 2010a, 2010b) generally do no such explaining. It is better to build a taxonomy of conscious versus unconscious processing on defining characteristics than on irrelevant ones. That has helped a lot in positioning dolphins in the taxonomy of species. It will also help a lot in positioning the wild amalgam of phenomena that the field of consciousness research has produced so far. And it will enable us to give consciousness its proper ontological status. But I have already contributed to that discussion extensively elsewhere (Lamme 2003, 2004, 2006, 2010a, 2010b), so I will lay that to rest here.

At the crack of dawn, something magical happens. Night turns into day, life springs, vibrations fill the air. We know, it is just the earth rotating. But a very fundamental transition it remains. Unconscious or conscious processing, it's all neurons doing their job, firing action potentials, exchanging chemicals, transferring information. But somehow, suddenly, they "turn on the light". You see. You have a conscious sensation of that dawn. Isn't it beautiful? You should take a picture of it.

Acknowledgements

References

Ahmed, B., Cordery, P. M., McLelland, D., Bair, W. & Krug, K. (2012). Long-range clustered connections within extrastriate visual area V5/MT of the rhesus macaque. *Cerebral Cortex*, *22* (1), 60-73. 10.1093/cercor/bhr072

Albert, M. K. & Hoffman, D. D. (2000). The generic-viewpoint assumption and illusory contours. *Perception*, *29* (3), 303-312.

Alkire, M. T., Hudetz, A. G. & Tononi, G. (2008). Consciousness and anesthesia. *Science*, *322* (5903), 876-80. 10.1126/science.1149213

Allman, J., Miezin, F. & McGuinness, E. (1985). Stimulus specific responses from beyond the classical receptive field: Neurophysiological mechanisms for local-global comparisons in visual neurons. *Annual Review of Neuroscience*, *8*, 407-30. 10.1146/annurev.ne.08.030185.002203

Almeida, J., Mahon, B. Z. Nakayama, K. & Caramazza, A. (2008). Unconscious processing dissociates along categorical lines. *Proceedings of the National Academy of Sciences of the United States of America*, *105* (39), 15214-15218. 10.1073/pnas.0805867105

Barbur, J. L., de Cunha, D., Williams, C. B. & Plant, G. (2004). Study of instantaneous color constancy mechanisms in human vision. *Journal of Electronic Imaging*, *13* (1), 15-28. 10.1117/1.1636491

Barbur, J. L. & Spang, K. (2008). Colour constancy and conscious perception of changes of illuminant. *Neuropsychologia*, *46* (3), 853-863. 10.1016/j.neuropsychologia.2007.11.032

Bekinschtein, T. A., Dehaene, S., Rohaut, B., Tadel, F., Cohen, L. & Naccache, L. (2009). Neural signature of the conscious processing of auditory regularities. *Proceedings of the National Academy of Sciences*, *106* (5), 1672-1677. 10.1073/pnas.0809667106

Block, N. (1996). How can we find the neural correlate of consciousness? *Trends in Neurosciences*, *19* (11), 456-459. 10.1016/S0166-2236(96)20049-9

—— (2005). Two neural correlates of consciousness. *Trends in Cognitive Sciences*, *9* (2), 46-52. 10.1016/j.tics.2004.12.006

—— (2007). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences*, *30* (5-6), 481-499. 10.1017/S0140525X07002786

Bosking, W. H., Zhang, Y., Schofield, B. & Fitzpatrick, D. (1997). Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *The Journal of Neuroscience*, *17* (6), 2112-2127.

Boyer, J. L., Harrison, S. & Ro, T. (2005). Unconscious processing of orientation and color without primary visual cortex. *Proceedings of the National Academy of Sciences*, *102* (46), 16875-16879. 10.1073/pnas.0505332102

Breitmeyer, B. G., Ogmen, H.G., Ro, T. & Singhal, N. S. (2004). Unconscious color priming occurs at stimulus- not percept-dependent levels of processing. *Psychological Science*, *15* (3), 198-202. 10.1111/j.0956-7976.2004.01503009.x

Breitmeyer, B. G., Ogmen, H. G., Ro, T., Ogmen, H. & Todd, S (2007). Unconscious, stimulus-dependent priming and conscious, percept-dependent priming with chromatic stimuli. *Perception & Psychophysics*, *69* (4), 550-557. 10.3758/BF03193912

Breitmeyer, B. G. & Ogmen, H. (2000). Recent models and findings in visual backward masking: A comparison, review, and update. *Perception & Psychophysics*, *62* (8), 1572-1595. 10.3758/BF03212157

Bullier, J., Hupé, J. M., James, A. C. & Girard, P. (2001). The role of feedback connections in shaping the responses of visual cortical neurons. *Progress in Brain Research*, *134*, 193-204.

Casali, A. G., Gosseries, O., Rosanova, M., Boly, M., Sarasso, S., Casali, K. R., Casarotto, S., Bruno, M. A., Laureys, S., Tononi, G. & Massimini, M. (2013). A theoretically based index of consciousness independent of sensory processing and behavior. *Science Translational Medicine*, *5* (198), 198ra105. 10.1126/scitranslmed.3006294

Chalmers D. J. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies*, *2* (3), 200-219.

Chopin, A. & Mamassian, P. (2012). Predictive properties of visual adaptation. *Current Biology*, *22* (7), 622-626. 10.1016/j.cub.2012.02.021

Churchland, P. M. (1981). Elimative materialism and the propositional attitudes. *Journal of Philosophy*, *78* (2), 67-90. 10.2307/2025900

Cohen, M. A. & Dennett, D. C. (2011). Consciousness cannot be separated from function. *Trends in Cognitive Sciences*, *15* (8), 358-364. 10.1016/j.tics.2011.06.008

Costello, P., Jiang, Y., Baartman, B., McGlennen, K. & He, S. (2009). Semantic and subword priming during binocular suppression. *Consciousness and Cognition*, *18* (2), 375-382. 10.1016/j.concog.2009.02.003

Cowey, A. & Heywood, C. A. (1997). Cerebral achromatopsia: Colour blindness despite wavelength processing. *Trends in Cognitive Sciences*, *1* (4), 133-139. 10.1016/S1364-6613(97)01043-7

Cox, M. A., Schmid, M. C., Peters, A. J., Saunders, R. C., Leopold, D. A. & Maier, A. (2013). Receptive field focus of visual area V4 neurons determines responses to illusory surfaces. *Proceedings of the National Academy of Sciences*, *110* (42), 17095-17100. 10.1073/pnas.1310806110

Crick, F. & Koch, C. (1998). Consciousness and neuroscience. *Cerebral Cortex*, *8* (2), 97-107. http://www.ncbi.nlm.nih.gov/pubmed/9542889

——— (2003). A framework for consciousness. *Nature Neuroscience*, *6* (2), 119-126. 10.1038/nn0203-119

Cumming, B. G. & Parker, A. J. (1997). Responses of primary visual cortical neurons to binocular disparity without depth perception. *Nature*, *389* (6648), 280-283. 10.1038/38487

Dehaene, S., Naccache, L., Le Clec'H, G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., Van de Moortele, P. F. & Le Bihan, D. (1998). Imaging unconscious semantic priming. *Nature*, *395* (6702), 597-600. 10.1038/26967. http://www.ncbi.nlm.nih.gov/pubmed/9783584

Dehaene, S., Jobert, A., Naccache, L., Ciuciu, P., Poline, J. B., Le Bihan, D. & Cohen, L. (2004). Letter binding and invariant recognition of masked words: Behavioral and neuroimaging evidence. *Psychological Science*, *15* (5), 307-313. 10.1111/j.0956-7976.2004.00674.x

Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J. & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: A testable taxonomy. *Trends in Cognitive Sciences*, *10* (5), 204-211. 10.1016/j.tics.2006.03.007

Dehaene, S. & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition*, *79* (1-2), 1-37. http://www.ncbi.nlm.nih.gov/pubmed/11164022

Del Cul, A., Baillet, S. & Dehaene, S. (2007). Brain dynamics underlying the nonlinear threshold for access to consciousness. *PLoS Biology*, *5* (10), e260. 10.1371/journal.pbio.0050260

Dennett, D. C. (1993). *Consciousness explained.* New York, NY: Penguin. (pp. 889-892). 10.2307/2108259

Dow, B. M., Snyder, A. Z., Vautin, R. G. & Bauer, R. (1981). Magnification factor and receptive field size in foveal striate cortex of the monkey. *Experimental Brain Research*, *44* (2), 213-228. http://www.ncbi.nlm.nih.gov/pubmed/7286109

Duncan, R. O. & Boynton, G. M. (2003). Cortical magnification within human primary visual cortex correlates with acuity thresholds. *Neuron*, *38* (4), 659-671. 10.1016/S0896-6273(03)00265-4

Edelman, D. B. & Seth, A. K. (2009). Animal consciousness: A synthetic approach. *Trends in Neurosciences*, *32* (9), 476-484. 10.1016/j.tins.2009.05.008

Enns, J. T. & Di Lollo V. (2000). What's new in visual masking? *Trends in Cognitive Sciences*, *4* (9), 345-352. 10.1016/S1364-6613(00)01520-5

Fahrenfort, J. J., Scholte, H. S. & Lamme, V. A. F. (2007). Masking disrupts reentrant processing in human visual cortex. *Journal of Cognitive Neuroscience*, *19* (9), 1488-1497. 10.1162/jocn.2007.19.9.1488

Fahrenfort, J. J., Snijders, T. M., Heinen, K., Van Gaal, S., Scholte, H. S. & Lamme, V. A. F. (2012). Neuronal integration in visual cortex elevates face category tuning to conscious face perception. *Proceedings of the National Academy of Sciences*, *109* (52), 21504-21509. 10.1073/pnas.1207414110

Fahrenfort, J. J. & Lamme, V. A. F. (2012). A true science of consciousness explains phenomenology: Comment on Cohen and Dennett. *Trends in Cognitive Sciences*, *16* (3), 138-140. 10.1016/j.tics.2012.01.004

Faivre, N. & Koch, C. (2014). Temporal structure coding with and without awareness. *Cognition*, *131* (3), 404-414. 10.1016/j.cognition.2014.02.008

Fang, F. & He, S. (2005). Cortical responses to invisible objects in the human dorsal and ventral pathways. *Nature Neuroscience*, *8* (10), 1380-1385. 10.1038/nn1537

Foster, K. H., Gaska, J. P., Nagler, M. & Pollen, D. A. (1985). Spatial and temporal frequency selectivity of neurones in visual cortical areas V1 and V2 of the macaque monkey. *Journal of Physiology*, *365*, 331-363.

Freiwald, W. A. & Tsao, D. Y. (2010). Functional compartmentalization and viewpoint generalization within the macaque face-processing system. *Science*, *330* (6005), 845-851. 10.1126/science.1194908

Gazzaniga, M. S. (2005). Forty-five years of split-brain research and still going strong. *Nature Reviews Neuroscience*, *6* (8), 653-659. 10.1038/nrn1723

Gilbert, C. D., Das, A., Ito, M., Kapadia, M. & Westheimer, G. (1996). Spatial integration and cortical dynamics. *Proceedings of the National Academy of Sciences*, *93* (2), 615-622.

Gilbert, C. D., Sigman, M. & Crist, R. E. (2001). The neural basis of perceptual learning. *Neuron*, *31* (5), 681-97. http://www.ncbi.nlm.nih.gov/pubmed/11567610

Gilbert, C. D. & Wiesel, T. N. (1989). Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. *The Journal of Neuroscience*, *9* (7), 2432-42. http://www.ncbi.nlm.nih.gov/pubmed/2746337

——— (1992). Receptive field dynamics in adult primary visual cortex. *Nature*, *356* (6365), 150-2. 10.1038/356150a0. http://www.ncbi.nlm.nih.gov/pubmed/1545866

Girard, P., Hupé J. M., James A. C. & Bullier, J. (2001). Feedforward and feedback connections between areas V1 and V2 of the monkey have similar rapid conduction velocities. *Journal of Neurophysiology*, *85* (3), 1328-1331. http://www.ncbi.nlm.nih.gov/pubmed/11248002

Gobbini, M. I., Gors, J. D., Halchenko, Y. O., Hughes, H. C. & Cipolli, C. (2013). Processing of invisible social cues. *Consciousness and Cognition*, *22* (3), 765-770. 10.1016/j.concog.2013.05.002

Goodale, M. A. & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, *15* (1), 20-25. 10.1016/0166-2236(92)90344-8

Gutnisky, D. A., Hansen, B. J., Iliescu, B. F. & Dragoi, V. (2009). Attention alters visual plasticity during exposure-based learning. *Current Biology*, *19* (7), 555-560. 10.1016/j.cub.2009.01.063

Harris, J. J., Schwarzkopf, D. S., Song, C., Bahrami, B. & Rees, G. (2011). Contextual illusions reveal the limit of unconscious visual vrocessing. *Psychological Science*, *22* (3), 399-405. 10.1177/0956797611399293

He, S., Cavanagh, P. & Intriligator, J (1996). Attentional resolution and the locus of visual awareness. *Nature*, *383* (6598), 334-337. 10.1038/383334a0

Hess, R. F., Beaudot, W. H. & Mullen, K. T. (2001). Dynamics of contour integration. *Vision Research*, *41* (8), 1023-1037. 10.1016/S0042-6989(01)00020-7

Hubel, D. H. (1982). Exploration of the primary visual cortex, 1955-78. *Nature*, *299* (5883), 515-524. http://www.ncbi.nlm.nih.gov/pubmed/6750409

Hubel, D. H. & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, *195* (1), 215-243.

Hupé, J. M., James, A. C., Payne, B. R., Lomber, S. G., Girard, P. & Bullier, J. (1998). Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. *Nature*, *394* (6695), 784-787. 10.1038/29537

Hupé, J. M., James, A. C., Girard, P. & Bullier, J. (2001). Response modulations by static texture surround in area V1 of the macaque monkey do not depend on feedback connections from V2. *Journal of Neurophysiology*, *85* (1), 146-63.

Hupé, J. M., James, A. C., Girard, P., Lomber, S. G., Payne, B. R. & Bullier, J. (2001). Feedback connections act on the early part of the responses in monkey visual cortex. *Journal of Neurophysiology*, *85* (1), 134-145.

Hwang, E., Kim, S., Han, K. & Choi, J. H. (2012). Characterization of phase transition in the thalamocortical system during anesthesia-induced loss of consciousness. *PLoS One*, *7* (12), e50580. 10.1371/journal.pone.0050580

Jiang, Y., Costello, P., Fang, F., Huang, M. & He, S. (2006). A gender- and sexual orientation-dependent spatial attentional effect of invisible images. *Proceedings of the National Academy of Sciences*, *103* (45), 17048-17052. 10.1073/pnas.0605678103

Kamermans, M., Kraaij, D. A. & Spekreijse, H. (1998). The cone/horizontal cell network: A possible site for color constancy. *Visual Neuroscience*, *15* (5), 787-797.

Kapadia, M. K., Ito, M., Gilbert, C. D. & Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: Parallel studies in human observers and in V1 of alert monkeys. *Neuron*, *15* (4), 843-856. 10.1016/0896-6273(95)90175-2

Karni, A. & Sagi, D. (1991). Where practice makes perfect in texture discrimination: Evidence for primary visual cortex plasticity. *Proceedings of the National Academy of Sciences of the USA*, *11* (88), 4966-4970. 10.1016/j.cub.2009.01.063

Kim, C. Y. & Blake, R. (2005). Psychophysical magic: Rendering the visible 'invisible'. *Trends in Cognitive Sciences*, *9* (8), 381-388. 10.1016/j.tics.2005.06.012

King, J. R. & Dehaene, S. (2014). A model of subjective report and objective discrimination as categorical decisions in a vast representational space. *Philosophical Transactions of the Royal Society B*, *369* (1641), 1471-2970. 10.1098/rstb.2013.0204

King, J. R., Sitt, J. D., Faugeras, F., Rohaut, B., El Karoui, I., Cohen, L., Naccache, L. & Dehaene, S. (2013). Information sharing in the brain indexes consciousness in noncommunicative patients. *Current Biology*, *23* (19), 1914-1919. 10.1016/j.cub.2013.07.075. http://www.ncbi.nlm.nih.gov/pubmed/24076243

Knierim, J. J. & Van Essen, D. C. (1992). Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *Journal of Neurophysiology*, *67* (4), 961-980.

Koch, C. & Tsuchiya, N. (2012). Attention and consciousness: Related yet different. *Trends in Cognitive Sciences*, *16* (2), 103-105. 10.1016/j.tics.2011.11.012

Kolb, F. C. & Braun, J. (1995). Blindsight in normal observers. *Nature*, *377* (6547), 336-338. 10.1038/377336a0

Kouider, S., Eger, E., Dolan, R. & Henson, R. N. (2009). Activity in face-responsive brain regions is modulated by invisible, attended faces: Evidence from masked priming. *Cerebral Cortex*, *19* (1), 13-23. 10.1093/cercor/bhn048

Kovács, I., Kozma, P., Fehér, A. & Benedek, G. (1999). Late maturation of visual spatial integration in humans. *Proceedings of the National Academy of Sciences*, *96* (21), 12204-12209. 10.1073/pnas.96.21.12204

Lamme, V. A. F. (1995). The neurophysiology of figure-ground segregation in primary visual cortex. *Journal of Neuroscience*, *15* (2), 1605-1615.

——— (2003). Why visual attention and awareness are different. *Trends in Cognitive Sciences*, *7* (1), 12-18. 10.1016/S1364-6613(02)00013-X

——— (2004). Separate neural definitions of visual consciousness and visual attention; a case for phenomenal awareness. *Neural Networks*, *17* (5-6), 861-872. 10.1016/j.neunet.2004.02.005

——— (2006). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences*, *10* (11), 494-501. 10.1016/j.tics.2006.09.001

——— (2010a). How neuroscience will change our view on consciousness. *Cognitive Neuroscience*, *1* (3), 204-220. 10.1080/17588921003731586

——— (2010b). What introspection has to offer, and where its limits lie. *Cognitive Neuroscience*, *1* (3), 232-240. 10.1080/17588928.2010.502224

Lamme, V. A. F., Van Dijk, B. W. & Spekreijse, H. (1993). Contour from motion processing occurs in primary visual cortex. *Nature*, *363* (6429), 541-543. 10.1038/363541a0

Lamme, V. A. F., Supèr, H. & Spekreijse, H. (1998a). Feedforward, horizontal, and feedback processing in the visual cortex. *Current Opinion in Neurobiology*, *8* (4), 529-35. http://www.ncbi.nlm.nih.gov/pubmed/9751656

Lamme, V. A. F., Zipser, K. & Spekreijse, H. (1998b). Figure-ground activity in primary visual cortex is suppressed by anesthesia. *Proceedings of the National Academy of Sciences of the United States of America*, *95* (6), 3263-8. http://www.ncbi.nlm.nih.gov/pubmed/9501251

Lamme, V. A. F., Rodriguez-Rodriguez, V. & Spekreijse, H. (1999). Separate processing dynamics for texture elements, boundaries and surfaces in primary visual cortex of the macaque monkey. *Cerebral Cortex*, *9* (4), 406-13. http://www.ncbi.nlm.nih.gov/pubmed/10426419

Lamme, V. A. F., Supèr, H., Landman, R., Roelfsema, P. R. & Spekreijse, H. (2000). The role of primary visual cortex (V1) in visual awareness. *Vision Research*, *40* (10-12), 1507-21. http://www.ncbi.nlm.nih.gov/pubmed/10788655

Lamme, V. A. F., Zipser, K. & Spekreijse, H. (2002). Masking interrupts figure-ground signals in V1. *Journal of Cognitive Neuroscience*, *14* (7), 1044-53. 10.1162/089892902320474490. http://www.ncbi.nlm.nih.gov/pubmed/12419127

Lamme, V. A. F. & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, *23* (11), 571-9. http://www.ncbi.nlm.nih.gov/pubmed/11074267

Lamme, V. A. F. & Spekreijse, H. (2000). Modulations of primary visual cortex activity representing attentive and conscious scene perception. *Frontiers in Bioscience*, *5*, D232-43. http://www.ncbi.nlm.nih.gov/pubmed/10704153

Li, F. F., Van Rullen, R., Koch, C. & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences*, *99* (14), 9596-9601. 10.1073/pnas.092277599

Macknik, S. L. & Livingstone, M. S. (1998). Neuronal correlates of visibility and invisibility in the primate visual system. *Nature Neuroscience*, *1* (2), 144-149. 10.1038/393

Maier J., Dagnelie G., Sprekreijse H. & Van Dijk B. W. (1987). Principal components-analysis for source localization of VEPs in man. *Vision Research*, *27* (2), 165-177. 10.1016/0042-6989(87)90179-9

Malach, R., Amir, Y., Harel, M. & Grinvald, A. (1993). Relationship between intrinsic connections and functional architecture revealed by optical imaging and in vivo targeted biocytin injections in primate striate cortex. *Proceedings of the National Academy of Sciences*, *90* (22), 10469-10473.

Maniscalco, B. & Lau, H. (2012). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and Cognition*, *21* (1), 422-430. 10.1016/j.concog.2011.09.021

Marcel, A. J. (1998). Blindsight and shape perception: Deficit of visual consciousness or of visual function? *Brain*, *121* (8), 1565-1588. 10.1093/brain/121.8.1565

Meng, M., Ferneyhough, E. & Tong, F. (2007). Dynamics of perceptual filling-in of visual phantoms revealed by binocular rivalry. *Journal of Vision*, *7* (13). 10.1167/7.13.8

Meuwese, J. D., Post, R. A., Scholte, H. S. & Lamme, V. A. F. (2013). Does perceptual learning require consciousness or attention? *Journal of Cognitive Neuroscience*, *25* (10), 1579-1596. 10.1162/jocn_a_00424

Meuwese, J. D., Scholte, H. S. & Lamme, V. A. F. (2014). Latent memory of unattended stimuli reactivated by practice: An FMRI study on the role of consciousness and attention in learning. *PLoS One*, *9* (3), e90098. 10.1371/journal.pone.0090098. http://www.ncbi.nlm.nih.gov/pubmed/24603676

Moore, T., Rodman, H. R., Repp, A. B. & Gross, C. G. (1995). Localization of visual stimuli after striate cortex damage in monkeys: Parallels with human blindsight. *Proceedings of the National Academy of Sciences*, *92* (18), 8215-8218. http://www.ncbi.nlm.nih.gov/pubmed/7667270

Moradi, F., Koch, C. & Shimojo, S. (2005). Face adaptation depends on seeing the face. *Neuron*, *45* (1), 169-175. 10.1016/j.neuron.2004.12.018

Moutoussis, K. & Zeki, S. (2002). The relationship between cortical activation and perception investigated with invisible stimuli. *Proceedings of the National Academy of Sciences*, *99* (14), 9527-9532. 10.1073/pnas.142305699

Mudrik, L., Breska, A., Lamy, D. & Deouell, L. Y. (2011). Integration without awareness: Expanding the limits of unconscious processing. *Psychological Science*, *22* (6), 764-770. 10.1177/0956797611408736

Nakayama, K. & Shimojo, S. (1992). Experiencing and perceiving visual surfaces. *Science*, *257* (5075), 1357-1363.

Nothdurft, H. C., Gallant, J. L. & Van Essen, D. C. (1999). Response modulation by texture surround in primate area V1: Correlates of "popout" under anesthesia. *Visual Neuroscience*, *16* (1), 15-34.

Oram, M. W. & Perrett, D. I. (1992). Time course of neural responses discriminating different views of the face and head. *Visual Neuroscience*, *68* (1), 70-84.

Overgaard, M., Rote, J., Mouridsen, K. & Ramsøy, T. Z. (2006). Is conscious perception gradual or dichotomous? A comparison of report methodologies during a visual task. *Consciousness and Cognition*, *15* (4), 700-708. 10.1016/j.concog.2006.04.002

Owen, A. M., Coleman, M. R., Boly, M., Davis, M. H., Laureys, S. & Pickard, J. D. (2006). Detecting awareness in the vegetative state. *Science*, *313* (5792), 1402. 10.1126/science.1130197

Panichello, M. F., Cheung, O. S. & Bar, M. (2012). Predictive feedback and conscious visual experience. *Frontiers in Psychology*, *3*. 10.3389/fpsyg.2012.00620

Pascual-Leone, A. & Walsh, V. (2001). Fast backprojections from the motion to the primary visual area necessary for visual awareness. *Science*, *292* (5516), 510-512. 10.1126/science.1057099

Rao, R. P. & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2* (1), 79-87. 10.1038/4580

Robichaud, L. & Stelmach, L. B. (2003). Inducing blindsight in normal observers. *Psychonomic Bulletin & Review*, *10* (1), 206-209.

Rock, I. & Palmer, S. (1990). The legacy of Gestalt psychology. *Scientific American*, *263* (6), 84-90. http://www.ncbi.nlm.nih.gov/pubmed/2270461

Roe, A. W., Lu, H. D. & Hung, C. P. (2005). Cortical processing of a brightness illusion. *Proceedings of the National Academy of Sciences*, *102* (10), 3869-3874. 10.1073/pnas.0500097102

Roelfsema, P. R. (2006). Cortical algorithms for perceptual grouping. *Annual Review of Neuroscience*, *29*, 203-227. 10.1146/annurev.neuro.29.051605.112939

Roelfsema, P. R., Lamme, V. A. F., Spekreijse, H. & Bosch, H. (2002). Figure-ground segregation in a recurrent network architecture. *Journal of Cognitive Neuroscience*, *14* (4), 525-537. 10.1162/08989290260045756

Rolls, E. T. (1992). Neurophysiological mechanisms underlying face processing within and beyond the temporal cortical visual areas. *Philosophical Transactions of the Royal Society*, *335* (1273), 11-21. 10.1098/rstb.1992.0002

Rolls, E. T. & Tovee, M. J. (1994). Processing speed in the cerebral cortex and the neurophysiology of visual masking. *Proceedings of the Royal Society B: Biological Sciences*, *257* (1348), 9-15. 10.1098/rspb.1994.0087. http://www.ncbi.nlm.nih.gov/pubmed/8090795

Rossi, A. F., Rittenhouse, C. D. & Paradiso, M. A. (1996). The representation of brightness in primary visual cortex. *Science*, *273* (5278), 1104-1107.

Rossi, A. F. & Paradiso, M. A. (1999). Neural correlates of perceived brightness in the retina, lateral geniculate nucleus, and striate cortex. *Journal of Neuroscience*, *19* (14), 6145-56.

Scholte, H. S., Witteveen, S. C., Spekreijse, H. & Lamme, V. A. F. (2006). The influence of inattention on the neural correlates of scene segmentation. *Brain Research*, *1076* (1), 106-115. 10.1016/j.brainres.2005.10.051

Schwiedrzik, C. M., Singer, W. & Melloni, L. (2011). Subjective and objective learning effects dissociate in space and in time. *Proceedings of the National Academy of Sciences*, *108* (11), 4506-4511. 10.1073/pnas.1009147108

Seghier, M. L. & Vuilleumier, P. (2006). Functional neuroimaging findings on the human perception of illusory contours. *Neuroscience & Biobehavioral Reviews*, *30* (5), 595-612. 10.1016/j.neubiorev.2005.11.002

Seitz, A. R., Kim, D. & Watanabe, T. (2009). Rewards evoke learning of unconsciously processed visual stimuli in adult humans. *Neuron*, *61* (5), 700-707. 10.1016/j.neuron.2009.01.016

Seitz, A. R. & Watanabe, T. (2003). Psychophysics: Is subliminal learning really passive? *Nature*, *422* (6927), 36. 10.1038/422036a

Self, M. W., Kooijmans, R. N., Supèr, H., Lamme, V. A. F. & Roelfsema, P. R. (2012). Different glutamate receptors convey feedforward and recurrent processing in macaque V1. *Proceedings of the National Academy of Sciences of the United States of America*, *109* (27), 11031-6. 10.1073/pnas.1119527109. http://www.ncbi.nlm.nih.gov/pubmed/22615394

Self, M. W., Van Kerkoerle, T., Supèr, H. & Roelfsema, P. R. (2013). Distinct roles of the cortical layers of area V1 in figure-ground segregation. *Current Biology*, *23* (21), 2121-2129. 10.1016/j.cub.2013.09.013

Sergent, C. & Dehaene, S. (2004). Is consciousness a gradual phenomenon? Evidence for an all-or-none bifurcation during the attentional blink. *Psychological Science*, *15* (11), 720-728. 10.1111/j.0956-7976.2004.00748.x. http://www.ncbi.nlm.nih.gov/pubmed/15482443

Seth, A. K. (2010). The grand challenge of consciousness. *Frontiers in Psychology*, *1*. 10.3389/fpsyg.2010.00005

Seth, A. K., Suzuki, K. & Critchley, H. D. (2011). An interoceptive predictive coding model of conscious presence. *Frontiers in Psychology*, *2*, 395. 10.3389/fpsyg.2011.00395. http://www.ncbi.nlm.nih.gov/pubmed/22291673

Shapley, R. & Hawken, M. J. (2011). Color in the cortex: Single- and double-opponent cells. *Vision Research*, *51* (7), 701-717. 10.1016/j.visres.2011.02.012

Sigman, M., Pan, H., Yang, Y., Stern, E., Silbersweig, D. & Gilbert, C. D. (2005). Top-down reorganization of activity in the visual pathway after learning a shape identification task. *Neuron*, *46* (5), 823-835. 10.1016/j.neuron.2005.05.014

Sigman, M. & Gilbert, C. D. (2000). Learning to find a shape. *Nature Neuroscience*, *3* (3), 264-269. 10.1038/72979

Silvanto, J., Lavie, N. & Walsh, V. (2005). Double dissociation of V1 and V5/MT activity in visual awareness. *Cerebral Cortex*, *15* (11), 1736-1741. 10.1093/cercor/bhi050

Simons, D. J. & Rensink, R. A. (2005). Change blindness: past, present, and future. *Trends in Cognitive Sciences*, *9* (1), 16-20. 10.1016/j.tics.2004.11.006. http://www.ncbi.nlm.nih.gov/pubmed/15639436

Snodderly, D. M. & Gur, M. (1995). Organization of striate cortex of alert, trained monkeys (Macaca fascicularis): Ongoing activity, stimulus selectivity, and widths of receptive field activating regions. *Journal of Neurophysiology*, *74* (5), 2100-2125.

Sperry, R. (1984). Consciousness, personal identity and the divided brain. *Neuropsychologia*, *22* (6), 661-673. 10.1016/0028-3932(84)90093-9

Stein, T., Hebart, M. N. & Sterzer, P. (2011). Breaking continuous flash suppression: A new measure of unconscious processing during interocular suppression? *Frontiers in Human Neuroscience*, *5*. 10.3389/fnhum.2011.00167

Steyn-Ross, M. L., Steyn-Ross, D. A., Sleigh, J. W. and Liley, D. T. (1999). Theoretical electroencephalogram stationary spectrum for a white-noise-driven cortex: Evidence for a general anesthetic-induced phase transition. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, *60* (6 Pt B), 7299-7311. 10.1103/PhysRevE.60.7299

Stoerig, P. & Cowey, A. (1989). Wavelength sensitivity in blindsight. *Nature*, *342* (6252), 916-918. 10.1038/342916a0

Straube, T., Mothes-Lasch, M. & Miltner, W. H. (2011). Neural mechanisms of the automatic processing of emotional information from faces and voices. *British Journal of Social Psychology*, *102* (4), 830-848. 10.1111/j.2044-8295.2011.02056.x

Sugase, Y., Yamane, S., Ueno, S. & Kawano, K. (1999). Global and fine information coded by single neurons in the temporal visual cortex. *Nature*, *400* (6747), 869-873. 10.1038/23703

Sugihara, T., Qiu, F. T. & Von der Heydt, R. (2011). The speed of context integration in the visual cortex. *Journal of Neurophysiology*, *106* (1), 374-385. 10.1152/jn.00928.2010

Supèr, H., Spekreijse, H. & Lamme, V. A. F. (2001). Two distinct modes of sensory processing observed in monkey primary visual cortex (V1). *Nature Neuroscience*, *4* (3), 304-310. 10.1038/85170

Supèr, H., Van der Togt, C., Spekreijse, H. & Lamme, V. A. F. (2003). Internal state of monkey primary visual cortex (V1) predicts figure-ground perception. *Journal of Neuroscience*, *23* (8), 3407-3414.

Supèr, H. & Lamme, V. A. F. (2007). Altered figure-ground perception in monkeys with an extra-striate lesion. *Neuropsychologia*, *45* (14), 3329-3334. 10.1016/j.neuropsychologia.2007.07.001

Swets, J. A., Green, D. M., Getty, D. J. & Swets, J. B. (1978). Signal detection and identification at successive stages of observation. *Perception & Psychophysics*, *23* (4), 275-289.

Sáry, G., Köteles, K., Kaposvári, P., Lenti, L., Csifcsák, G., Frankó, E., Benedek, G. & Tompa, T. (2008). The representation of Kanizsa illusory contours in the monkey inferior temporal cortex. *European Journal of Neuroscience*, *28* (10), 2137-2146. 10.1111/j.1460-9568.2008.06499.x

Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annual Review of Neuroscience*, *19*, 109-139. 10.1146/annurev.ne.19.030196.000545. http://www.ncbi.nlm.nih.gov/pubmed/8833438

Tononi, G. (2004). An information integration theory of consciousness. *BMC Neuroscience*, *5* (42). 10.1186/1471-2202-5-42. http://www.ncbi.nlm.nih.gov/pubmed/15522121

——— (2008). Consciousness as integrated information: A provisional manifesto. *Biological Bulletin*, *215* (3), 216-242.

——— (2012). Integrated information theory of consciousness: an updated account. *Archives Italiennes de Biologie*, *150* (4), 293-329.

Tononi, G. & Massimini, M. (2008). Why does consciousness fade in early sleep? *Annals of the New York Academy of Sciences*, *1129*, 330-334. 10.1196/annals.1417.024

Tsuchiya, N. & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature Neuroscience*, *8* (8), 1096-1101. 10.1038/nn1500

Van der Togt, C., Kalitzin, S., Spekreijse, H., Lamme, V. A. F. & Supèr, H. (2006). Synchrony dynamics in monkey V1 predict success in visual detection. *Cerebral Cortex*, *16* (1), 136-148. 10.1093/cercor/bhi093

Van Gaal, S. & Lamme, V. A. F. (2012). Unconscious high-level information processing: Implication for neurobiological theories of consciousness. *Neuroscientist*, *18* (3), 287-301. 10.1177/1073858411404079

Van Loon, A. M., Fahrenfort, J. J., Van der Velde, B., Lirk, P. B., Vulink, N. C. C., Hollmann, M. W., Scholte, H. S. and Lamme, V. A. F. (submitted). NMDA receptor antagonist ketamine distorts object recognition by reducing feedback to early visual cortex. *The Journal of Neuroscience*

Von der Heydt, R., Peterhans, E. & Baumgartner, G. (1984). Illusory contours and cortical neuron responses. *Science*, *224* (4654), 1260-1262.

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M. and Von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization. *Psychological Bulletin*, *138* (6), 1172-1217. 10.1037/a0029333

Wang, L., Weng, X. & He, S. (2012). Perceptual grouping without awareness: Superiority of Kanizsa triangle in breaking interocular suppression. *PLoS One*, *7* (6), e40106. 10.1371/journal.pone.0040106

Weiskrantz, L. (1996). Blindsight revisited. *Current Opinion in Neurobiology*, *6* (2), 215-220. 10.1016/S0959-4388(96)80075-4

Wokke, M. E., Sligte, I. G., Scholte, H. S. & Lamme, V. A. F. (2012). Two critical periods in early visual cortex during figure-ground segregation. *Brain and Behavior*, *2* (6), 763-777. 10.1002/brb3.91

Yamahachi, H., Marik, S. A., McManus, J. N., Denk, W. & Gilbert, C. D. (2009). Rapid axonal sprouting and pruning accompany functional reorganization in primary visual cortex. *Neuron*, *64* (5), 719-729. 10.1016/j.neuron.2009.11.026

Zipser, K., Lamme, V. A. F. & Schiller, P. H. (1996). Contextual modulation in primary visual cortex. *Journal of Neuroscience*, *16* (22), 7376-7389.

Consciousness as Inference in Time

A Commentary on Victor Lamme

Lucia Melloni

Unraveling the neural correlates of conscious remains one of the great challenges of our time. Victor Lamme proposes that neural integration through feedback loops is what differentiates conscious from unconscious processing. Here, I review his hypothesis, focusing on the spatial scale of integration as well as the possible neural mechanisms involved. I go on to show that any theory of the neural correlates of consciousness is incomplete if it cannot account for how prior knowledge shapes perception and how this form of integration occurs. Finally, I propose that integration across moments in time is a crucial but hitherto neglected aspect of conscious perception, which creates the "flow" of conscious experience.

Commentator

Lucia Melloni
lucia.melloni@brain.mpg.de

Max Planck Institute for Brain
Research
Frankfurt a. M., Germany

Target Author

Victor Lamme
victorlamme@gmail.com

Universiteit van Amsterdam
Amsterdam, Netherlands,

Editors

Thomas Metzinger
metzinger@uni-mainz.de

Johannes Gutenberg-Universität
Mainz, Germany

Jennifer M. Windt
jennifer.windt@monash.edu

Monash University
Melbourne, Australia

1 Qualia 2.1: Integration is key but is it all?

Why do we see the way we *see*? How is our perception different from the way a photograph is acquired on the sensor chip of a digital camera? It seems obvious that we do not see an image made of individual pixels but an integrated, smooth, colourful, and vivid image. What is the neural substrate of this marvellous capacity that makes us feel and experience the way we do? These are the central questions that Victor Lamme sets out to address in his paper *The Crack of Dawn: Perceptual Functions and Neural Mechanisms that Mark the Transition from Unconscious Processing to Conscious Vision.*

This is by no means an easy task, even when one stays away from the difficult problem of qualia or "what it is like to be" (Nagel 1974). The question of how awareness arises has preoccupied philosophers and scientists for centuries, and while significant progress has been made in recent decades we are still far from reaching a conclusion (Dehaene 2014; Koch 2004). One thing is clear however: success in understanding the neural machinery that instantiates consciousness rests on identifying the fundamental features that characterise a state as conscious and that distinguish it from unconscious states.

A remarkable discovery of the past century is that a significant portion of all mental operations, including fairly complex ones such as decision-making and perceptual categorisation, can be carried out unconsciously. Take the case of language: while it seems effortless to understand the words that you are currently reading, you do not have conscious access to the syntactic processes that ultimately allow you to grasp the relations between the elements of this sentence and thus its meaning. These complex mental operations occur "behind the scene" of consciousness. Given that so many intricate processes can operate unconsciously, one cannot but wonder what consciousness is good for. Which mental processes *require* consciousness, if any? And if so, what really distinguishes conscious from unconscious cognition? Victor Lamme offers a stimulating and comprehensive review of processes in vision that can be performed outside the realm of awareness. The list is long and may be surprising (also see Kouider & Dehaene 2007), ranging from detection of simple (e.g., oriented lines) and complex features (e.g., faces; Almeida et al. 2013; de Gardelle et al. 2011; Del Zotto et al. 2013), to mathematical operations such as abstract comparisons between quantities (Greenwald et al. 2003), to triggering of motor plans (Dehaene et al. 1998), and even error-related responses to stimuli that fully escape our consciousness (Cohen et al. 2009).

What do we need consciousness for, then? Lamme proposes that consciousness is required when all sources of information need to be integrated. For instance, when we see a face, we can not only *detect* that is a face, a process that can be performed unconsciously, but also *identify* it as that of our friend Billy, whom we have not seen in ten years and that we *remember* warmheartedly from our childhood. Consciousness brings this *unified* moment in which all comes together: previous experiences are retrieved from memory (e.g., do we have reason to like Billy?) and unified with the context of the current experience (e.g., where are we now?), but also intertwined with predictions for future actions (e.g., would we like to engage in a conversation?). Thus, in one single moment, past,

present and future come together and form a unified conscious experience. Many scientists nowadays agree that conscious experience provides an abstract summary of all available sources of information, from which many features are filtered out and reinterpreted in a format that is most useful for further actions, thoughts, deliberations, and chain operations that cannot be processed by non-conscious processors (Lamme this collection; Baars 2002; Dehaene 2014; Melloni & Singer 2010). Hence, what reaches our perception is a highly processed, "interpreted" version of the world. One key intuition is that the unification and "interpretation" of the experience that reaches our consciousness is achieved through the activation of myriads of neurons that signal individual features, but that it is by virtue of *integrating* their information through dynamic interactions (for example via synchronous coordination of their activity or via feedback processes) that a coherent experience across senses, space, and time comes about.

An important caveat is that integration of information *per se* is unlikely to distinguish conscious from unconscious processing as integration of many features can also proceed unconsciously (Dehaene et al. 1998; Gaillard et al. 2009; Lin & He 2009; Melloni et al. 2007; Melloni & Rodriguez 2007; Mudrik et al. 2014). In fact, integration through convergence is a key principle of the wiring of the brain, which explains the mere existence of feature-selective neurons that respond to motion, shape, or complex stimuli such as faces, and that process information in an unconscious manner. If it is not integration *per se*, then what kind of integration are we talking about?[1] We and others (Melloni & Singer 2010; Thompson & Varela 2001;

---

1 Giulio Tononi (2004; Tononi & Koch 2008) argues that not only integration but also differentiation/segregation (e.g., distinguishing a particular state from all possible other states) is characteristic of conscious states. However, even when both conditions are met, say integration through convergence is observed in FFA and differentiated from other states, e.g., there is no activation in PPA, an area selective to processing places, and thus there is no guarantee that this would constitute a conscious state. In fact, experimental evidence suggests that such feature-selective processing can indeed proceed unconsciously, for example in the case of face processing under conditions of masking (de Gardelle et al. 2011), continuous flash suppression (Almeida et al. 2013), and in blindsight patients (Del Zotto et al. 2013).

Varela et al. 2001) have previously argued for a distinction between *local* and *global* integration, and proposed that the spatial scale of integration differentiates between unconscious and conscious states: unconscious processing is observed when local integration occurs within the divergent-convergent feedforward architecture; conscious processing however requires long-range integration through neural synchronization, which integrates information across the various levels of the cortical processing hierarchy.

Indeed, in recent years, a wealth of experimental studies (Aru et al. 2012; Gaillard et al. 2009; Hipp et al. 2011; Melloni et al. 2007; Melloni & Rodriguez 2007) have provided support to the idea that long-range integration through synchronous coupling is a mechanism for conscious perception, and that the spatial scale of synchronisation strongly correlates with the perceptual outcome. For example, we have shown that masked words are only consciously perceived when accompanied by a burst of long-distance synchronization in the gamma band, while unconscious processing, even up to a semantic level, elicits only local gamma oscillations (Melloni et al. 2007; Melloni & Rodriguez 2007). Although controversy still persists as to whether long-range integration necessitates the involvement of particular brain areas (Dehaene 2014; Edelman & Tononi 2000) or not (Lamme this collection; Melloni & Singer 2010), it is reassuring to witness some convergence on the results that have even led to clinical applications (e.g., coma classification, King et al. 2013). In his most recent work, Victor Lamme now also assigns a central role to the spatial scale of the integration for consciousness, joining an ever-increasing number of researchers proposing long-range integration as key to consciousness (Dehaene & Changeux 2011; Edelman & Tononi 2000; Melloni & Singer 2010; Thompson & Varela 2001). An interesting point of divergence from other theories is that while Lamme assigns a particular role to feedback and horizontal connections in the integration of information for consciousness, other theories, including our own, hypothesise that it is the synchronisation of neural populations that glues all experiences into one, thereby instantiating con-

sciousness. As empirical data and theoretical considerations continue to accumulate, we expect that this and other pressing challenges such as identifying how far is "long" in the brain, or whether "long" involves the activation of specific neural cell populations, specific areas, and/or a specified number of nodes will become addressable.

However, imagine those questions have been addressed and we know that integration on a particular spatial scale is key to consciousness; would we have understood what consciousness is or how it comes about? Here I propose that we would not, as any theory that does not account for two fundamental, hitherto neglected aspects of conscious experience will fall short of explaining consciousness. In particular, our experience is never an island in isolation, but instead is shaped by previous knowledge, by priors that stem from the preceding context or from our history of learning. These priors determine our perception; and thus understanding how they become integrated is paramount to explaining consciousness. However, an even more pressing problem is that conscious experience unfolds over time, whereby the recent past moulds the current moment, which in turn creates predictions for moments to come, i.e., the future. How all those temporal processes intertwine and define our experience (the flow of consciousness) is something that most research has neglected. In the following sections I will review current research that we and others have undertaken with the purpose of raising awareness of these overlooked integrative properties of conscious experience and the challenges that they entail for the study of consciousness.

2 Consciousness as an inferential process and the consequences for the neural mechanism of conscious perception

One central and characteristic feature of conscious perception is its constructive nature. In contrast to unconscious cognition, which is directly driven by sensory stimulation, the images that reach consciousness often bear little resemblance to reality. Indeed, percepts in our

mind can be understood as useful distortions of reality in which only specific parts of the physical input are represented while being enriched with a model of the world that has been learned and that provides context to the current moment. In the words of Heinz von Foerster (1984), "the world, as we perceive it, is our own invention". To provide a striking example of this, consider the image on the right (Figure 1) and try to figure out what it shows. Most people at first see a collection of black and white blobs, much like the input that strikes our retina—a raw, uninterpreted signal. Now, rotate the page upside down. Voila! You will clearly see a face (do you recognize whose face it is?). Remarkably, you can turn the page back and you will continue seeing the face. Once you have recognized the image, the visual system has created a *prior*, an expectation that enriches perception. This example is not mere curiosity. Most of our behaviour and perceptions are based on predictions: we do not wait for visual input to impinge our eyes, we *actively* look for it. We cannot, however, initiate a rational search for an object without making *predictions* about "what" it is, "where" it is likely to be, and even "when" it is likely to be there. The brain's ability to make predictions and to mould its data gathering accordingly is thus essential for its ability to evaluate options, make life-critical decisions, and generate adaptive behaviour.

While the constructive nature of perception is undeniable and may even appear as one of its defining features, surprisingly little research has been carried out to understand how previous experience interacts with consciousness. Most importantly, the scientific community has not embraced an understanding of consciousness in the context of a flow of experience in which every moment is integrated with past moments and interfaced with expectations about what will happen in the future (but see Varela 1999). A possible reason for neglecting the contribution of previous experience is that this integration of past with present moments has been understood as a process of "unconscious inference" (following von Helmholtz 1866/1962), or, in Victor Lamme's words, in the context of the "automaticity of the many ex-

pectation effects." However, this inferential process is carried out in the backstage of consciousness, and it is only the result that we consciously experience. This bears resemblance to syntactic analysis, which is also carried out automatically and unconsciously, but is paramount to conscious access to meaning. Without unconscious syntactic analysis we would not be able to "consciously" understand text; nor is its automatic activation under our control. In the same vein, our conscious perception would be totally different if prior knowledge did not help us enrich or even construct our experience, endowing it with meaning. In fact, it has been proposed that alterations in perception, i.e., the defragmented sensory experience observed in schizophrenics and autistic people can be the result of a deficit in this inferential process (Jardri & Deneve 2013; Pellicano & Burr 2012), underscoring the fundamental role that perceptual inference plays in conscious perception.

One promising framework within which the influence of previous experience through unconscious inference can be understood is the Bayesian framework. When applied to perception, each mathematically-formulated ingredient of this framework can be assigned a perceptual counterpart, with previous experience referring to the prior, the current moment referring to the likelihood, unconscious inference referring to Bayes rule (which combines the prior with the likelihood in an optimal way), and the result—our perception—referring to (the peak of) the posterior distribution. This idea has recently proven to be a powerful tool for understanding perception not only in terms of modelling behaviour, but also as a theoretical framework for understanding how perception arises in the brain. A prominent implementation of the latter is Predictive Coding (Friston 2010). This theory postulates that the brain builds models (priors) of the world based on previous experience, which are used to explain the current inputs. This occurs iteratively *across all levels in the cortical hierarchy* with the goal of minimising predictions errors, i.e., the difference between what is expected and the incoming sensory input, which are energetically costly. This minimization process can either be achieved by chan-

ging the way the system samples its environment, or by changing its models. Relevant for this discussion is the idea that perceptual inference, in the Predictive Coding framework, implies that all levels in the hierarchy reach an agreement, i.e., minimise all prediction errors, much like the idea of a unified/integrative moment as proposed by Victor Lamme and others (Dehaene 2014; Edelman & Tononi 2000; Melloni & Singer 2010). While Predictive Coding by itself is currently agnostic as to whether such unified agreement represents a conscious state, the central tenet that integration across all levels is what the system strives for still holds. This allows for the formulation of interesting, testable predictions about the Neural Correlates of Consciousness (NCC).

In recent years research in my lab has focused on understanding how previous experience enriches perception, how expectations alter the NCC, and how this can be understood within the Predictive Coding framework. The central idea that motivated these studies was to test whether or not the NCC are context independent, i.e., impervious to the influence of expectations, as many theories implicitly postulate. To test this hypothesis we presented subjects with illusory letters, that is letters whose borders where not explicitly defined but instead required the activation of figure–ground segregation cues. We reasoned that providing subjects with a prior, i.e., knowing which letter would be presented next, would facilitate the figure–ground segregation process, making an initially invisible letter clearly visible. In line with our expectations, we observed that the threshold of conscious perception is not fixed but instead changes depending on the availability of previous knowledge: subjects are able to perceive a stimulus on the basis of minimal sensory information when they have a clear expectation. We were able to confirm this result in a series of different paradigms in which expectations could be generated online from recent experience as in the example of the letter given above (Melloni et al. 2011; Schwiedrzik et al. 2014), drawn from memory based on prior exposure to clearly visible natural images (Aru et al. 2012), stem from a life-long history of association between letters

and colour as in grapheme–colour synaesthesia (van Leeuwen et al. 2013), or result from systematic training as in perceptual learning (Schwiedrzik et al. 2009, 2011). These studies allowed us to test not only whether the behavioural threshold of conscious perception is fixed, but also how previous knowledge would affect the neural "construction" of conscious percepts.

Figure 1: Can you recognize what this is? If not, rotate the image. Note that once you turn it back around the object is now clear.

A first hypothesis we derived from the Predictive Coding framework was that the presence of strong priors should have an effect of how quickly content reaches awareness. If conscious perception is the result of a process that iterates until information is consistent between the different levels of the hierarchy (Di Lollo et al. 2000), i.e., until all prediction errors are minimised, then having a better model of the input based on prior knowledge may speed up this process. Indeed and contrary to the com-
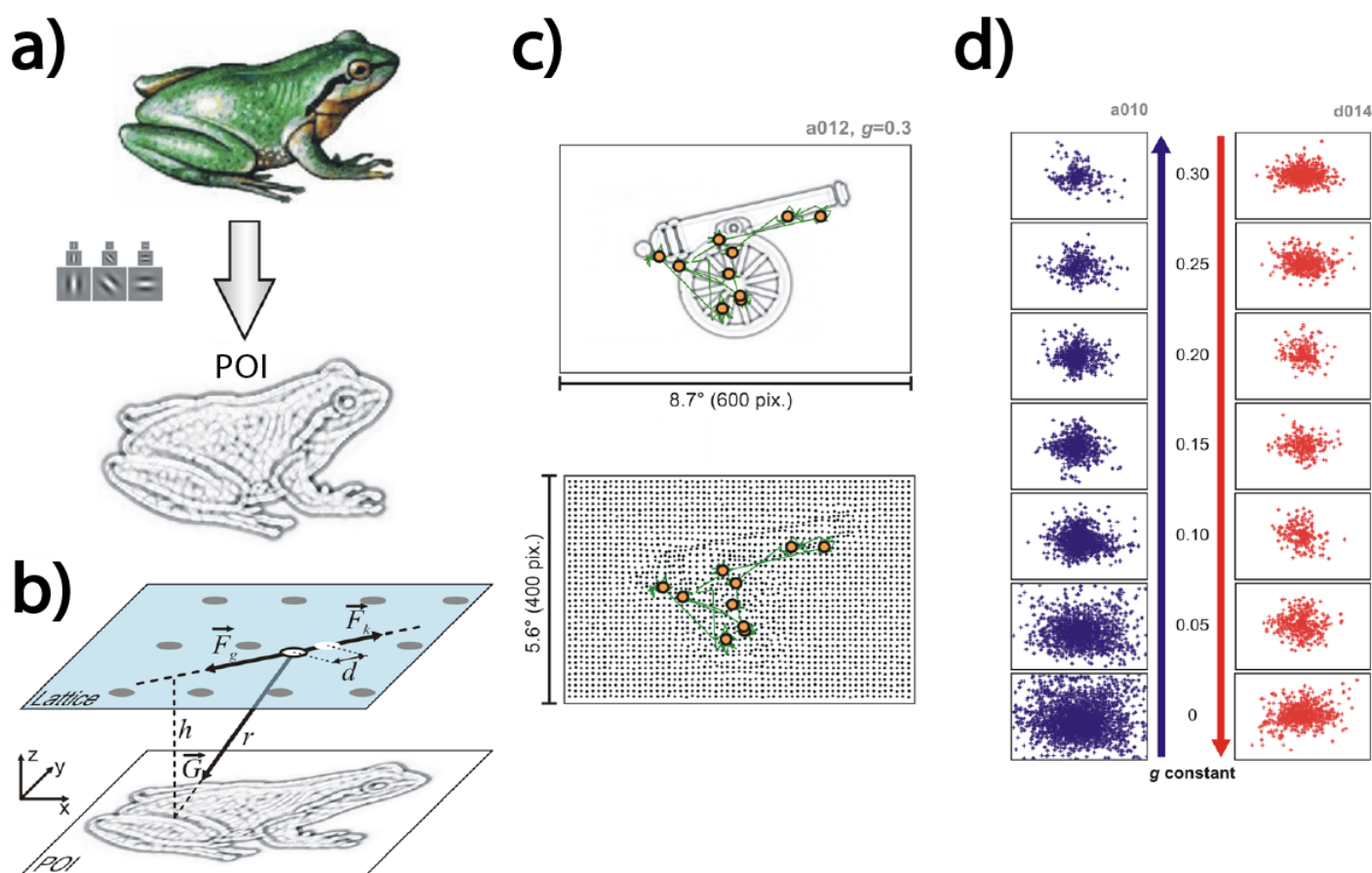
Figure 2: (a) Original images are filtered through a series of gabor wavelets, which allows the estimation of the points of maximal local information (Points of Maximal Information, POI) in the source image. (b) Dots of an elastic lattice are created by mapping the POI in the projection plane, and attracting them by the projection F0 of a gravitational force G. (c) Pattern of saccades/fixations when subjects recognise a stimulus and its underlying POI map. (d) Pattern of fixations for stimuli of different degradation levels from high degradation (0) to low degradation (0.30). Dots in blue correspond to fixations when subjects do not have an expectation of the stimuli, dots in red correspond to patterns of fixation observed in the presence of expectations. Note that in the presence of expectations, the distribution of fixations are much less scattered. From Moca et al. (2011).

mon belief that information processing in the brain has a fixed latency, we observed that the NCC shifts in time when a prior is available. While the electrophysiological difference between seen and unseen letters occurred around 300ms when it exclusively depended on sensory evidence, it occurred as early as 200ms when priors were available (Melloni et al. 2011). Thus, priors sped up information processing by 100ms. These results have important implications for the search for the NCC as they show that conscious processing is not bound to a particular time, but can flexibly adjust its timing depending on the task at hand, the readiness of the system, or the presence of expectations.

They also pose a challenge to theories that postulate that the NCC always occur late, as proposed by Victor Lamme (this collection) or Stanislas Dehaene (2014).

A second prediction that follows from the principle of minimising prediction errors is that in the presence of priors, activity in lower areas can be "explained away" by priors in higher brain areas (Murray et al. 2004); this entails that when inputs can be fully predicted based on previous experience, they do not elicit prediction errors. To test this hypothesis, we took the same study to the MEG and performed source localisation. Here, we found that priors sparsify the networks involved in processing the

stimulus, such that when a prior is present only the brain areas that are most diagnostic to the stimulus features are activated (Mayer et al. in preparation). All alternative interpretations of the stimulus are thus "explained away". Thus, consciousness and its neural correlates appear as mobile targets, which adjust their locus in the presence of expectations. This poses a further challenge to the search for the NCC, as not only the timing, but also the location of neural activation does not appear as a diagnostic feature for the NCC.

Finally, Predictive Coding also suggests that priors may be used to change the way information is sampled, as the models derived from previous experience can be used to optimise the search for the most relevant information (Friston et al. 2012). Only rarely do we keep our gaze still and wait for the world to bring novel information; instead, we scan images through rhythmic patterns of eye movements accompanied by fixations. This active sensing view implies that perception is not a passive phenomenon in which the system waits for information to hit the sensory transducers, but instead an active process that seeks information through exploratory routines (Melloni et al. 2009; Schroeder et al. 2010). To test whether and how priors affect the sampling of information we developed stimuli for which we could quantify the local information content at each point (Figure 2) and determined the efficiency of information extraction based on eye movements in the presence or absence of expectations. Figure 2 shows that when subjects have prior knowledge of the object they are trying to perceive, they can immediately orient their eyes to areas of most diagnostic information for the perception of an object. At the same time, the sampling of information becomes sparser, concentrating eye movements to maximally informative areas (Moca et al. 2011). This implies that priors direct our exploratory motor routines, thus optimising perception.

Overall, these studies show that previous experience enriches the contents of consciousness and fundamentally changes the way information is processed in our brain, enhancing speed and efficiency. This raises questions for theories that propose a fixed latency or neural locus for conscious access, but also complicates the quest for the NCC, as they turn out to differ in time and location depending on the precision and accuracy of expectations. Although current formulations of Predictive Coding do not make specific predictions about consciousness, this framework may nevertheless prove to be an important starting point in trying to understand these effects. In fact, more explicit theoretical links between Predictive Coding and consciousness are now being worked out (e.g., Clark 2013; Hohwy 2013; Seth et al. 2011)—after all, Predictive Coding has been framed as a unifying theory of the brain (Friston 2010), which would fall short if consciousness was left unexplained.

3 The neglected dimension of consciousness: Time and the flow of consciousness

But is that all? One dimension of our experience that is often neglected is time. Of course, time is an implicit component of previous experience, however, it may also be revealing to consider time by itself. In fact, living organisms seldom encounter a static image in isolation, but are instead confronted with a flow of temporally-correlated sensory inputs (Schwartz et al. 2007). Imagine for instance a tennis match, and picture the tennis ball flying over the field. If queried, you could easily estimate where the ball is, but also where it was a second ago and where it will be in a few milliseconds. Event-objects of the conscious mind[2] thus per definition unfold in time and we also act in time: we make use of current and previous input to figure out the most appropriate response predicting their consequences. There is thus a continuum of interdependencies along the time dimension whereby every past moment is *integrated* with the present and projected into the future, giving rise to the flow of consciousness. The same way

---

2 We are usually conscious of objects, and become so by virtue of their being differentiated from the background, but also because their internal features are linked or bound in some way. Objects and their internal features do not need to be static entities but can have temporal dynamics, i.e., they develop or change in time. In this case, they become events (and thus event-objects of the conscious mind).

we have been thinking about the integration of multiple source of information *within* a given moment of time, such as multiple features of a single object, there is thus integration *across* time. A case in point is strikingly vivid perceptual aftereffects, such as the waterfall illusion, where viewing motion in one direction for several seconds causes a subsequently presented static image to move in the opposite direction (Purkinje 1820). Such effects are not limited to basic perceptual features such as motion direction, colour, or orientation, but also affect high-level percepts such as the perceived gender of faces (Webster et al. 2004), numerosity (Burr & Ross 2008), or gaze direction (Jenkins et al. 2006); and they are not limited to fleeting illusions that vanish almost instantaneously, but may persist for days or even weeks (Jones & Holding 1975). This indicates that our current experience is embedded into a continuous flow of previous experience at multiple time scales, ranging from lifelong experience with our environment to short-term, moment-by-moment effects that arise from our most recent encounters, even if just milliseconds ago.

The past thus leaves traces (predictions) that determine the current contents of consciousness. This has the consequence that the contents of consciousness represent an aggregate of imprints from the past and the present moment that jointly promote a sense of *stability over time*. However, through which mechanism these interdependencies affect our perception is currently unclear. Experimentally, the multiple time-scales of previous experience are particularly evident when subjects are confronted with sequences of multistable stimuli such as the Necker cube.[3] Because the sensory information these stimuli provide by themselves is insufficient to determine perception, they are particularly susceptible to the effects of previous experience. Under these conditions, one can observe two different effects that temporal dependencies entail: on the one hand, an *attractive* effect, which increases the likelihood of continuing to perceive the same stimulus, and on the other hand a *repulsive* effect, which increases the like-

lihood of perceiving something different. The former is often referred to as hysteresis, priming, stabilisation, or perceptual memory, while the latter is commonly known as perceptual adaptation.

Recently, Chopin & Mamassian (2012) studied the temporal dynamics of these serial dependencies, addressing the question of which part of the perceptual history the system retains and how remote and recent experiences differentially determine perception. They observed a remarkable dissociation between long stretches of time that occurred in the remote past (in their case several minutes) and short stretches of time that had just recently occurred (a few seconds ago): while the former had a positive correlation with perception, and thus ensured stability over time (hysteresis), the latter had a negative correlation to perception, that is, it promoted alternative interpretations (adaptation). These two timescales indicate that previous experience can act along at least two separate timescales and hence, that there may be several mechanisms at work. Using functional magnetic resonance imaging, we set out to further elucidate how these effects are implemented in the brain, how the brain entertains these two opposing processes without mutual interference, and what determines their direction (Schwiedrzik et al. 2014). Presenting multistable visual stimuli sequentially, we found that although affecting our perception concurrently, hysteresis and adaptation map into distinct cortical networks: a widespread network of higher-order visual and fronto-parietal areas was involved in hysteresis, while adaptation was confined to early visual areas (areas V2/V3). Importantly, hysteresis and adaptation bear a differential relation with whether or not the stimuli were consciously perceived: while adaptation was present even if the adapting interpretation was not consciously perceived (in agreement with previous reports, e.g., Hock et al. 1996), hysteresis depended on what was previously consciously perceived. Hence, conscious experiences in the past affected the present experience, preserving continuity in time, while unconscious processing had the opposite effect, bringing change and novelty to perception.

---

[3] But they are by no means limited to ambiguous stimuli (Fischer & Whitney 2014; Treisman 1984).

This brings us back to the question of neural integration, indicating that even in the case of *integration over time,* the spatial scale at which neuronal processing occurs determines whether content enters awareness or not: in the case of hysteresis, a conscious moment is integrated in time with another conscious moment, which involves a widespread cortical network, while in the case of adaptation, prior information is only integrated within a local module, which happens irrespective of whether this prior information is consciously experienced or not, similar to Lamme's "base grouping". This interpretation fits with results that have been obtained in the auditory domain in which short temporal regularities can be detected unconsciously eliciting a locally generated event-related potential (ERP), termed mismatch negativity (MMN), while detection of long-term regularities depends on conscious perception, which elicits an electrophysiological response known as P300 from a widespread network of brain areas (Bekinschtein et al. 2009; Faugeras et al. 2011).

Together, I propose that these results mesh well with the idea that one of the functions of consciousness is to interpret the world in *long timescales*, bringing together the *now* with the past beyond the simple and automatic input-output relations rooted in unconscious processors, thus allowing for the extraction of more complex and abstract regularities. Brain areas with longer time constants such as the prefrontal cortex (Fuster 1973) would extract the world's statistics from the remote past, creating a model of the world that keeps a stable picture. In contrast, early sensory areas with short time constants act on shorter timescales, sampling the world for alternative interpretations, thus allowing the system to stay tuned to deviations from the long-term statistics (Clifford 2012; Snyder et al. under review).[4]

While previous studies and established experimental paradigms have mostly focused on the "nowness" of conscious perception, it ap-

pears that much remains to be learned about consciousness and its fundamental phenomenological characteristics such as its flow and our sense of stability over time. In fact, considering that much of what we currently know about the NCC stems from "static" paradigms, and by those I mean paradigms that do not take the temporal context in which the stimuli unravel into account and thus only inform us about what has "changed" in consciousness, we in fact only have access to the neural processes related to the *update* of contents in consciousness, while the mechanisms at play in the *maintenance* or continuity of our experience remain obscure (but see Kleinschmidt et al. 2002). The *present* might be known, but the *flow* is still a mystery!

Thus I propose that a full account of consciousness requires a reappraisal of our object of study in which we incorporate the temporal flow of consciousness as another fundamental property that needs to be explained. This calls for a dynamic view in which a train of conscious states (the flow) would be captured as successions of neuronal meta-assemblies, each with a particular relaxation time, followed by phase transitions, which determine the time of emergence, dominance, and dissolution of a state that leads to another perceptual cycle (Melloni & Singer 2010; Varela 1999). In this framework, the rate-limiting factor for the formation of a new meta-assembly would correspond to the time needed to establish stable phase relations; while the different time constants promoting stability vs. change may be implemented by different oscillatory frequency bands, in addition to the intrinsic time window of integration of a given area (Chaudhuri et al. 2014).

In summary, much remains to be discovered about consciousness and its neural correlates, but significant progress has already been made since the seminal paper by Crick & Koch (1990) that got the field going about twenty-five years ago. Victor Lamme's experimental work and theoretical proposals on the role of feedback connections and reentrant activity in conscious perception have been central to bringing us closer to an understanding of the neural processes that allow us to "see". His paper in this volume contains an erudite review

---

4 Similarly, higher areas have larger receptive fields than lower areas, allowing integration over larger regions of space, and are often more broadly tuned (i.e., allow for more variability in the stimulus, e.g., different views of the same object). This resonates well with psychophysical evidence that hysteresis is spatially less specific and more broadly tuned than adaptation (Gepshtein & Kubovy 2005; Knapen et al. 2009).

of the present knowledge against a background of thought provoking hypotheses, e.g., that the function of consciousness is to solve difficult perceptual problems. In Lamme's view, consciousness is there to create, while unconscious processes are there to utilise. In close analogy to any creative process, consciousness in Lamme's framework is slow and takes time and resources to develop. In a way, his proposal is that it is all about distance, or time. This is a powerful intuition, and an idea worth exploring, yet its contribution does not end there—more than that, it serves as a reminder of a central characteristic of consciousness that is not yet fully explored, namely that conscious experience unfolds at a characteristic spatio-temporal scale, and that it is this flow in space/time that brings the strong sense of experiential stability and continuity. The interwoven temporal scales of the flow of consciousness that bring about the "unity of experience" remain the next challenge, and maybe the one that will finally unlock the mystery of consciousness.

3.1 Acknowledgements

References

Almeida, J., Pajtas, P. E., Mahon, B. Z., Nakayama, K. & Caramazza, A. (2013). Affect of the unconscious: Visually suppressed angry faces modulate our decisions. *Cognitive, Affective, & Behavioral Neuroscience*, *13* (1), 94-101. 10.3758/s13415-012-0133-7

Aru, J., Axmacher, N., Do Lam, A. T., Fell, J., Elger, C. E., Singer, W. & Melloni, L. (2012). Local category-specific gamma band responses in the visual cortex do not reflect conscious perception. *The Journal of Neuroscience*, *32* (43), 14909-14914. 10.1523/JNEUROSCI.2051-12.2012

Baars, B. J. (2002). The conscious access hypothesis: Origins and recent evidence. *Trends in Cognitive Sciences*, *6* (1), 47-52. 10.1016/S1364-6613(00)01819-2

Bekinschtein, T. A., Dehaene, S., Rohaut, B., Tadel, F., Cohen, L. & Naccache, L. (2009). Neural signature of the conscious processing of auditory regularities. *Proceedings of the National Academy of Sciences of the United States of America*, *106* (5), 1672-1677. 10.1073/pnas.0809667106

Burr, D. & Ross, J. (2008). A visual sense of number. *Current Biology*, *18* (6), 425-428. 10.1016/j.cub.2008.02.052

Chaudhuri, R., Bernacchia, A. & Wang, X. J. (2014). A diversity of localized timescales in network activity. *Elife*, *3*, e01239. 10.7554/eLife.01239.

Chopin, A. & Mamassian, P. (2012). Predictive properties of visual adaptation. *Current Biology*, *22* (7), 622-626. 10.1016/j.cub.2012.02.021

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36* (3), 181-204. 10.1017/S0140525X12000477

Clifford, C. W. (2012). Visual perception: knowing what to expect. *Current Biology*, *22* (7), 223-225. 10.1016/j.cub.2012.02.019

Cohen, M. X., van Gaal, S., Ridderinkhof, K. R. & Lamme, V. A. (2009). Unconscious errors enhance prefrontal-occipital oscillatory synchrony. *Frontiers in Human Neuroscience*, *3* (54), 1-12. 10.3389/neuro.09.054.2009

Crick, F. & Koch, C. (1990). Towards a neurobiological theory of consciousness. *Seminars in the Neurosciences*, *2*, 263-275.

de Gardelle, V., Charles, L. & Kouider, S. (2011). Perceptual awareness and categorical representation of faces: Evidence from masked priming. *Consciousness and Cognition*, *20* (4), 1272-1281. 10.1016/j.concog.2011.02.001

Dehaene, S. (2014). *Consciousness and the brain: Deciphering how the brain codes our thoughts.* New York, NY: Viking Penguin.

Dehaene, S. & Changeux, J. P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron, 70* (2), 200-227. 10.1016/j.neuron.2011.03.018

Dehaene, S., Naccache, L., Le Clec, H. G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., van de Moortele, P.-F. & Le Bihan, D. (1998). Imaging unconscious semantic priming. *Nature, 395* (6702), 597-600. 10.1038/26967

Del Zotto, M., Deiber, M. P., Legrand, L. B., De Gelder, B. & Pegna, A. J. (2013). Emotional expressions modulate low alpha and beta oscillations in a cortically blind patient. *International Journal of Psychophysiology, 90* (3), 358-362. 10.1016/j.ijpsycho.2013.10.007

Di Lollo, V., Enns, J. T. & Rensink, R. A. (2000). Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General, 129* (4), 481-507.

Edelman, G. M. & Tononi, G. (2000). *A universe of consciousness: How matter becomes imagination.* New York, NY: Basic Books.

Faugeras, F., Rohaut, B., Weiss, N., Bekinschtein, T. A., Galanaud, D., Puybasset, L., Bolgert, F., Sergent, C., Cohen, L., Dehaene, S. & Naccache, L. (2011). Probing consciousness with event-related potentials in the vegetative state. *Neurology, 77* (3), 264-268. 10.1212/WNL.0b013e3182217ee8

Fischer, J. & Whitney, D. (2014). Serial dependence in visual perception. *Nature Neuroscience, 17* (5), 738-743. 10.1038/nn.3689

Friston, K. J. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience, 11* (2), 127-138. 10.1038/nrn2787

Friston, K., Adams, R. A., Perrinet, L. & Breakspear, M. (2012). Perceptions as hypotheses: Saccades as experiments. *Frontiers in Psychology, 3,* 1-20. 10.3389/fpsyg.2012.00151

Fuster, J. M. (1973). Unit activity in prefrontal cortex during delayed-response performance: Neuronal correlates of transient memory. *Journal of Neurophysiology, 36* (1), 61-78.

Gaillard, R., Dehaene, S., Adam, C., Clémenceau, S., Hasboun, D., Baulac, M., Cohen, L. & Naccache, L. (2009). Converging intracranial markers of conscious access. *PLoS Biology, 7* (3), e61. 10.1371/journal.pbio.1000061

Gepshtein, S. & Kubovy, M. (2005). Stability and change in perception: spatial organization in temporal context. *Experimental Brain Research, 160* (4), 487-495. 10.1007/s00221-004-2038-3

Greenwald, A. G., Abrams, R. L., Naccache, L. & Dehaene, S. (2003). Long-term semantic memory versus contextual memory in unconscious number processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29* (2), 235-247. 10.1037/0278-7393.29.2.235

Hipp, J. F., Engel, A. K. & Siegel, M. (2011). Oscillatory synchronization in large-scale cortical networks predicts perception. *Neuron, 69* (2), 387-396. 10.1016/j.neuron.2010.12.027

Hock, H. S., Schoner, G. & Hochstein, S. (1996). Perceptual stability and the selective adaptation of perceived and unperceived motion directions. *Vision Research, 36* (20), 3311-3323. 10.1016/0042-6989(95)00277-4

Hohwy, J. (2013). *The predictive mind.* Oxford, UK: Oxford University Press.

Jardri, R. & Deneve, S. (2013). Circular inferences in schizophrenia. *Brain, 136* (11), 3227-3241. 10.1093/brain/awt257

Jenkins, R., Beaver, J. D. & Calder, A. J. (2006). I thought you were looking at me: direction-specific aftereffects in gaze perception. *Psychological Science, 17* (6), 506-513. 10.1111/j.1467-9280.2006.01736.x

Jones, P. D. & Holding, D. H. (1975). Extremely long-term persistence of the McCollough effect. *Journal of Experimental Psychology: Human Perception and Performance, 1* (4), 323-327. 10.1037/0096-1523.1.4.323

King, J. R., Sitt, J. D., Faugeras, F., Rohaut, B., El Karoui, I., Cohen, L. & Dehaene, S. (2013). Information sharing in the brain indexes consciousness in noncommunicative patients. *Current Biology, 23* (19), 1914-1919. 10.1016/j.cub.2013.07.075

Kleinschmidt, A., Büchel, C., Hutton, C., Friston, K. J. & Frackowiak, R. S. (2002). The neural structures expressing perceptual hysteresis in visual letter recognition. *Neuron, 34* (4), 659-666. 10.1016/S0896-6273(02)00694-3

Knapen, T., Brascamp, J., Adams, W. J. & Graf, E. W. (2009). The spatial scale of perceptual memory in ambiguous figure perception. *Journal of Vision, 9* (13), 11-12. 10.1167/9.13.16

Koch, C. (2004). *The quest for consciousness: A neurobiological approach.* Englewood, CO: Roberts & Company.

Kouider, S. & Dehaene, S. (2007). Levels of processing during non-conscious perception: A critical review of visual masking. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *362* (1481), 857-875. 10.1098/rstb.2007.2093

Lamme, V. (2015). The crack of dawn: Perceptual functions and neural mechanisms that mark the transition from unconscious processing to conscious vision. In T. Metzinger & J. M. Windt (Eds.) *Open MIND*. Frankfurt a. M., GER: MIND Group.

Lin, Z. & He, S. (2009). Seeing the invisible: The scope and limits of unconscious processing in binocular rivalry. *Progress in Neurobiology*, *87* (4), 195-211. 10.1016/j.pneurobio.2008.09.002

Mayer, A., Schwiedrzik, C. M., Singer, W. & Melloni, L. (in preparation). *Expectations sparsify networks for letter recognition.*

Melloni, L., Molina, C., Pena, M., Torres, D., Singer, W. & Rodriguez, E. (2007). Synchronization of neural activity across cortical areas correlates with conscious perception. *The Journal of Neuroscience*, *27* (11), 2858-2865. 10.1523/JNEUROSCI.4623-06.2007

Melloni, L., Schwiedrzik, CM., Rodriguez, E. & Singer, W. (2009). (Micro)Saccades, corollary activity and cortical oscillations. *Trends in Cognitive Sciences*, *13* (6), 239-245. 10.1016/j.tics.2009.03.007

Melloni, L., Schwiedrzik, C. M., Muller, N., Rodriguez, E. & Singer, W. (2011). Expectations change the signatures and timing of electrophysiological correlates of perceptual awareness. *The Journal of Neuroscience*, *31* (4), 1386-1396. 10.1523/JNEUROSCI.4570-10.2011

Melloni, L. & Rodriguez, E. (2007). Non-perceived stimuli elicit local but not large-scale neural synchrony. *Perception*, *36* (ECVP Abstract Supplement)

Melloni, L. & Singer, W. (2010). Distinct characteristics of conscious experience are met by large-scale neuronal synchronization. In E. K. Perry, D. Collerton, F. E. N. LeBeau & H. Ashton (Eds.) *New horizons in the neuroscience of consciousness* (pp. 17-28). Amsterdam, NL: John Benjamins.

Moca, V. V., Tincas, I., Melloni, L. & Muresan, R. C. (2011). Visual exploration and object recognition by lattice deformation. *PLoS One*, *6* (7), e22831. 10.1371/journal.pone.0022831

Mudrik, L., Faivre, N. & Koch, C. (2014). Information integration without awareness. *Trends in Cognitive Sciences*, *18* (8), 414-421. 10.1016/j.tics.2014.04

Murray, S. O., Schrater, P. & Kersten, D. (2004). Perceptual grouping and the interactions between visual cortical areas. *Neural Networks*, *17* (5-6), 695-705. 10.1016/j.neunet.2004.03.010

Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review*, *83* (4), 435-450. 10.2307/2183914

Pellicano, E. & Burr, D. (2012). When the world becomes 'too real': A Bayesian explanation of autistic perception. *Trends in Cognitive Sciences*, *16* (10), 504-510. 10.1016/j.tics.2012.08.009

Purkinje, J. E. (1820). Beiträge zur näheren Kenntnis des Schwindels aus heautognostischen Daten. *Medicinische Jahrbücher des kaiserl.-königl. österreichischen Staates*, *6*, 79-125.

Schroeder, C. E., Wilson, D. A., Radman, T., Scharfman, H. & Lakatos, P. (2010). Dynamics of active sensing and perceptual selection. *Current Opinion in Neurobiology*, *20* (2), 172-176. 10.1016/j.conb.2010.02.010.

Schwartz, O., Hsu, A. & Dayan, P. (2007). Space and time in visual context. *Nature Reviews Neuroscience*, *8* (7), 522-535. 10.1038/nrn2155

Schwiedrzik, C. M., Singer, W. & Melloni, L. (2009). Sensitivity and perceptual awareness increase with practice in metacontrast masking. *Journal of Vision*, *9* (10), 11-18. 10.1167/9.10.18

——— (2011). Subjective and objective learning effects dissociate in space and in time. *Proceedings of the National Academy of Sciences of the United States of America*, *108* (11), 4506-4511. 10.1073/pnas.1009147108

——— (2011). Subjective and objective learning effects dissociate in space and in time. *Proceedings of the National Academy of Sciences of the United States of America*, *108* (11), 4506-4511. 10.1073/pnas.1009147108

Schwiedrzik, C. M., Ruff, C. C., Lazar, A., Leitner, F. C., Singer, W. & Melloni, L. (2014). Untangling perceptual memory: Hysteresis and adaptation map into separate cortical networks. *Cerebral Cortex*, *24* (5), 1152-1164. 10.1093/cercor/bhs396

Seth, A. K., Suzuki, K. & Critchley, H. D. (2011). An interoceptive predictive coding model of conscious presence. *Front in Psychology*, *2* (395), 1-16. 10.3389/fpsyg.2011.00395

Snyder, J., Schwiedrzik, C. M., Vitela, D. & Melloni, L. (forthcoming). *How previous experience shapes perception across sensory modalities.*

Thompson, E. & Varela, F. J. (2001). Radical embodiment: Neural dynamics and consciousness. *Trends in Cognitive Sciences*, *5* (10), 418-425. 10.1016/S1364-6613(00)01750-2

Tononi, G. (2004). An information integration theory of consciousness. *BMC Neuroscience*, *5* (45), 1-22. 10.1186/1471-2202-5-42

Tononi, G. & Koch, C. (2008). The neural correlates of consciousness: An update. *Annals of the New York Academy of Sciences, 1124* (1), 239-261. [10.1196/annals.1440.004](#)

Treisman, M. (1984). A theory of criterion setting: An alternative to the attention band and response ratio hypotheses in magnitude estimation and cross-modality matching. *Journal of Experimental Psychology: General, 113* (3), 443-463. [10.1037/0096-3445.113.3.443](#)

van Leeuwen, T. M., Wibral, M., Sauer, A., Uhlhaas, P., Singer, W. & Melloni, L. (2013). Neural synchronization during bottom-up and top-down visual processing in grapheme- color synesthetes and schizophrenia patients. *Poster at the 43rd Meeting of the Society for Neuroscience (SfN), San Diego, USA.*

Varela, F. (1999). The specious present: A neurophenomenology of time consciousness. In J. Petitot, J. Varela, J.-M. Roy & B. Pachoud (Eds.) *Naturalizing phenomenology* (pp. 266-314). Stanford, CA: Stanford University Press.

Varela, F., Lachaux, J. P., Rodriguez, E. & Martinerie, J. (2001). The brainweb: Phase synchronization and large-scale integration. *Nature Reviews Neuroscience, 2* (4), 229-239. [10.1038/35067550](#)

von Foerster, H. (1984). On constructing a reality. In P. Watzlawick (Ed.) *The invented reality: How do we know what we believe we know* (pp. 41-62). New York: W.W.Norton & Co.

von Helmholtz, H. (1962). *Handbuch der physiologischen Optik.* New York, NY: Dover.

Webster, M. A., Kaping, D., Mizokami, Y. & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature, 428* (6982), 557-561. [10.1038/nature02420](#)

Predictive Coding Is Unconscious, so that Consciousness Happens *Now*

A Reply to Lucia Melloni

Victor Lamme

Conscious percepts depend strongly on past events. Expectations, primes, and prior experiences all shape the percept we have at any moment in time. Yet does this imply that conscious experience should be viewed as extended in time—as "flowing"—instead of as just happening now?

Author

Victor Lamme
Victorlamme@gmail.com
Universiteit van Amsterdam
Amsterdam, Netherlands

Commentator

Lucia Melloni
lucia.melloni@brain.mpg.de
Max Planck Institute for Brain Research
Frankfurt a. M., Germany

Editors

Thomas Metzinger
metzinger@uni-mainz.de
Johannes Gutenberg-Universität
Mainz, Germany

Jennifer M. Windt
jennifer.windt@monash.edu
Monash University
Melbourne, Australia

1 To infer or to integrate, that is the question

In her commentary, Lucia Melloni argues that consciousness unfolds in time: there is a stream of consciousness. What I see now is intricately linked to what I have seen before. And what I see now is what I expect to see—much along the lines of predictive coding. A full understanding of consciousness should not neglect this point. There is even a stronger claim that somehow the process of inference over time is crucial to understanding consciousness.

I appreciate the boldness of linking the framework of Bayesian predictive coding to specific stages in the process of generating consciousness:

> One promising framework within which the influence of previous experience can be understood is the Bayesian framework. When applied to perception, each mathematically-formulated ingredient of this framework can be assigned a percep-

tual counterpart, with previous experience referring to the prior, the current moment referring to the likelihood, unconscious inference referring to Bayes rule (which combines the prior with the likelihood in an optimal way), and the result—our perception—referring to (the peak of) the posterior distribution. (Melloni this collection, p. 4)

To my knowledge, this is the first time this has been so explicitly laid out—writers on predictive coding thus far have always stayed a little vague on where exactly consciousness sits in the Bayesian framework.

Yet at the same time, there is the suggestion of long temporal range integration being the key ingredient of consciousness:

Event-objects of the conscious mind thus per definition unfold in time and we also act in time: we make use of current and previous input to figure out the most appropriate response predicting their consequences. There is thus a continuum of interdependencies along the time dimension whereby every past moment is integrated with the present and projected into the future, giving rise to the flow of consciousness. The same way we have been thinking about the integration of multiple sources of information within a given moment of time, such as multiple features of a single object, there is thus integration across time. (Melloni this collection, pp. 7-8)

This makes intuitive sense, particularly in the case of moving objects, such as the tennis ball Melloni uses as an example. Indeed it is hard —if not impossible—to pinpoint the exact *now* of conscious experience of such a ball.[1]

Yet the two points seem contradictory. In the Bayesian predictive coding framework, consciousness is the *result* of the unconscious inferential processes. Previous knowledge and

experience (the priors) play an important role, but they are combined with current input to produce the posterior, which is conscious sensation. In the second account, however, consciousness seems to be something that is stretched out over time, so that both prior and posterior are smelted into a "flow" of consciousness. I find it hard to reconcile these two views.

2 The latency of visual consciousness is variable

Melloni discusses some impressive experiments that show the crucial importance of prior information and expectation in shaping or simply altering conscious experience (and her example, figure 1, is enlightening and flattering at the same time). In all these cases, however, consciousness is portrayed as the *outcome* or *result* of an otherwise unconscious inferential process. The result may come earlier or later, as in the experiment on letter priming that Melloni describes, resulting in earlier (200ms) or later (300ms) electrophysiological correlates of conscious recognition depending on the presence or absence of appropriate priors. Further experiments are discussed, showing that neural correlates of consciousness may shift (neural) location, depending on expectation and priors. Yet still, the end result —consciousness—occurs at the end of a cascade of neural operations. Consciousness, in this account, may occur at variable moments and locations, but *moments* they are.

These results complement earlier findings that the latency of recurrent processing—and hence the emergence of a conscious sensation —may vary. Super et al. (2001) showed that degrading stimulus quality may increase the latency of recurrent signals to V1 in the monkey visual cortex (see figure 5c of Supèr et al. 2001), and that this affects the latency of behavioral responses of animals that are consciously reporting the presence or absence of the stimuli. Latency of recurrent signals may also vary spontaneously between trials, which correlates with the latency of memory-guided —but not reflexive—saccades to the targets

---

1 Although some have argued that consciousness unfolds in time as a succession of static frames, more or less like the single frames of a movie—even at specific frequencies, namely 10Hz and 40Hz (Van Rullen & Koch 2003).

that elicit these recurrent signals (Supèr et al. 2004). In humans, the latencies of electrophysiological correlates of recurrent processing also vary, either spontaneously or depending on stimulus properties (Jolij et al. 2011), or depending on the IQ of the subject (Jolij et al. 2007). Likewise, this has consequences for the latency of conscious sensations. The Jolij 2011 study, for example, found that variations in the latency of recurrent EEG signals covary with variations in subjective simultaneity of the stimuli evoking these signals.[2] These results invariably imply that consciousness arises at a particular *moment* in time. That moment may vary from stimulus to stimulus, from trial to trial, from person to person, from prior to prior. But nothing is flowing or stretched out over time.

3 Consciousness is not streaming, but taking snapshots

One may argue that these findings are all obtained with stimuli that are presented *de novo,* using the classic stimulus-onset paradigms. In normal vision, things don't suddenly appear out of nowhere. Or do they? We naturally make about three saccadic eye movements per second, and each time the eye lands on a "new" scene which is—from a retinotopic point of view—radically different from the previous one. In between, we are blind due to saccadic suppression. Moreover, little information seems to be transferred from one view to the next, although some (attended) neural representations seem to be *remapped* across saccades (see Bays & Husain 2007, for an overview of trans-saccadic memory and neural remapping). Such a remapping may allow for a more efficient saccade from one object to the next, when both were already present before the first saccade was made. The predictive coding framework seems to re-emerge in this context: objects that were present or attended on a first fixa-

tion form a sort of prior for the representation that is built during the second fixation (which may then arise more rapidly).

Melloni further claims that previous experience has different effects on what is perceived now depending on the temporal interval between prior and current experience. Bistable percepts show hysteresis or adaptation depending on these temporal intervals, or depending on whether the previous experience was conscious or not. But again, I fail to see how these findings support the idea that consciousness is stretched out over time instead of just happening *now.*[3]

So I appreciate the importance of the predictive coding framework. Previous experience plays a very important role in the conscious sensations we have, and the why and how of this is extremely important for fully understanding vision. But these contributions are unconscious. Consciousness happens now, and its neural correlates are likewise limited in time. Consciousness of the past we call memory.

Acknowledgements

---

[2] For this reason, I don't quite understand why Melloni suggests that I am claiming that consciousness arises at a particular and *fixed* moment in time. My claim is only that it comes *after* feedforward processing, and as soon as recurrent processing emerges—which may vary.

[3] Of course there are visual percepts that are more or less defined by their temporal sequence, the prime example being motion. But this does not imply that the perception of motion is flowing. The first thing the brain does in detecting motion is to convert the flow of motion into a discrete and momentary signal, indistinguishable from how the brain represents other features such as orientation, color, or shape. As a result we see motion now, and instantaneously, which is also crucial for our survival: perceiving something moving in the shadows of a bush (e.g., a snake) needs to be translated into action as soon as possible (e.g., running away). No time for any flow there.

References

Bays, P. M. & Husain, M. (2007). Spatial remapping of the visual world across saccades. *Neuroreport*, *18* (12), 1207-1213. 10.1097/WNR.0b013e328244e6c3

Jolij, J., Huisman, D., Scholte, H. S., Hamel, R., Kemner, C. & Lamme, V. A. F. (2007). Processing speed in recurrent visual networks correlates with general intelligence. *Neuroreport*, *18* (1), 39-43. 10.1097/01.wnr.0000236863.46952.a6

Jolij, J., Scholte, H. S., Van Gaal, S., Hodgson, T. L. & Lamme, V. A. F. (2011). Act quickly, decide later: Long-latency visual processing underlies perceptual decisions but not reflexive behavior. *Journal of Cognitive Neuroscience*, *23* (12), 3734-3745. 10.1162/jocn_a_00034

Melloni, L. (2015). Consciousness as inference in time. In T. Metzinger & J. M. Windt (Eds.) *Open MIND*. Frankfurt a. M., GER: MIND Group.

Supèr, H., Spekreijse, H. & Lamme, V. A. F. (2001). Two distinct modes of sensory processing observed in monkey primary visual cortex (V1). *Nature Neuroscience*, *4*, 304-310. 10.1038/85170

Supèr, H., Van der Togt, C., Spekreijse, H. & Lamme, V. A .F. (2004). Correspondence of presaccadic activity in the monkey primary visual cortex with saccadic eye movements. *Proceedings of the National Academy of Sciences of the United States of America*, *101* (9), 3230-3235. 10.1073/pnas.0400433101

Van Rullen, R. & Koch, C. (2003). Is perception discrete or continuous? *Trends in Cognitive Sciences*, *7* (5), 207-213. 10.1016/S1364-6613(03)00095-0