MODULATIONS OF PRIMARY VISUAL CORTEX ACTIVITY REPRESENTING ATTENTIVE AND CONSCIOUS SCENE PERCEPTION

Victor A.F. Lamme [1,2] and Henk Spekreijse [1]

[1] Graduate School of Neurosciences, Dept. Visual System Analysis, AMC, University of Amsterdam, P.O. Box 12011, 1100 AA Amsterdam, The Netherlands, [2] The Netherlands Ophthalmic Research Institute, AMC, Meibergdreef 47, 1105 BA Amsterdam, The Netherlands

TABLE OF CONTENTS

1. ABSTRACT

In the visual cortex, information is transferred from one area to the next by means of feedforward connections. These connections shape the receptive field properties of neurons in subsequent visual areas. Horizontal and feedback connections modulate this neuronal activity, resulting in the phenomenon of contextual modulation. In area V1, where receptive field properties reflect only low level processing, contextual modulation can be observed that represents fully evaluated perceptual saliency of the features within the receptive field. Here, we discuss to what extent these modulations are related to high level visual processes like perceptual organization, attention and visual awareness. Contextual modulation appears to reflects a process very distinct from receptive field based processing. This process seems to integrate information from distant areas in visual cortex to neurophysiologically 'highlight' those neurons that represent image elements or features of objects that stand out perceptually. Moreover, similar modulations are observed in relation to whether objects are attended to or not. Finally, these modulations are only present when subjects are aware of the visual input.

2. INTRODUCTION

2.1. Visual areas are defined by receptive field tuning properties

The receptive field of a neuron in visual cortex is the part of the visual field from which action potential responses can be elicited by presenting a stimulus. This stimulus has to meet a number of requirements for the cell to respond; the receptive field is 'tuned' to particular features. For example, cells in primary visual cortex (V1) respond better to some orientations of luminance contrast than to others (1,2). Traditionally, the function of a visual area is derived from the set of features to which the neurons in the area are tuned. MT is called a 'motion' area because many of its cells are tuned to direction, speed, or other aspects of motion (3-7), V4 is a 'color' or a 'form' area because cells are tuned to certain wavelengths (8,9) or elementary shapes (10,11), etc. This view is further corroborated by the fact that lesions of these areas typically cause deficits that are related to the processing function we attribute to that area by means of its receptive field tuning properties (7,12-17). Thus we arrive at a view where visual processing is subdivided in specific modules, each solving a particular subproblem of vision (18,19). Although this view is much disputed in its strictness, it has nevertheless profoundly pervaded our thinking of the roles of cortical areas in vision.

2.2. Combining the distributed information

When the visual field is analyzed by limited receptive fields within functionally separate modules, extensive and dynamic interactions are required to combine the distributed information. Anatomical connections provide the framework for those interactions (20). Within each cortical area, horizontal interactions integrate information from separate parts of the visual field (21,22). Between areas, information is transferred in a feedforward fashion from low level areas to higher level ones. But in addition, feedback connections transfer information in the reverse direction (23). When going upstream through the

hierarchy of visual areas, receptive fields obtain increasingly complex tuning properties and rapidly increase in size (7). Receptive field properties thus strongly reflect the convergent-divergent feedforward cascade of information processing. Feedback connections are highly diverging, but their influence on cells in lower areas is not reflected by the small receptive fields of an area such as V1 (23). Also horizontal connections within V1 spread over much larger distances than the size of receptive fields would neccessitate (22,24). This indicates that the receptive field properties of neurons do not take into account the interactions that are mediated by a set of connections that numerically even outweighs the set of feedforward connections.

There are other response properties of neurons, however, that do take advantage of these sets of connections. Once a receptive field contains some stimulus, the response to this stimulus may be modulated by surrounding stimuli. A key feature of this phenomenon is that the modulating stimuli do not evoke a response when presented alone; they are outside the 'classical' receptive field. The early experiments, in area 17 of anesthetised cats, typically used bars or gratings to stimulate both the receptive field and its surround. Modulatory effects could be evoked from large distances, but were strongest at short distances. Both fascilitory and inhibitory effects were found, that could be either non-specific or any combination of orientation and direction of motion specific (25-30). Also in areas beyond V1, modulation from outside the receptive field was reported (4,30,31). The phenomenon thus seems to be a general property of visual cortical cells.

In this paper, we will review the phenomenon of contextual modulation, and focus on primary visual cortex of the monkey in particular. It will be shown that contextual modulation indeed reflects widespread interactions within and between cortical areas, interactions that are related to high level processes like perceptual organisation, attention and visual awareness (figure 1). V1 is an interesting area in this respect. Its receptive fields are very small and their tuning properties are simple. On that account, activity reflecting perceptual interpretation of the scene as a whole is least expected here. On the other hand, it can be said that the area is at the top of the hierarchy in terms of feedback connections. If some feedback related activity would be present here that represents the convergence of information from all visual areas, this activity would be expected to clearly reflect a fully evolved perceptual interpretation. In V1 therefore, we can expect receptive field processing of a sort most detached from scene perception, while contextual modulation might be closely related to it.

3. PHENOMENOLOGY OF CONTEXTUAL MODULATION IN V1

3.1. Single versus multiple image elements
Line segments form ideal stimuli for V1 cells. V1 receptive fields are typically tuned for their orientation, direction of motion, disparity, size, contrast or color (2,19,32-34). V1 neurons are therefore viewed as encoding information about the line segments or contrast edges within their receptive fields. However, a single line segment on a blank background is a visual scene only rarely encountered. In natural scenes, line segments or contrast edges are combined with many others, forming edges, textures, object boundaries, the one occluded by the other etc. In those natural situations, the perceptual interpretation of a single line segment strongly depends on its context. So what happens to the V1 responses when the perceptual context of a line segment within the receptive field is manipulated?

Compare the single line segment of figure 2a with the multiple line segments of figure 2b. The isolated line segment strongly draws attention to itself while the set of line segments rather draws attention as a group, where each individual line segment is of much less importance. In V1, this is expressed in the neuronal responses by means of contextual modulation. If the line segment of figure 2a would fall on a V1 receptive field, and the surrounding line segments that are added in figure 2b would fall outside, the response of the neuron is much stronger in the first situation than in the second (35).

The reduced saliency of the center line segment in figure 2b can be alleviated by having its orientation differ from those of the surrounding elements, as in figure 2c. This results in a perceptual 'pop-out' of the center element (36,37). This (partial) restoration of perceptual saliency is again expressed by a modulation of V1 responses, which are larger in case of figure 2c than of 2b, although not as large as the response would be for a lone element (2a) (35). That this is not a purely local phenomenon, i.e. only governed by the orientation difference between center and surround stimuli, is illustrated by figure 2d. Here the center line segment is surrounded by line segments of different orientations, as in figure 2c. There is no perceptual pop-out, however, because the orientation difference between center and surround is not different from orientation differences amongst elements in general (37,38,39). Likewise, responses of V1 neurons to stimuli like those of figure 2d are not different from the responses to figure 2b stimuli (35,40).

Another factor that makes line segments segregate from a background of randomly oriented line segments is when they group into elongated chains (Fig. 2e). Such groupings depend on the relative alignment of the line segments, and important factors are colinearity, relative distance, angle, and axial offset (40-42). Remarkably, these very same factors influence contextual modulation in V1 neurons. Line segments that are flanked by collinear ones elicit larger responses (40,43).

3.2. Boundaries, surfaces and figure-ground segregation
In figure 3a, the more or less one dimensional chain of lines of figure 2e has been extended to two dimensions. The grouped line segments, that now group on the basis of orientation similarity, segregate from a background of line segments of another orientation. The stimulus is richer in some aspect, however; we can distinguish a boundary and a surface of the segregating
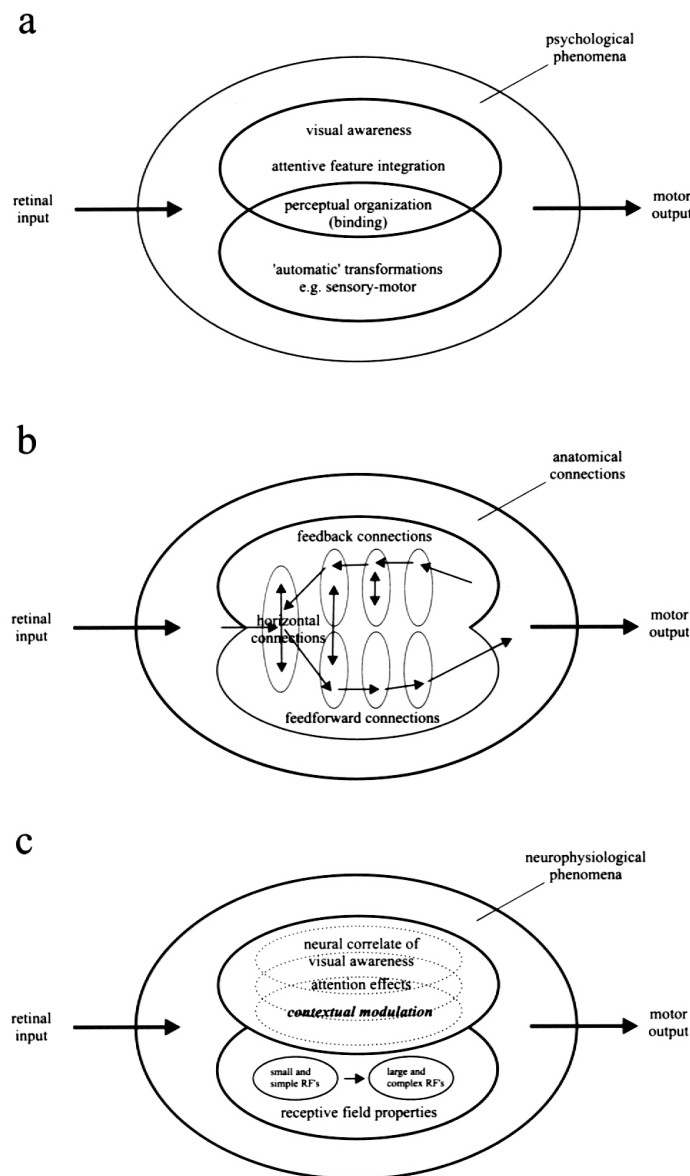
a

b

c

Figure 1: Visual processing at the psychological, neuro-anatomical, and neurophysiological level. (*a*) The psychological level: More or less automatic processes transform visual input into motor output. An example would be the grasping of an object that suddenly falls from a table. This does require some form of perceptual organization or binding, that occurs automatically, without attention. Other transformations require selective attention, such as in visual search tasks that do not occur in parallel. Perceptual organization is therefore conceptualized at the site of overlap between pre-attentive and attentive processes. Visual awareness is a phenomenon that belongs exclusively to the attentive domain (and maybe to some other domains that are left undefined here). (*b*) At the neuroanatomical level, feedforward connections go from one visual area (the vertical ellipses) to the next. Feedback connections go in the reverse direction. Horizontal connections link areas at the same level, or connect distant retinotopic (or whatever topology exists within the area) sites within the same area. (*c*) Receptive field tuning properties evolve very fast, even in the highest visual areas, and mainly reflect feedforward, automatic processing. For example, areas of parietal cortex transform visual input, framed in a retinal coordinate system, into information framed in head, body and object centered coordinates (refs 86,87). Thus transformed, the visual input could effectively guide eye, arm or body movements to grasp objects (that suddenly fall from a table).Contextual modulation, together with neural effects of selective attention, belongs to a different class of neurophysiological phenomena. As will be shown in this paper, this class will harbor the neural substrates of perceptual organization, attention and visual awareness. The figures can roughly be overlaid to link psychological, anatomical and neurophysiological levels with each other.
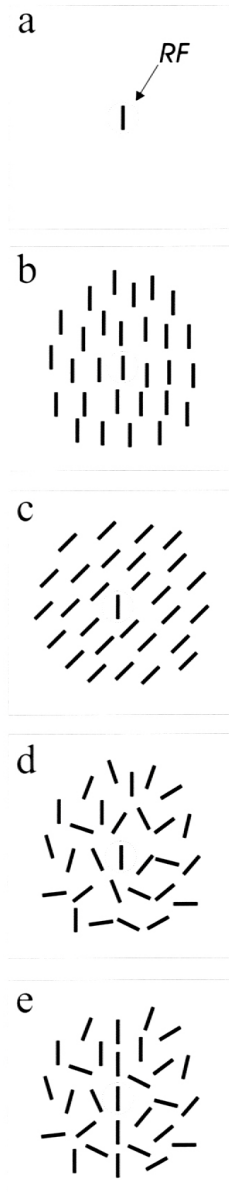
Figure 2: Context of a line segment changes its perceptual saliency. (*a*) A lone line segment is perceptually most salient. (*b*) The same line segment, now embedded in similar ones, draws much less attention as an individual. (*c*) The line segment may 'pop-out' when its orientation differs from that of the surrounding line segments, restoring its perceptual saliency to some extent. (*d*) Global stimulus aspects are taken into consideration, since a similar local orientation difference does not produce 'pop-out' when surrounding elements all have different orientations. (*e*) Also perceptual grouping by co-axial alignment may cause line segments to segregate, increasing their perceptual saliency. When a V1 neuron is stimulated by presenting these displays, so that its receptive field only covers the same center line segment in all cases, contextual modulation of its responses signals the perceptual saliency of the center line segment.

figure. Where line segments of the one orientation juxtapose line segments of the orthogonal orientation we observe a sharp boundary between the square figure and the background. The line segments within this boundary form a figure surface, that is perceived as if lying in front of the background, that is assumed to continue behind it. The boundary is asymmetrical in that it 'belongs' to the figure and not to the background.

To analyse how contextual modulation modifies responses depending on whether the receptive field contained line segments belonging to either the boundary or surface of the figure or to the background, complementary stimulus pairs need to be used. In this way, receptive field stimulation can be kept identical in all cases (see ref 44), so that contextual modulation can be studied separately from receptive field processing as such. Figure 3b shows population responses from neurons in V1 of the awake monkey for three different receptive field contexts; background, figure boundary or figure surface. The response to line segments that belong to the background consists of a transient followed by a slow decay. When the same neurons respond to identical line segments that now compose the boundary between figure and ground, the same initial transient is observed. However, from about 80 milliseconds past stimulus onset, the response is higher than the response to background elements. The same initial response is also obtained when the receptive field is at the center of the figure, and now at about 100 milliseconds figure and background responses start to diverge. In figure 3c, these data, combined with data from 12 other positions of the receptive field relative to figure and ground, are represented in three dimensions. The different positions of receptive field relative to figure and ground are represented on the front horizontal axis. Contextual modulation can now be observed for 9 positions of the receptive field within the boundaries between figure and ground. Initially, at 80 milliseconds, contextual modulation is only present for the boundaries between figure and ground. But from about 100 milliseconds on, all elements within the figure elicit a stronger response than the same elements of the background. It is as if a neural image of the figure is stamped out of the neural image of the background, closely reflecting our figure-ground percept (45).

These data illustrate an important point about contextual modulation that is not observed in the results discussed in the previous section. In the previous section (figure 2) we saw contextual modulation reflecting local grouping and segregation criteria. But using stimuli like in figure 3a, we see that these local criteria can be overridden: At the center of the figure, the line segments are surrounded by similar ones, but nevertheless contextual modulation is the same as close to the edge, where elements are flanked by orthogonal ones (although the modulation at the edge occurs at an eralier latency). At the immediate outside of the boundary between figure and ground elements are flanked by orthogonal ones, but contextual modulation is absent (that is, responses are identical to background positions further away). Apparently, contextual modulation is not limited to reflecting local discontinuities, or differences between
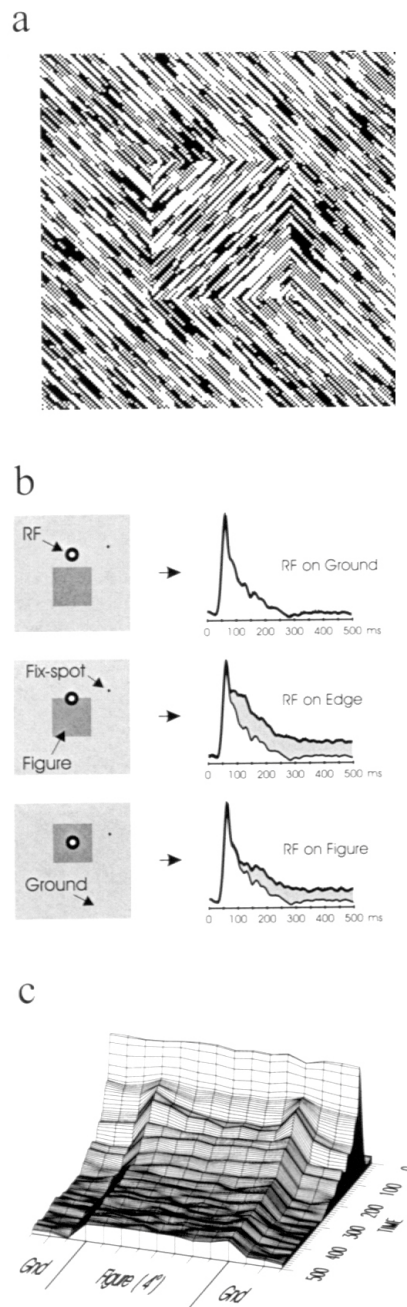
235

a

b

c

Figure 3: The neural correlate of figure-ground segregation in V1. (*a*) Line segments of similar orientation perceptually group together and segregate from line segments of another orientation. The center square is considered a textured figure on a textured background, that perceptually seems to continue behind it. The boundary between figure and ground belongs to the figure surface. (*b*) Responses in V1 to the figure-ground display are larger when the receptive field (open circle; RF) of a neuron is on the boundary or on the surface of the figure than when it is on the background (while receptive field stimulation is left identical in these three cases, see ref 44). Note that contextual modulation (gray shading) only develops after 80 to 100 milliseconds after stimulus onset; the initial transient is identical in all three cases, showing that only what happens within the receptive field determines the responses up to 80 milliseconds. (*c*) Responses in V1 with the receptive field at 15 different positions relative to figure and ground, such that the contextual modulation is scanned across a line passing over and through the figure. The 15 positions are on the x-axis (in front), time is on the y-axis (side) and response strength on the vertical axis. Responses are identical up to about 80 milliseconds after stimulus onset (note the horizontal wave a the back of the plot, which is at 50 ms). Then, responses are 'highlighted' at the boundary between figure and ground first. This is followed by an equal response enhancement for all positions of the receptive field within the figure, compared to responses for positions of the receptive field on the background.

receptive field center and surround stimuli. Contextual modulation in these experiments reflects the figure-ground relationships of the surfaces in the scene. An alternative interpretation would be that the neurons representing the perceptually most salient elements of the scene, in this case the whole figure, are highlighted relative to neurons representing less important elements.

These kinds of effects can be elicited by figure-ground displays defined by a variety of cues, like differences in orientation, direction of motion, disparity, colour or luminance. At the population response level, contextual modulation is of the same magnitude for figures defined by these various cues, while also at the single unit level some cells show complete cue-invariance (44,46). Surprisingly, figure-ground related contextual modulation bears no relation to the receptive field properties of the neurons recorded from. For example, modulation for a motion defined figure can be recorded in cells without direction selectivity (44). Also, when cues are combined, modulation is not additive, but identical to the one-cue alone situation. This indicates that the modulation signals figure-ground relationships instead of feature specific differences; a figure is a figure, no matter how it is defined. It also shows that contextual modulation is mediated by mechanisms that are far removed from those that shape and tune the local receptive field.

Also when more complicated scenes are used, contextual modulation reflects the figure-ground arrangements of the surfaces in that scene. Particularly nice examples can be found in Zipser et al. (46) and Lee et al. (47), using surfaces with holes in them, multiple overlying surfaces etc. Under specific conditions, it can sometimes be observed that there is a stronger modulation right at the geometrical center of the figures in those scenes (47). It has been proposed that this plays a role in representing the medial axes of objects (42,47).

Modulations of neuronal activity in V1 can also be observed in relation to perceived brightness. The perceived brightness of a surface can be modulated by changing the brightness of surrounding surfaces. V1 neurons modulate their activity according to such changes in perceived brightness [48]. Perceived brightness can also be changed by presenting a surface, followed in time (for example 50 ms later) by the presentation of a surrounding surface of equal luminance, a phenomenon called metacontrast masking. At some time intervals between the presentation of the first surface and the second surrounding surface, the perceived brightness of the first surface is diminished, and at some intervals (and under ideal conditions) perception of the appearance of the first surface may totally vanish (49). The responses of V1 cells, whose receptive fields were centered on the center disk of a metacontrast stimulus, were modulated according to the reduced apparent brightness of that center disk. This modulation is not present in an early part of the response (<100 ms), only in a later one (50).

4. FUNCTIONS OF CONTEXTUAL MODULATION

4.1. Contextual modulation reflects perceptual saliency
In the above results, contextual modulation is shown to reflect three different aspects of image elements; segregation or pop-out, detectability, and apparent brightness. An element is said to pop-out when its detection is independent of the number of surrounding elements. For example in figure 2c, the time it would take to recognise that there is one element different from the rest would be equal for this amount of distractors or for a far larger amount. In other words, it does not take serial or attentive scrutiny of the individual elements of the scene to recognise that one is different from the rest. Surfaces may also pop-out, as in figure 3a, which is called texture segregation.

Detectability of image elements refers to a different psychophysical measure. By presenting an element at various low contrasts, its detection threshold (for instance 50% detection chance) can be determined. In this way, for example, it was found that detection threshold lowers, i.e. detectacility increases, when line segments are flanked by collinear ones (Fig. 2e) (40-42). It was also found that detectability is larger within an enclosed area of the scene than outside (51), which is reminiscent of the contextual modulation effects for surfaces (44,47). Apparent brightness is typically determined by comparing the brightness of a test surface with that of another one. In a two alternatives forced choice procedure, subjects are asked to tell which of two surfaces appears brightest (50).

These three measures have some essential differences. For example, one analyzes performance at detection thresholds, whereas the other two may be used at more normal levels of contrast. Also, whether elements that pop-out always have lower detection thresholds, or increased apparent brightness is questionable. They do share common ground however, which may be best described as a generalised notion of preceptual saliency. In this context the term is defined as a measure of how well an image element is capable of drawing attention to itself. Elements that pop-out, or have lower detection thresholds, or are brighter than others, all share a common property; they draw attention. This intuitive relation between the three measures is underscored by the findings of contextual modulation. When contexts either produce pop-out, or increase detectability, or increase apparent brightness, contextual modulation always results in an increase in the neuronal responses. When single unit responses are analyzed, cells might in some cases exhibit opposite effects, i.e. some proportion showing an increase while another proportion shows a decrease (40,52). On the multi-unit or population level, however, the outcome is always very clear; response amplitude reflects the proposed generalized notion of perceptual saliency.

4.2. Contextual modulation reflects perceptual grouping
The computations performed within the separate cortical areas have to be combined to produce a coherent output. The role of vision will be to segregate objects from each other, and to select particular ones for behavioural responses. Image elements and features of the objects that we encounter are processed by cells at different locations, and these have to be combined so that we are able to manipulate them (the binding problem (53,54)). To some extent this might be explained by the feedforward cascade
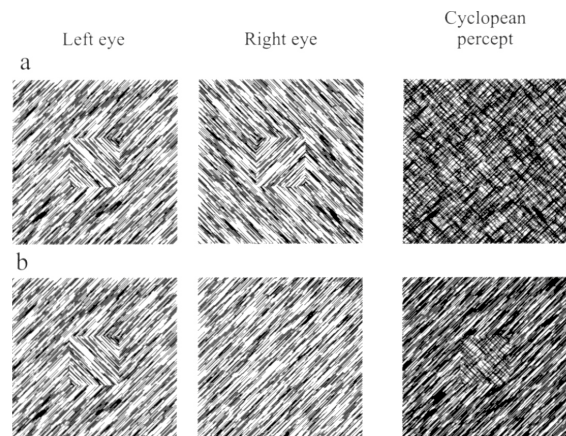
Left eye   Right eye   Cyclopean percept

Figure 4: (*a*) Orthogonal textures presented to the two eyes, each containing a figure on a background yield a cyclopean percept of a homogenous texture, with no visible figure (at excentric fixation). (*b*) When the figure is present in one eye only, the cyclopean percept is that of a figure on a background. The stimulus in *b* yields figure-ground related contextual modulation, the stimulus in *a* does not (46).

of information processing. The increasingly complex receptive field tuning properties that one observes when going upstream through the hierarchy of visual areas suggest that low level features are combined into complex constellations of features in higher areas (7,55). Strict feedforward processing, however, has its limitations, for example in terms of the number of neurons that is needed to map every possible combination of features onto all possible outputs (the combinatorial explosion). In this context, many have advocated population coding as opposed to grandmother cell hypotheses (53,56,57). In population coding, cells might at some time be engaged in a processing task with one set of neurons (an assembly), but at another time in a different task with a different set. For this mechanism to work, cells have to be labelled as belonging to the same assembly, with a label that can rapidly be switched on and off. For example, in processing an image such as figure 3a, such a label should tag the neurons that code for the elements of the figure separately from the neurons that code for the background.

Synchrony of firing between cells has been proposed to act as such a label (53). Neurons in V1 fire their action potentials in relative synchrony when their receptive fields are stimulated with a coherently moving bar, and this synchrony is reduced when the neurons are co-stimulated with separate bars (58,59). On the basis of the hypothesis that synchrony labels neurons belonging to the same assembly, one would predict that neurons whose receptive fields fall within the figure region of an image such as figure 3a, fire in synchrony with other neurons whose receptive fields fall within the figure, and fire not in synchrony with neurons whose receptive fields fall on the background. We tested this, by recording from multiple sites simultaneously in V1, but did not find that to be the case. Synchrony between neurons at considerable distance

was found. However, this synchrony could be equally strong between neuron pairs that had their receptive fields both within the figure as between pairs that had their receptive fields on either side of the boundary between figure and ground (60). The results suggested that synchrony reflects the interactions mediated by local and horizontal connections, but does not take into account more global (feedback) interactions that form the basis of the figure-ground percept. Apparently, synchrony in V1 does not operate as a label tagging neurons to belong to an assembly coding for the figure-ground percept.

An alternatively binding tag could be an enhanced firing rate. In that case, cells engaged in the processing of features or elements of the same object would have an enhanced firing rate compared to other cells. The results discussed above (section 3) provide evidence that the neural system might use enhanced firing rate in this way. All neurons responding to elements of the same figure have an equal amount of response enhancement (Fig. 3c). Also, colinear line segments that group together share an enhanced firing rate (40). A drawback of firing rate as a binding tag is that it is difficult to separate several assemblies from each other, while this can easily be achieved with synchrony as a tag (53). It is however questionable whether the visual system is indeed capable of representing many objects simultaneously. Experiments on change blindness indicate that not more than two to three objects are represented by the visual system at a time (61,62). Visual search experiments show that multiple features of objects can only be linked going from one object to the next in a serial manner (63,64). These and many other experiments show that feature binding is a process that is mediated by attentional mechanisms(63). The relation between contextual modulation and attentional feature binding will be further discussed below.

4.3. Contextual modulation is related to object based attention

Perceptually salient elements (section 4.1) are strong bottom-up attention grabbers. In that sense, contextual modulation can also be viewed as representing the amount of attention that is drawn by elements in the image. A distinction can be made between this kind of bottom-up attention and top-down attention. Top-down attention is considered to be an influence on early processing on the basis of motivation, behavioural setting or other central factors.

Numerous studies have reported modulatory effects of top-down attention on neuronal activity in many cortical areas (for a review see ref 65). Whether these effects can also be observed in V1 has been somewhat controversial but several recent studies have shown clear attentional effects in V1 (66). We have recorded activity from V1 neurons while monkeys were doing a curve tracing task. In this task, several curves are projected on a screen, and the animals are required to mentally trace one of them (the target curve) without making eye movements. Concurrent psychophysical studies indicate that in such a task more attention is allocated to the curve that is traced than to the distractor curves. The V1 responses reflected the

attentional enhancement of the target curve; responses to line segments of the target curve were enhanced relative to responses to line segments of the distractor curve (67). In other words, the neural image of the whole target curve is 'highlighted' relative to the other curves.

This is of course very similar to the 'highlighting' of the whole figure surface that is observed in relation to figure-ground segregation (figure 3). Another similarity is that both effects occur after a considerable delay (as do many attentional effects, both in human ERP as well as in monkey single unit studies (65)). Apparently figure-ground segregation and figure-figure separation (one curve from the others) share many aspects of their neural representations. If we assume that in the figure-ground stimuli the figure draws more attention than the background, an important further conclusion is that bottom-up and top-down attention have very similar neural correlates.

A further distinction is between spatial and object based attention (54,63,68,69). In Feature Integration Theory (63), attentional mechanisms are required to turn features into objects. Two types are distinguished, spatial and object based attention. Features of the same object may be represented in separate maps (e.g. colour and orientation) and by focusing attention to the location in space where the two coincide these features can be tagged as belonging to the same object. Objects can also overlap in space. Features therefore also need to be tagged as belonging to a particular object. This requires object based attention (63,68,69).

It is unlikely that contextual modulation is a reflection of focal spatial attention; two separate and distant figures evoke the same amount of modulation as one (70). Also the attentional enhancements observed in relation to curve tracing can only be attributed to object based attention; when curves overlap, the traced curve still is enhanced relative to the distractor curve, even though they share some space (67). There are good reasons to consider the modulations to be a neural correlate of object based attention. These modulations precisely highlight the elements of an object that segregates from background or from another object. These segregations must operate on the basis of perceptual grouping criteria like proximity, similarity, and colinearity. The modulations may thus serve to encode for the attentional binding of features into objects. This is in fact the same role as suggested in the previous section (contextual modulation as a binding tag) with attention serving as the binding mechanism (63,64).

4.4. Contextual modulation reflects visual awareness

Attention is intimately related to visual awareness yet cannot be fully equated to it (71,72). To further elucidate this relation, it might be useful to compare their neural substrates (73). The neural correlate of visual attention at the single unit level seems to be an enhanced activity of neurons representing the attended features, objects, spatial locations etc, at the expense of non-attended ones (65). The neural correlate of visual awareness is however still a matter of much debate (74), and the role of V1 is at the core of the controversy (75).

It has been argued that V1 should be excluded from the neural substrate of visual awareness, because many cells in V1 respond to stimulus attributes of which we are not aware (74,75,76). Apparently, the activity of these neurons does not suffice for the stimulus attribute to reach awareness. However, that does not exclude the possibility that other V1 neurons do show activity that is correlated to perception and whose activity might be sufficient for that percept.. Using fMRI, it has been shown that V1 is activated during visual imagery (74). Rivalry experiments show that at least the activity of a small proportion of cells correlates with what stimulus is perceived (77). More importantly, a comparison between receptive field tuning properties and contextual modulation shows that some types of activity are more intimately related to perception than others and that both types of activity may co-exist in a single area. Moreover, a neuron whose activity does not correlate with how a scene is perceived before 100 ms, might correlate very well with perception at longer latencies. This makes it rather difficult to attribute a particular function or role in visual awareness to a particular neuron, let alone to a whole cortical area. It will be more fruitful to look for the *processes* that constitute awareness than for the *areas* that do so.

Contextual modulation might be a good candidate for the neural manifestation of such a process. There are several converging pieces of evidence to support this. First, there are stimulus manipulations that have strong effects on whether stimuli are perceived or not, and these seem to effect late onset modulations in particular (50). We did an experiment that was inspired by the stimuli used by Kolb and Braun to demonstrate blindsight in normal observers (78). When two figure-ground displays with orthogonal orientations are presented each to one of the two eyes, the fused Cyclopean percept is that of a homogenous texture with no figure present in it (figure 4a). When one eye is presented with a homogenous texture and the other eye with a figure-ground display, the figure is visible (figure 4b). The latter stimulus evokes contextual modulation signalling the presence of the figure, but the former stimulus does not (46); so while the figures are present in either eye alone the modulation signals the percept (no figure present) rather than this information.

Of course the above example is a manipulation of the stimulus rather than of visual awareness. A more critical experiment would be to leave the stimulus identical and manipulate awareness instead. A (rather crude) way of manipulating awareness is anaesthesia. While receptive field tuning properties of V1 neurons are little or not affected by anaesthesia (32,79), contextual modulation is affected by anaesthesia to different degrees. Short latency modulations, evoked by surround stimuli that may exert their effects through local or horizontal connections within V1 (figure 2), can be recorded in awake as well as anesthetized animals (35,40,80,81). However, modulations reflecting perceptual pop-out are stronger in awake than in anesthetized animals (82). Finally, contextual modulation

related to figure-ground segregation is fully suppressed by anaesthesia (70). The latter type of modulation seems to depend strongly on feedback from extrastriate areas (83). It seems as if the longer information has to 'travel' over the network of local, horizontal, and feedback connections to evoke the modulatory effects, the more susceptible these effects are to anaesthesia.

Anesthesia will not only affect visual awareness but probably many other processes as well. A more direct link to visual awareness would only be established when a trial by trial comparison of perceived versus not perceived stimuli, that are otherwise identical, is performed. This was recently done by using figure-ground displays like those shown in figure 3. Contextual modulation was recorded in monkeys that had to report whether the figure was perceived or not. This was done in a manner very similar to the one used to demonstrate that monkeys do not perceive stimuli in blindsight (84). Catch-trials, in which no figure was present at all, were presented in combination with figure-present trials, where a figure appeared at one of three possible locations. The animal's task was to indicate the position of the figure, when present, by making a saccadic eye movement towards it. On catch-trials, the animal was rewarded when it remained fixating. The key feature of the paradigm is that when for some reason the figure in a figure-present trial is not perceived, the monkey will signal this by maintaining fixation. Neural responses recorded during figure-present trials that resulted in a correct saccade were compared to responses from figure-present trials that were classified as figure-absent (catch-trial) by the monkey. Figure-ground related contextual modulation was strongly reduced or absent in the case where the monkey did not perceive the figure (85). In other words, contextual modulation only reflects the figure-ground relationships when these are (consciously) perceived by the animal.

5. CONCLUSIONS

The responses of V1 cells appear to be a mixture of the well established low level activity related to the detection of elementary features of the image, and activity that correlates with aspects of perceptual organization, attention, and visual awareness. The latter type of activity, contextual modulation, takes into account information from very distant parts of the visual scene, to signal a generalized notion of perceptual saliency of the image elements that fall on the receptive field of the neuron that is recorded from. It thus highlights those neurons in the brain that represent features that are in some way of more importance to our behavioural decisions than others. In that sense it is strongly related to neurophysiological correlates of attention. In particular, it might play a role in attentive feature binding. Finally, contextual modulation appears to be associated with processes that make it possible for a visual stimulus to reach visual awareness.

6. REFERENCES

1. Hubel D.H., T.N. Wiesel: Receptive fields and functional architecture of monkey striate cortex. *J Physiol (London)* 195, 215-243 (1968)

2. Hubel D.H., T.N. Wiesel: Ferrier Lecture. Functional architecture of macaque monkey visual cortex. *Proc R Soc Lond B* 198, 1-59 (1977)

3. Maunsell J.H.R., D.C. Van Essen: Functional properties of neurons in the middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed and orientation. *J Neurophysiol* 49, 1127-1147 (1983)

4. Allman J.M., F. Miezin, E. McGuiness: Direction and velocity-specific responses from beyond the classical receptive field in the middle temporal visual area (MT). *Perception* 14, 105-126 (1985)

5. Newsome W.T., A. Mikami, R.H. Wurtz: Motion selectivity in macaque visual cortex. III. Psychophysics and physiology of apparent motion. *J Neurophysiol* 55, 1340-1351 (1986)

6. Movshon J.A., E.H. Adelson, M.S. Gizzi, W.T. Newsome: The analysis of moving visual patterns. In *Pattern recognition mechanisms*, ed. C. Chagas, R. Gatass, C. Gross, pp 117-151. New York. Springer-Verlag (1986)

7. Maunsell J.H.R., W.T. Newsome: Visual processing in monkey extrastriate cortex. *Ann Rev Neurosci* 10, 363-401 (1987)

8. Zeki S.M.: Colour coding in rhesus monkey prestriate cortex. *Brain Res* 53, 422-427 (1973)

9. Zeki S.M.: The representation of colours in the cerebral cortex. *Nature* 284, 412-418 (1980)

10. Desimone R., S.J. Schein: Visual properties of neurons in area V4 of the macaque: Sensitivity to stimulus form. *J Neurophysiol* 57, 835-868 (1987)

11. Gallant J.L., C.E. Connor, S. Rakshit, J.W. Lewis, D.C. Van Essen: Neural responses to polar, hyperbolic, and cartesian gratings in area V4 of the macaque monkey. *J Neurophysiol* 76, 2718-2739 (1996)

12. Dean P.: Visual cortex ablation and thresholds for succesively presented stimuli in rhesus monkeys: II. Hue. *Exp Brain Res* 35, 69-83 (1979)

13. Newsome W.T., R.H. Wurtz, M.R. Dürsteler, A. Mikami: Deficits in visual motion perception following ibotenic acid lesions of the middle temporal visual area of the macaque monkey. *J Neurosci* 5, 825-840 (1985)

14. Heywood C.A., A. Cowey: On the role of cortical area V4 in the discrimination of hue and patern in macaque monkeys. *J Neurosci* 7, 2601-2617 (1987)

15. Wild H.M., S.R. Butler, D. Carden, J.J. Kulikowski: Primate cortical area V4 important for color constancy but not wavelength discrimination. *Nature* 313, 133-135 (1985)

16. Schiller P.H., K. Lee: The role of the primate extrastriate area V4 in vision. *Science* 251, 1251-1253 (1991)

17. Schiller P.H.: The effects of V4 and middle temporal (MT) area lesions on visual performance in the rhesus monkey. *Visual Neurosci* 10, 717-746 (1993)

18. DeYoe E.A., D.C. Van Essen: Concurrent processing streams in monkey visual cortex. *Trends Neurosci* 11, 219-226 (1988)

19. Livingstone M.S., D.H. Hubel: Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science* 240, 740-749 (1988)

20. Felleman D.J., D.C. Van Essen: Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex* 1, 1-47 (1991)

21. Gilbert C.D., T.N. Wiesel: Columnar specificity of intrinsic horizontal and cortico-cortical connections in cat visual cortex. *J Neurosci* 9, 2432-2442 (1989)

22. Gilbert C.D.: Horizontal integration and cortical dynamics. *Neuron* 9, 1-13 (1992)

23. Salin P., J. Bullier: Corticocortical connections in the visual system: Structure and function. *Physiol Reviews* 75, 107-154 (1995)

24. Gilbert C.D.: Circuitry, architecture and functional dynamics of visual cortex. *Cereb Cortex* 3, 373-386 (1993)

25. Jones B.H.: Responses of single neurons in cat visual cortex to a simple and a more complex stimulus. *Am J Physiol* 218, 1102-1107 (1970)

26. Blakemore C., E.A. Tobin: Lateral inhibition between orientation detectors in the cat's visual cortex. *Exp Brain Res* 15, 439-440 (1972)

27. Maffei L., A. Fiorentini: The unresponsive regions of visual cortical receptive fields. *Vision Res* 16, 1131-1139 (1976)

28. Nelson J.I., B. Frost: Orientation selective inhibition from beyond the classic visual receptive field. *Brain Res* 139, 359-365 (1978)

29. Albus K., W. Fries: Inhibitory sidebands of complex receptive fields in tha cat's striate cortex. *Vision Res* 20, 369-372 (1980)

30. Allman J.M., F. Miezin, E. McGuiness: Stimulus specific responses from beyond the classical receptive field: Neurophysiological mechanisms for local-global comparisons in visual neurons. *Ann Rev Neurosci* 8, 407-430 (1985)

31. Desimone R., J. Moran, S.J. Schein, M. Mishkin: A role for the corpus callosum in visual area V4 of the macaque. *Visual Neurosci* 10, 159-171 (1993)

32. Schiller P.H., B.L. Finlay, S.F. Volman: Quantitative studies of single cell properties in monkey striate cortex. I-V. *J Neurophysiol* 39, 1288-1374 (1976)

33. DeValois R.L., D.G. Albrecht, L.G. Thorell: Spatial frequency selectivity of cells in macaque visual cortex. *Vision. Res* 22, 545-559 (1982)

34. Poggio G.F.: Mechanisms of stereopsis in monkey visual cortex. *Cereb Cortex* 3, 193-204 (1995)

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

36. Nothdurft H.C.: Texture segmentation and pop-out from orientation contrast. *Vision Res* 31, 1073-1078 (1991)

37. Nothdurft H.C.: Common properties of visual segmentation. In: *Higher-order processing in the visual system*, eds. R. Bock, and J.A. Goode, Ciba Foundation Symposium 184, pp 245-268, Wiley, Chichester (1994)

38. Nothdurft H.C.: Sensitivity for structure gradient in texture discrimination tasks. *Vision Res* 25, 1957-1968 (1985)

39. Landy M.S., J.R. Bergen: Texture segregation and orientation gradient. *Vision Res* 31, 679-691 (1991)

40. Kapadia M.K., M. Ito, C.D. Gilbert, G. Westheimer: Improvement in visual sensitivity by changes in local context: parallel studies in human observers and in V1 of alert monkeys. *Neuron* 15, 843-856 (1995)

41. Field D.J., A. Hayes, F. Hess: Contour integration by the human visual system: evidence for a local 'association field'. *Vision Res* 33, 173-193 (1993)

42. Kovacs I.: Gestalten of today: early processing of visual contours and surfaces. *Behav Brain Res* 10, 100-110 (1996)

43. Nelson J.I., B. Frost: Intracortical facilitation among co-oriented, co-axially aligned simple cells in cat striate cortex. *Exp Brain Res* 6, 54-61 (1985)

44. Lamme V.A.F.: The neurophysiology of figure-ground segregation in primary visual cortex. *J Neurosci* 15, 1605-1615 (1995)

45. Lamme V.A.F., V. Rodriguez, H. Spekreijse: Separate processing dynamics for texture elements, boundaries and surfaces in primary visual cortex. *Cerebral Cortex* 9, 406-413 (1999)

46. Zipser K., V.A.F. Lamme, P.H. Schiller: Contextual modulation in primary visual cortex. *J Neurosci* 16, 7376-7389 (1996)

47. Lee T.S., R. Mumford, R. Romero, V.A.F. Lamme: The role of the primary visual cortex in higher level vision. *Vision Res* 38, 2429-2454 (1998)

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

49. Schiller P.H., M. Smith: Monoptic and dichoptic metacontrast. *Percept Psychophys* 3, 237-239 (1968)

50. Bridgeman B.: Temporal response characteristics of cells in monkey striate cortex measured with metacontrast masking and brightness discrimination. *Brain Res* 196, 347-364 (1980)

51. Kovacs I., B. Julesz: Perceptual sensitivity maps within globally defined visual shapes. *Nature* 370, 644-646 (1994)

52. Levitt J.B., J.S. Lund: Contrast dependence of contextual effects in primate visual cortex. *Nature* 387, 73-76 (1977)

53. Singer W., C.M. Gray: Visual feature integration and the temporal correlation hypothesis. *Annu Rev Neurosci* 18, 555-586 (1995)

54. Treisman A.: The binding problem. *Curr Opin Neurobiol* 6, 171-178 (1996)

55. Barlow H.B.: The neuron doctrine in perception. *The cognitive neurosciences*, pp 415-435, ed. M.S. Gazzaniga, MIT press, Cambridge Ma (1995)

56. Abeles M.: *Local cortical circuits*. Springer-verlag, Berlin (1982)

57. Georgopoulos A.P.: Higher order motor control. *Ann Rev Neurosci* 14, 361-377 (1991)

58. Gray C.M., A.K. Engel, P. König, W. Singer: Oscillatory responses in cat visual cortex exhibit intercolumnar synchronization which reflects global stimulus properties. *Nature* 338, 334-337 (1989)

59. Freiwald W.A., A.K. Kreiter, W. Singer: Stimulus dependent intercolumnar synchronization of single unit responses in cat area 17. *Neuroreport* 6, 2348-2352 (1995)

60. Lamme V.A.F., H. Spekreijse: Neuronal synchrony does not represent texture segregation. *Nature* 396, 362-366 (1998)

61. Irwin D.E.: Eye movements and scene perception: memory for things observed. *Invest Ophthalmol Vis Sci* 38, S707 (1999)

62. Rensink R.A.: How much of a scene is seen? The role of attention in scene perception. *Invest Ophthalmol Vis Sci* 38, S707 (1997)

63. Treisman A.: The perception of features and objects. In: *Attention: Selection, awareness and control: A tribute to Donald Broadbent.* pp. 5-35, ed. Baddeley, A., Weiskrantz, L., Clarendon Press, Oxford (1993)

64. Wolfe J.M., S.C. Bennett: Preattentive object files: shapeless bundles of basic features. *Vision Res* 37, 25-43 (1997)

65. Desimone R., J. Duncan: Neural correlates of selective visual attention. *Ann Rev Neurosci* 18, 193-222 (1995)

66. Motter B.C.: Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *J Neurophysiol* 70, 909-919 (1993)

67. Roelfsema P.R., V.A.F. Lamme, H. Spekreijse: Object based attention in primary visual cortex of the macaque monkey. *Nature* 395, 376-381 (1998)

68. Duncan J.: Selective attention and the organization of visual information. *J Exp Psychol Gen* 113, 501-517 (1984)

69. Vecera S.P., M.J. Farah: Does visual attention select objects or locations? *J Exp Psychol Gen* 123, 146-160 (1994)

70. Lamme V.A.F., K. Zipser, H. Spekreijse: Figure-ground activity in primary visual cortex is suppressed by anaesthesia. *Proc Natl Acad Sci* 95, 3263-3268 (1998)

71. Posner M.I.: Attention: the mechanisms of consciousness. *Proc Natl Acad Sci USA* 91, 7398-7403 (1994)

72. Block N.: How can we find the neural correlate of consciousness? *Trends Neurosci* 19, 456-459 (1996)

73. Newsome W.T.: Visual attention: Spotlights, highlights and visual awareness. *Curr Biology* 6, 357-360 (1996)

74. Crick F., C. Koch: Consciousness and neuroscience. *Cerebral Cortex* 8, 97-107 (1998)

75. Crick F., C. Koch: Are we aware of neural activity in primary visual cortex? *Nature* 375, 121-123 (1995)

76. He S., P. Cavanagh, J. Intriligator: Attentional resolution and the locus of visual awareness. *Nature* 383, 334-336.

77. Leopold D.A., N.K. Logothetis: Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature* 379, 549-553 (1996)

78. Kolb F.C., J. Braun: Blindsight in normal observers. *Nature* 377, 336-338 (1995)

79. Snodderly D.M., M. Gur: Organization of striate cortex of alert, trained monkeys (Macaca fascicularis): ongoing activity, stimulus selectivity, and widths of receptive field activating regions. *J Neurophysiol* 74, 2100-2125 (1995)

80. Kastner S., H.C. Nothdurft, I.N. Pigarev: Neuronal correlates of pop-out in cat striate cortex. *Vision Res* 37, 371-376 (1997)

81. Polat U., K. Mizobe, M.W. Pettet, T. Kasamatsu, A.M. Norcia: Collinear stimuli regulate visual responses depending on cell's contrast threshold. *Nature* 391, 580-584 (1998)

82. 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. *Vis Neurosci* 16, 15-34 (1999)

83. Lamme V.A.F., H. Supèr, H. Spekreijse: Feedforward, horizontal, and feedback processing in the visual cortex. *Curr Opin Neurobiol* 8, 529-535 (1998)

84. Moore T., H.R. Rodman, A.B. Repp, C.G. Gross: Localization of visual stimuli after striate cortex damage in monkeys: Parallels with human blindsight. *Proc Natl Acad Sci USA* 92, 8215-8218 (1995)

85. Supèr H., V.A.F. Lamme, H. Spekreijse: Contextual modulation in monkey primary visual cortex (V1) matches figure-ground perception. *Invest Ophthalmol Vis Sci* 40, S357 (1999)

86. Andersen R.A., L. Snyder, C.S. Li, B. Stricanne: Coordinate transformations in the representation of spatial information. *Curr Opin Neurobiol* 3, 171-176 (1993)

87. Jeannerod M., M.A. Arbib, G. Rizolatti, H. Sakata: Grasping objects: the cortical mechanisms of visuomotor transformation. *Trends Neurosci* 18, 314-320 (1995)