UNIVERSITÀ DEGLI STUDI DI FIRENZE

Dottorato di Ricerca in
Psicologia e Neuroscienze
ciclo XXVI
Settore Scientifico Disciplinare M-PSI/02

Object segmentation and enumeration

Dottorando
Dott. Castaldi Elisa

Tutore
Prof. Burr David Charles
Prof.ssa Morrone Maria Concetta

Anni 2011/2013

UNIVERSITÀ DEGLI STUDI DI FIRENZE
Dottorato di Ricerca in Psicologia e Neuroscienze
ciclo XXVI
Settore Scientifico Disciplinare M-PSI/02

Object Segmentation and enumeration

Candidate:
Dott. Castaldi Elisa

Supervisor:
Prof. Burr David Charles
Prof.ssa Morrone Maria Concetta

The human brain is able to reconstruct a symbolic representation of the scene starting from highly complex and variable visual information falling on the retina. Without any effort we can instantaneously understand the content of the image: which objects are there and how many of them there are.

In order to understand which objects are represented in a visual scene and to segment them from the background, the visual system has to identify the contour through the identification of edges and lines. These salient visual features are identified by peaks in local energy that occur at those points in the space where the Fourier components of the luminance profile come into phase.

A key role in features identification process has been attributed to the achromatic simple and complex cells in V1, because of their oriented and spatial phase selective receptive fields. On the contrary, chromatic opponent cells with their circular receptive fields were thought to be minimally sensitive to shape coding, but particularly useful in coding chromatic information. However, recent research on monkeys has demonstrated the existence of chromatic neurons with oriented and phase-sensitive receptive fields in the primary visual area.

The first part of this thesis evaluates the role of colour information in contour detection (edges and lines) and shape perception. BOLD response selectivity both to edge and line stimuli (both stimuli with a congruent phase spectrum) and to an amplitude-matched random noise control stimulus modulated both in luminance or red‑green equiluminant colour contrast has been measured. All these stimuli had identical power but different phase spectra, either highly congruent (even or odd symmetry stimuli) or random (noise). At equiluminance, V1 BOLD activity showed no preference between congruent- and random-phase stimuli, as well as no preference between

even- and odd-symmetric stimuli. Areas higher in the visual hierarchy, both along the dorsal pathway (caudal part of the intraparietal sulcus, dorsal LO and V3A) and the ventral pathway (V4), responded preferentially to odd-symmetric over even-symmetric stimuli, and to congruent over random phase stimuli. Interestingly, V1 showed an equal increase in BOLD activity at each alternation between stimuli of different symmetry, suggesting the existence of specialised mechanisms for the detection of edges and lines such as even- and odd-chromatic receptive fields.

Overall the results indicate that the analysis of the energy of the image necessary for object segmentation starts from the primary visual cortex both for luminance and for colored stimuli and then spreads up to the dorsal and ventral stream.

As soon as shapes are defined, it is also possible to count how many objects are contained in an image. The "Number Sense", the human ability to give an approximate estimate of the number of elements in a scene at a single glance, is thought to be a basic visual dimension. Increasing evidence is characterizing the biological substrate of the perception of numerosity. Neurophysiological experiments on monkeys have described number selective neurons, and their existence has been hypothesized for humans as well from many neuroimaging studies. Indeed neuroimaging studies investigating brain regions encoding numerosity have reported fMRI habituation signals in the intra-parietal sulcus (IPS), a region where activation patterns for different numerosities can be accurately discriminated and generalized across changes in low-level stimulus parameters and formats. A numerotopic map has recently been described at this level.

Psychophysical experiments have shown that numerosity, like the other visual properties of the image, is susceptible to adaptation after-effects. In the second part of this thesis an experiment testing the effect of adaptation on numerosity decoding in human brain is presented. Importantly, to further investigate the role of the primary visual area in the perception of numerosity, the overall contrast energy across numbers has been balanced, since, as shown in the first experiment presented in this thesis, the energy is the fundamental property modulating V1 activity.

Moreover a novel adaptation paradigm is presented, wherein rapid adapting and testing stimuli can be separated by more than 20 seconds. This procedure is ideally suited for measuring BOLD adaptation with rapid event related design. Indeed it allows for temporal dissociation between activity evoked by the adaptation and test stimuli. The novel adaptation paradigm produces psychophysical adaptation to number, causing a decrease in the perceived numerosity of stimuli presented.

The BOLD signal differences between Pre- and Post- Adaptation where present only in the intra parietal sulcus bilaterally, leaving the primary visual cortex unaffected.

Multivariate pattern recognition on functional imaging data was used to classify brain activity evoked by non-symbolic numbers before and after adaptation to the highest numerosity. Classifiers were trained on one adaptation condition (either using Pre or Post Adaptation trials) and then tested with unseen trials left out from the training data as well as unused trials of the other adaptation condition.

The results show that only classifiers in IPS and not in V1 were able to decode numerosities. Interestingly adaptation causes an increase in the IPS classifier's ability to decode numbers, suggesting an increased tuning refinement of the numerosity detectors after adaptation, which is supported also by the decrease in JND.

Overall this evidence supports previous findings showing that numerosity adaptation does not rely on low-level stimulus parameters, but that it alters higher-order representation of magnitude.

Index

1. General Introduction

Whenever we look at any visual scene we instantaneously perceive in a single glance the identity and the number of the objects contained in it.

Although we quickly and effortlessly understand the content of an image, this is not a trivial task.

Indeed images falling on the retina are very complex and highly variable in average luminance, chromatic content, contrast and size.

The visual system does not simply transmit this complex information passively from the eye to the brain, but it actively analyses the image to produce a symbolic representation of the visual scene.

The only way the visual system has to reconstruct a coherent representation of the visual scene is to extract the most salient information from an image, such as lines and edges, determining contours of an object.

The importance of features extraction process was understood by David Marr (1976) who suggested that the visual system starts to reconstruct the visual scene from a *primal sketch*, that is a caricature composed of lines, edges and other simple forms that capture the essence of a scene neglecting many details. Lines and edges are indeed rich source of information that we use to delimit object extent, like it is shown by artistic sketches in Figure 1.

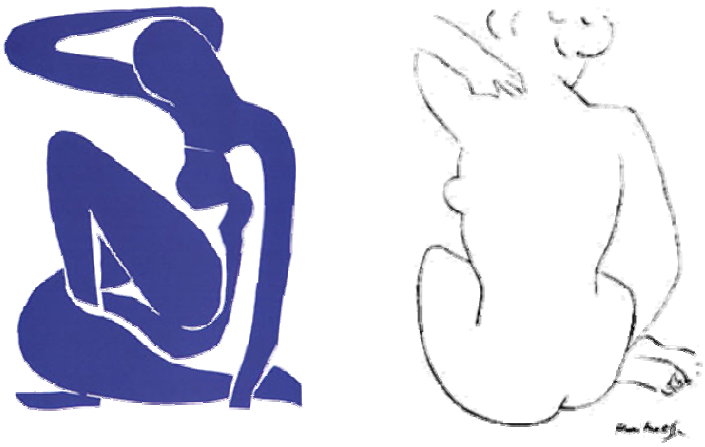

Figure 1: Women's bodies defined simply by edges and lines. Artistic sketches: "Blue nude" and "Black Sketch" by Henri Matisse, 1952.

A classical way to demonstrate that the visual system does not simply take a picture of the visual scene and map it into a mental representation, but instead it deals with a reconstruction process, is provided by visual illusions.

Indeed in the image reconstruction process the visual system can get confused, producing an illusory perception. Knowing which physical properties of the stimulus caused such illusion brings often to a better understanding of the coding mechanism itself.

Two examples of visual illusions are shown in Figure 2.

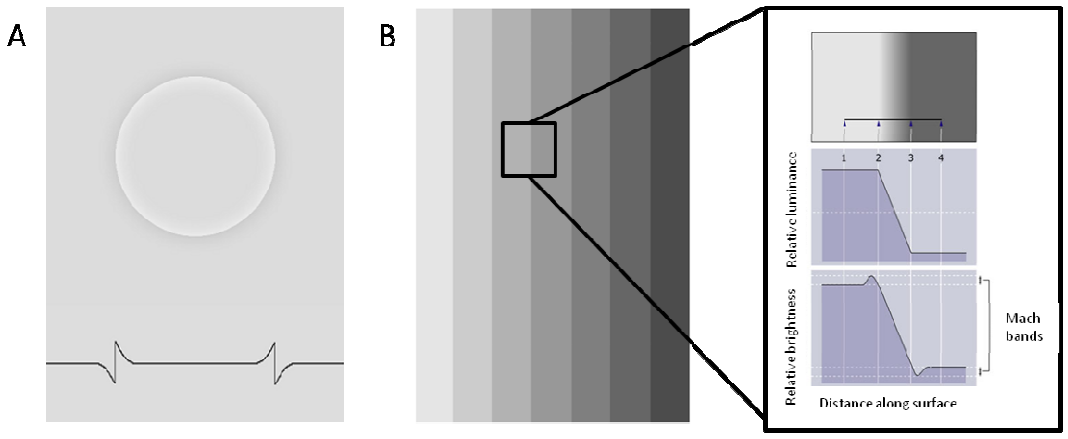

Figure 2: A. Craik- O'Brein Cornsweet illusion B. Mach bands

In the Craik-O'Brein Cornsweet illusion (Fig. 2A) the area enclosed within the edges seems brighter then the surround although the luminance is the same. Infact occluding the edges abolishes the perception of the lighter surface in the center, demonstrating that the brighter circle is indeed illusory. This illusion shows how identifying an edge defined by a contrast change can causes a different perception of brightness, even if the luminance is the same*.

*Luminance is a physical quantity, which gives a measure of the light intensity per unit area radiated from a surface in a certain direction. It is expressed in candelas per square meter. Brightness is the perceptual correlate of luminance, defined as an attribute of visual sensation according to which an area appears to emit more or less light.

Another example is provided by the Mach bands (Fig. 2B) where ramped changes in luminance give the illusion that there is a pair of thin dark and light bands at any luminance transition. Lightness profile shows instead that there are no actual bands.

In everyday life we constantly encounter such illusions. Indeed natural scenes contain many luminance profile associated to lines, edges, triangular waves, ramps and hybrids of them (Perona and Malik 1990).

Nevertheless the visual system has to identify simultaneously and quickly all the important features contained in the image in order to reconstruct the primal sketch.

Moreover shape is not the only source of information that the visual system has to analyse.

In any handbook dealing with visual perception it is explained that one way to successfully integrate all the different information into a coherent percept can be to decompose the complexity of the visual scene into different sources of information such as form, colour, movement, position, to analyse them separately and then to integrate them into a complete percept (Levine, Warach et al. 1985). In order to perform this task two distinct parallel pathways are dedicated to analyse the different components of visual scenes.

The different parallel information processing begins when visual information, after an initial transformation performed by the retinal circuitry, is conveyed to the cortex along two main pathways: the magnocellular (M) and the parvocellular (P) pathways.

From the M- and P- ganglion cells of the retina, these pathways, after passing through different parts of V1 and V2, spread respectively dorsally

until mediotemporal and posterior parietal cortex and ventrally until the inferior temporal cortex.

Traditionally the M- and P- pathway are often called “where” and “what” pathways respectively in order to underline the kind of information they analyse: position and motion of objects in the space from one hand, and identification of objects from the other.

This sharp separation of information processing is quite artificial and it is widely accepted that these two pathways interact and contribute in the analysis of the different kind of information. For example it is known that both M- and P- pathways contribute in analyse movement, being the first specialized in the high temporal frequencies and the latter mainly sensitive to the lower temporal frequencies.

However, in spite of the demonstrated cooperation between the two pathways, the colour and the shape of an object, and the analysis underlying their detection, have always been considered two separate entities.

The reason is that the major difference between M- and P- pathways is their selectivity to the chromatic properties of the image. Indeed both M- and P- pathways contribute in analyzing the form detecting the luminance contrast, having respectively low and high spatial frequencies resolution, but only the P-pathway can detect colour being the M-pathway essentially achromatic.

Hence the traditional view has always assigned the detection of form to achromatic mechanisms. To reinforce the link between form and achromaticity there is also the evidence that, since they were found by Hubel and Wiesel, oriented receptive field of V1 cells had been characterized as operators that detect luminance (and not colour!) contrasts as those

typically associated with edges and contours in an image, contributing in this way to the perception of shape of the objects.

These achromatic mechanisms could achieve the goal in detecting contours thanks to their sensitivity to spatial phase. Indeed according to local energy model of feature detection of Morrone and Burr (1988), borders typically correspond to those points in the space where the Fourier components of the luminance profile tend to come into phase, that is in other words, that several harmonics have the same phase alignment determining edges or lines.

The oriented organization of the achromatic mechanisms' receptive fields allow them to match with the luminance profile of lines or edges, depending on their even- or odd- symmetry respectively, while this match is not performable for colour being the chromatic mechanisms' receptive field organized in concentric, and not oriented, structure.

Indeed at the retinal and NGL level, the receptive fields of visual neurons are arranged in symmetric concentric antagonistic excitatory and inhibitory regions both for luminance and for colour detectors.

At V1 level many of the achromatic circular NGL receptive fields probably converge on individual simple cell of V1 forming a straight oriented receptive field with excitatory and inhibitory subregions that detect luminance contrast. Probably different simple cells project on complex cells that have an orientation axis too but there are no well defined excitatory and inhibitory subregions.

On the contrary, colour detectors maintain a circularly organized receptive field from the ganglion retinal cells to the V1 chromatic opponent cells.

Hence the sensitivity to phase of the oriented simple and complex V1 cells' receptive fields allow them to analyse object shape on the basis of the luminance contrast.

In this frame one can easily understand why the perception of form has always been kept separate from the perception of colour, the latter not having even the detectors necessary for analyzing form, given that cells in V1 that analyse colour does not have orientation tuning.

However recent evidence suggests that also colour contrast might contribute in the construction of the form representation.

Indeed recently Johnson et al. (2008) demonstrated that even chromatic neurons in monkey primary visual area have a similar straight and oriented receptive field structure as the luminance one and that they are sensitive to the phase of the stimulus. This opens the possibility that there might be also for colour a congruence spatial phase coding between the harmonic components that compose a stimulus and hence that also the chromatic information could contribute to detect contours and to the shape perception as well as the luminance one.

In the first part of this thesis I will concentrate on how human visual system extracts the most salient features from an image, such as lines and edges, determining contours of an object and how this form information is related with the colour one.

In agreement with studies on monkeys (Johnson, Hawken et al. 2001; Johnson, Hawken et al. 2004; Johnson, Hawken et al. 2008) and with previous psychophysical studies (Girard and Morrone 1995; Martini, Girard et al. 1996) I will show that the information related to the shape and the one related to the colour are not necessary distinct and processed separately.

As soon as objects are defined, it is also possible to count them, since form and number information are inextricably linked.

Humans possess a non-verbal capacity to give an approximate estimate of the number of elements contained in a visual scene. This ability is called Number Sense and is thought to be an ancient and biological encoded ability that does not depend on either language nor education (Dehaene and Cohen 1997). Being able to rapidly have an idea of how many objects are contained in a visual scene is important for the individual's fitness and survival, regulating social interactions and ultimately decision making (for example deciding to attack or to join a group of individuals, or to choose the place with more food).

Indeed the ability to discriminate numerosity is not only present in humans, but also animals are able to count (Gallistel and Gelman 2000; Huntley-Fenner 2001; Jordan and Brannon 2006; Jordan and Brannon 2006). Moreover infants of only a few months of age (Starkey, Spelke et al. 1990; Xu and Spelke 2000; Lipton and Spelke 2003; Brannon, Libertus et al. 2008; Izard, Dehaene-Lambertz et al. 2008; Hyde and Spelke 2011) and adults without formal mathematical training (Gordon 2004; Pica, Lemer et al. 2004; Frank, Everett et al. 2008) can discriminate numerosities. This approximate non verbal numerosity system is probably the biological precursor for the abstract and symbolical perception of number and it let us to estimate the number of objects in a visual scene without counting them one by one (Dehaene and Cohen 1997; Feigenson, Dehaene et al. 2004; Piazza, Pinel et al. 2007; Cantlon, Libertus et al. 2009; Nieder and Dehaene 2009).

Studies from primate neurophysiology are assessing the neural foundation of the basic approximate numerical competence, showing that single neurons are able to encode the number of items in a visual display (Nieder,

Freedman et al. 2002; Nieder and Miller 2004; Nieder and Miller 2004; Nieder and Merten 2007; Nieder and Dehaene 2009; Nieder 2013) .

Recordings in monkeys trained to discriminate numerosities have revealed numerosity-selective neurons in the lateral prefrontal cortex (31% of all randomly selected cells), in the fundus of the intraparietal sulcus (18%) and a small proportion also in the anterior inferior temporal cortex (Nieder, Freedman et al. 2002; Nieder and Miller 2004) . These neurons showed a maximum activity to one presented quantity (which defined the neuron's preferred numerosity) and a progressive drop in response as the displayed numerosity changed from the preferred one.

Interestingly these number-selective neurons are showing a compressive logarithmic pattern of behavior with increasingly coarser encoding for high numerosities. This match with the observation that as the numerosity increases, a larger difference between the two quantities is needed in order to maintain the same discrimination performance.

This means that the number-selective neurons obey Weber's and Fechner's law*, a behavior that typically characterize the representations of sensory magnitude.

As a consequence it has been proposed that numerosity is a basic sensory visual dimension like colour, contrast, spatial frequency, orientation, size and speed. Supporting this idea it has been demonstrated that, as for the perception of these classical visual properties of the image, the perception of numerosity is susceptible to adaptation after effect (Burr and Ross 2008).

---

*The Weber's law stated that the just noticeable difference (JND) between two stimuli is proportional to the magnitude of the stimuli. The Fechner's law states that the subjective sensation is proportional to the logarithm of the stimulus intensity.

Visual after effect are distortions of perception produced by perceptual adaptation and they are found for dimension for which neural "detectors" are thought to exists.

However it has been argued that the perception of numerosity could derive from the perception of other visual cues, like the overall area or the texture density, defined as the number of elements per unit of area (Durgin 1995; Durgin and Huk 1997; Durgin 2008; Dakin, Tibber et al. 2011; Tibber, Greenwood et al. 2012). According to these authors numerosity is not sensed independently, but it is derived from texture density. Nevertheless other psychophisical studies from Ross and Burr (2010) together with a hierarchical generative model of number perception (Stoianov and Zorzi 2012) are demonstrating that selectivity to visual numerosity could develop naturally within visual neural structures, indipendently from texture perception (Ross and Burr 2012).

The second part of this thesis deals with how the human visual system processes numerosity and how adaptation effects can shed lights on its cortical organization, confirming its independency from texture computation.

2. Part I: Object segmentation

2.1 Introduction

2.1.1 Selectivity to spatial phase in luminance

David Marr was the first who tried to explain how the visual system can extract salient features to form a symbolic representation of the scene, creating a *primal sketch*. In this primal sketch the luminance intensity value at each point of the image should be computed separately at different scales by operators of different sizes. These operators are filters that compute the convolution between luminance intensity value and a Mexican-hat weight function obtaining a new function. The points where this new function change from positive to negative values or vice versa (zero crossing) indicate the points where in the original image luminance varied more, typically underlining an edge.

However simulation of this model gives rise to spurious features that do not correspond to the perceived position of the features in the scene. Attempts to minimize these artefacts are being criticised as they are often ineffective.

At the moment the most accepted model that provides an explanation of how the visual system could detect salient information of a scene is the local energy* model of feature detection of Morrone and Burr (Morrone and Burr 1988).

---

*Stimulus energy is formally defined as the integral of the power spectrum (the Fourier energy) of the stimulus.

In this model authors give a physical definition of visual features that are identified by peaks in local energy that occur in those points in the space where the Fourier components of the luminance profile tend to come into phase.

Infact, the boundaries that demarcate an object and distinguishing it from other objects and from the background are often associated with changes in the luminance profile and typically correspond to portions of the area where several harmonics have the same phase alignment determining edges or lines.

Hence salient spatial points in the luminance distribution of an image correspond to points where the various harmonics have the same phase congruency. The absolute value of the phase of the harmonics depends on the type of features, being 0° or 180° for bright and dark lines and ±90° for edges. Intermediate phase indicate the presence of both edges and lines.

The visual system achieves the detection of these important features in two steps. In the first stage a local energy function is computed from the sum of the squared responses of visual operators that are pairs of matched filters, with even- and odd-symmetry of identical amplitude spectrum but differing in phase of 90° (tuned respectively to lines and edges). Peaks in the output (local energy) mark all salient features indifferently (lines, edges and combination of them). In the second stage the features are then coded as a line or an edge (or both), depending on the relative strengths of the odd- and even-symmetric operators response. In this way by evaluating the relative response of the two operators (that correspond to evaluate the phase) it is possible to classify the type of feature. A strong even-symmetric response indicates a line while a strong odd-symmetric response indicates an edge.

Human visual system computes local energy separately over several orientation and spatial scales producing 'feature maps' that provide independent descriptions of brightness of the image. For most images the feature tends to be consistent at all scales and the feature maps are simply summed together.

The authors suggested also biologically plausible correlates to their model hypothesizing that the squaring operation corresponds to the activity of the complex cells of V1, whose receptive fields are usually modelled as the quadrature pair summation of two subunits with Gabor receptive fields at sine and cosine phase respectively. These Gabor subunits resemble the output of a pair of V1 simple cells with opposite contrast polarities.

The oriented organization of the V1 achromatic mechanisms' receptive fields would allow them to match with the luminance profile of lines or edges, depending on the even- or odd- symmetry respectively.

Electro-physiological investigations measuring visual evoked potentials (VEPs) support the hypothesis of the existence of odd-symmetric receptive fields as visual mechanisms involved in detecting the polarity of edges (Burr, Morrone et al. 1992). Jittered contrast reversed ramp stimuli were used making the local luminance at each point on the waveform varying randomly over time, in this way the phase-locked VEPs observed must result from the change in edge polarity and not from local luminance fluctuations. Moreover because the ramp stimuli were both leftward and rightward, the response of an hypothetical even-symmetric receptive field should not change at all with edge-polarity. Therefore the average over the time was expected to be constant while the response profile of an hypothetical odd-symmetric receptive fields would not have to change with edge polarity. This generates

the synchronized phase-locked VEPs that were indeed measured presumably originated from simple cells activity.

Studying the response of macaque V1 neurons to compound grating stimuli, Mechler et al. (2002; 2007) found that they were tuned to line-like, edge-like and intermediate one-dimensional features. Although the authors claimed that the model of Morrone and Burr was too simple, they reached the same conclusion that the phase coherence is already coded at V1 level. Infact they compared responses in V1 single cells for grating stimuli with different congruence phase and reported that both simple and complex cells were able to code the congruence phase. Although Mechler et al. suggested that the response of both simple and complex cells to Fourier components implies the presence of high order non-linearities (order≥3) and that features selectivity and specificity of individual V1 neurons strongly depend on speed, the phase of the harmonics remains the fundamental parameter to detect salient feature.

The importance of phase regularities (rather than the spatial power spectra) in defining a natural image was assessed also by Felsen et al. (2005) who measured response of cat V1 complex cells and found it higher for natural image than for random stimuli obtained by scrambling the phase spectra. However this phase sensitivity was not found for simple cells.

Many fMRI studies compared the sensibility to phase coherence in human and monkeys leading to apparent discordant results.

Rainer et al. (2002) investigated BOLD signal modulation associated with natural or scrambled images in macaque cortex. They found that BOLD activity in V1 increased with scrambling but then diminished for a very highly scrambled images. BOLD modulation in extrastriate areas (V2,V3,V3A,V4) was essentially unable to distinguish between natural and scrambled images

but also in these areas very highly scrambled images were associated with a decrease in activity. Finally in the superior temporal sulcus (STS), as well as in inferior temporal (IT), BOLD decreased systematically with increasing scramble level until a minimum for highly scrambled. Hence from this study it seems that the preference for natural images is established only later in visual analysis with primary visual cortex preferring slightly the scrambled images.

However in this study image scrambling was obtained by rearranging the position of the pixels within an image, thus leaving the overall identity of the pixels invariant. Scrambling images in such a way has the effect of introducing new edge and of adding high spatial frequency in the image and this effect, rather than a shift of phases that is of interest here, could account for these results. This is consistent with previous studies that show that V1 preferred stimuli with medium spatial frequency that correspond to the peak found in this article of Rainer et al. (2002) also if it corresponds to the scrambled image rather than to the natural one.

The preference for phase coherent stimuli was already found in the primary visual cortex by Rainer et al. (2001) who studied the BOLD signal modulation in anesthetized monkeys showing them natural images as opposed to degraded version of them obtained by scrambling the phase spectrum of various amounts. They found that natural images (100% phase coherence) generally elicited more BOLD activity than the noise pattern (0% phase coherence). This was true for the primary visual cortex, extrastriate visual cortex (V2, V3, V3A and V4) and for the anterior bank of the superior temporal sulcus (STS).

Olman et al. (2004) raised to a different conclusion showing that congruence in spatial phase had not an effect on BOLD fMRI response of V1

which seemed indeed to be influenced by the amplitude spectrum and hence by the contrast. But this study has two problems. The first is that they compared natural or scrambled images (pink and white noise) against an equiluminant background. In this way the response modulation driven by changes in phase spectrum is very small relative to that produced in the amplitude one. The second problem is that Olman et al. worked at high RMS contrast level where the local energy model of feature detection predicts a small or null differences between noise and coherent patterns.

This last problem influences also Tjan et al. (2006) experiment. These authors investigated the BOLD response change in function of the SNR (signal-to-noise ratio) in human subject that had to decide if two pictures were taken from the same scene. They used pictures embedded with different level of pink noise generated by perturbing the phase spectrum of a scene while maintaining mean luminance and RMS contrast as the original picture. The result showed that the modulation of the BOLD response by stimulus SNR increase from posterior (V3d, V3v/VP, V4) to anterior (V3a, LO, pFs -posterior fusiform sulcus-) ROIs with no noticeable effect on V1 and V2. They concluded that the BOLD responses in V1/V2 were not dependent on the presence of any forms or edges in the image because the BOLD response in these areas did not depend on the amount of noise embedded in the picture. The increase of pink noise affected most of all the BOLD response especially in two regions of the lateral occipital complex (LOC), i.e. LO and pFs, that are involved in the analysis of visual shapes and object recognition. Unfortunately the RMS contrast of the images used in this experiment had not been reported in this study and it is not possible to exclude whether the absence of the V1 BOLD modulation could be due to the saturating RMS contrast of the images.

The ability in perceiving edge conveys also the possibility to detect differences in the surface brightness perception. Infact the presence of edges generate brightness perception of the area they enclose or delimit. The brightness seems to be determined by channel tuned at 1 c/deg (Perna and Morrone 2007) with no need for pooling or integration across spatial channels.

This phenomenon was studied by Perna et al. (2005) who used the Cornsweet illusion (COC) to study the perception of brightness in humans. To generate the illusion they used edges containing only high spatial frequencies that enclosed a square that seems brighter or darker then the surround although they were of the same luminance. Infact occluding the edges abolished the perception of the lighter (or darker) surface in the center, demonstrating that the square was indeed illusory.

The authors compared BOLD activation to this stimulus with one where the edge was changed to line via the Hilbert transform and same RMS contrast. This last stimulus generated the perception of brightness and raised in depth.

Comparing both edges and lines against noise elicited a strong BOLD response in most visual areas, including V1. There was also a small negative activation of some regions within the central visual field representation (marked in blue in Figure 3).

The edge versus line stimuli produced no positive response in primary visual cortex or in most other visual areas, probably because the responses were identical and cancelled each other out.

Hence V1 responded equally to features with the same amplitude spectra and RMS contrast, but with different phase profiles and different Michelson

contrasts despite the perceptual difference in the apparent brightness of the surface.

There was a strong positive response to line stimuli in two associative areas within the dorsal pattern: one extending from the caudal region of the intraparietal sulcus (CIP, Talairach coordinates: ±31, -82, 25) up to the median portion of the transversal occipital sulcus without crossing it, the other extending more caudally and laterally, below the transversal occipital sulcus along the lateral occipital sulcus (LO, Talairach coordinates: ±43, -79, 11).

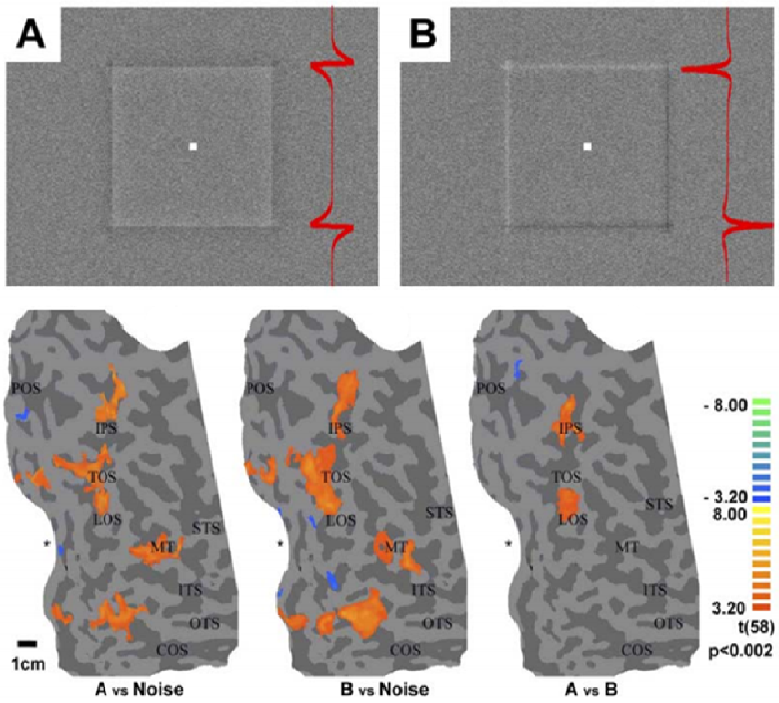

Figure 3: Examples of the stimuli used (A: edge stimulus; B: line stimulus) and of the BOLD activity measured in one subject when the edge stimulus was presented against only noise, the line stimulus against noise, and the edge stimulus against the line stimulus. The star represents the fovea. TOS, transverse occipital sulcus; LOS, lateral occipital sulcus; IPS, intraparietal sulcus; POS, parieto-occipital sulcus; ITS, inferior temporal sulcus; STS, superior temporal sulcus; COS, collateral sulcus. From Perna 2005.

The authors suggested the presence of non linear mechanisms in these two associative areas. Infact the response to edge pattern was only marginally stronger than the response to the line stimuli when presented against noise alone. Nevertheless, the direct contrast between line and edge showed a slightly smaller response than the response to edge or line versus noise alone, but far stronger than it would have been predicted from their differences (about half of the size of the individual edge or line responses).

In V1 there was a small but significant modulation in response to edge versus line stimulation, but the modulation occurred with each change in stimulus type, showing no preference for edge over line stimuli, as shown in Figure 4, bottom row.

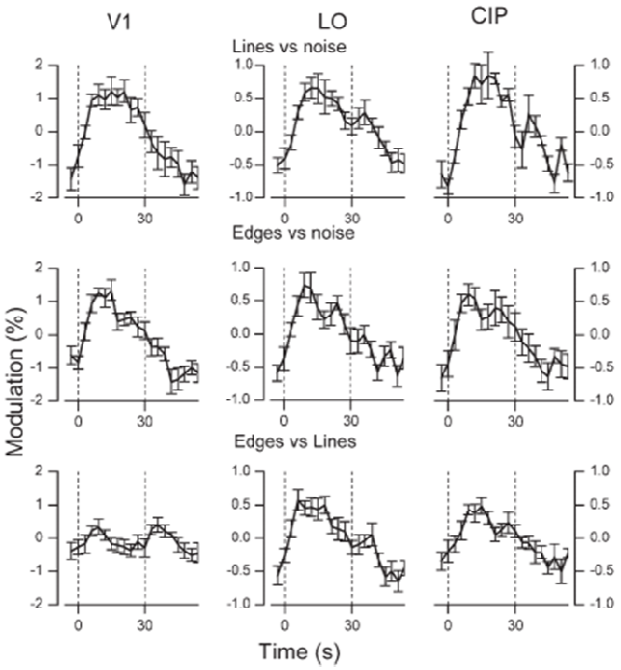

Figure 4. Time Courses of BOLD Responses in V1, LO, and CIP in response to edge versus noise stimuli (top row), to line versus noise stimuli (middle row), and to edge versus line stimuli (bottom row). The dashed lines show the time of the stimulus transition. From Perna 2005.

Further analysis in this study also established that the primary visual cortex only detects features and does not specifically encode the illusory brightness which is instead computed by the two associative areas LO and

CIP that probably have a rule in the 3D reconstruction of the volume and space.

Congruent results were found by the same authors in a more recent (2008) experiment where they compared the BOLD response to periodic band-pass images of matched amplitude spectrum and same RMS contrast but with different phase spectra and different Michealson contrast. They created two phase congruent stimuli: pure edges and pure lines that had same global and local energy but phase shifted of 90 deg being one the Hilbert transformation of the other. Moreover they created a non coherent phase random noise stimulus with different spatial distribution of local energy and in which no salient features could be detected because the Fourier components were summed together with random phases.

Alternation between coherent and non coherent phase stimuli activated several areas including V1 and many extrastriate areas (Figure 5).

The comparison between congruence phase stimuli did not elicit activity in V1 nor in other early retinotopic visual areas while a differential activation emerged from two higher areas found in the precedent experiment of Perna (2005): one in the dorsal LOC, the other in the more caudal part of the occipital branch of the intraparietal sulcus (CIP).

The authors also tested the dependency of BOLD response on the RMS contrast of the stimuli. The result showed that at low and medium contrasts V1 and dLO-CIP response was higher for coherent phase stimuli respect to random noise, but at a higher contrast the V1 response for edge and line decreased, while the one of the two higher order areas remained stable. The V1 BOLD response decrement probably reflects the response at single neuron level that saturates at that high contrast level.

To simulate such response the authors needed to correct the computation of the local energy with a subsequent non-linear gain stage. The local energy at each image location is obtained by the sum of the outputs of the even- and odd-symmetric receptive field for each given orientation and spatial frequency.

But if the BOLD response correlates with the average energy response given by the integral of all local energy values weighted with their frequency, since for coherent stimuli the most frequent energy value is zero, the model predicts a higher response to the random pattern.

So to simulate the V1 response the neural activity should be related to the local-energy of the visual stimuli through a gain function:

$$\underset{\theta,\omega}{R} = \alpha \cdot \sum_{S} \frac{E_S^{\beta}}{E_S^{\beta} + g_0^{\beta}},$$

where R is the neural activity at a certain spatial scale and orientation, Es is the local energy, g0 is a constant related to the saturating property of the response neuron and α and β are parameters. Different value of β are tested (different dotted lines in Figure 5) concluding that β=3 can fit well the V1 BOLD signal versus contrast (continuous curve). This gain function introduces a non-linearity (whose strength is given by the parameter β) that amplifies the value of Es (peaks are higher) allowing to the average value of the energy to be higher for the congruent phase stimuli than for the random phase stimuli and reproducing in this way the V1 BOLD response.

This model of neural response that involve the computation of local energy and a non linear response gain function was not satisfactory for the dLO+CIP response because these areas did not show dependence on contrast and probably performed a qualitatively different analysis, like structuring the image from the individual feature.

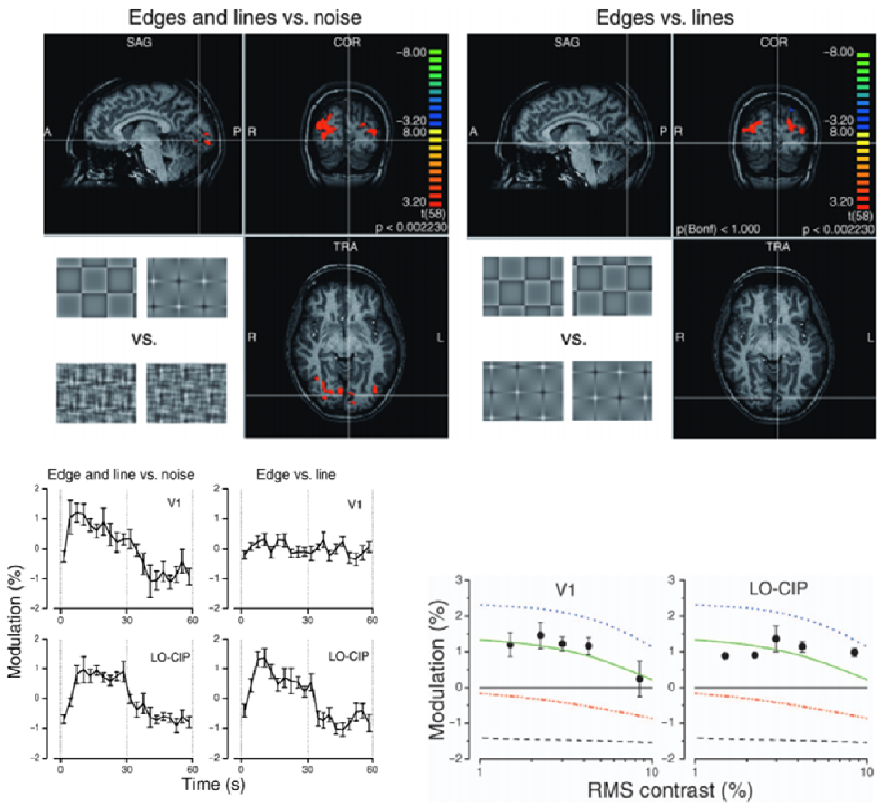

Figure5: Pattern of activity elicited by the alternation of edge and line vs. noise stimuli (left) and by the alternation to edge vs. line stimuli (right) in one subject.
The graphs on the left show the time courses of BOLD responses for V1 (upper row) and the union of the dLO and CIP (lower row). The left column shows the response to edge and line vs. noise stimuli, the right column to the alternation of edge versus. line stimuli. The dashed lines show the time of stimulus transition.
The graphs on the right show the amplitude of BOLD-signal modulation in V1and dLO and CIP as a function of RMS stimulus contrast for the alternation of the edge and line vs. noise. The lines are predictions of a local energy model with a non-linear response gain given by the equation. The curves correspond to different values of the β exponent equal to 1, 2, 3 and 4 from bottom to top. The continuous curve correspond to β=3.

It is important to note that in these last two studies artificial stimuli were used where the phase congruency was total and this condition is difficult to find in otherwise natural images. Moreover the stimuli of Perna et al. (2008) did not contain low spatial frequencies that are instead present in natural images and that are less influenced from phase congruency than higher spatial components. So these differences, other than the RMS contrast dependence, could account for the absence of preference of natural images

respect to noise in previous studies. Finally in some of those studies natural images were presented against a blank field causing a large contrast jump that may amplify non-linear components of BOLD response masking the selectivity for phase congruency that is instead appreciable in the Perna et al. study where contrast of stimuli was constant through each scan.

Perna et al. (2008) results are supported also by Henriksson et al. (2009) experiment which used fMRI to explore phase-sensitive pooling of spatial frequencies. They presented grating patches simultaneously in all four quadrants at 7.6° eccentricity while subject performed a fixation task. In the first part of the experiment the subject was adapted to a grating stimulus composed by the fundamental and the third harmonic (with 0° or 180° phase difference), then examined with test stimuli in which the third harmonic was relocated 0°, 45°, 90°, 135°, 180° with RMS contrast kept constant to 14%. The change in phase difference activated all visual areas V1, V2, V3AB, hV4, LO, pFus and intraparietal sulcus. Nor difference in local Michelson contrast (caused by the progressive shift between the two spatial frequency components) nor difference in absolute phase (position) or of the third harmonic can explain this result. The second part of the experiment was a block design where fMRI response to congruent (ranging from 0° to 90°) and random phase stimuli were compared. Activation maps showed that gratings with congruent phase alignment evoke stronger responses in V1, V2, V3AB, hV4, LO, pFus and intraparietal sulcus. Moreover the authors computed a selectivity index for each area from the response to congruent stimuli showing that the index increase through the ventral streams areas V1, V2, V3, hV4, LO pFsu and also in the intraparietal sulcus.

Hence also this experiment shows that several visual areas, including V1, are sensitive to spatial phase difference in luminance grating stimuli and

probably the phase-specific spatial frequencies pooling starts already in V1, phase congruence across multiple spatial frequencies are identified later in higher-level visual areas. Indeed the selectivity of these areas for congruent phase structure increased along the ventral stream and is high along the intraparietal sulcus.

2.1.2 Selectivity to spatial phase in equiluminance

All of the studies mentioned until now have been concerned with lines and edges defined by changes in luminance. However, luminance contrast is not the only source of spatial information in a visual image. Another rich source of information is provided by colour.

The traditional view of colour processing considers that V1 chromatically opponent cells, as those of the retina and NGL, have a concentric receptive field without a well defined orientation selectivity and arguing that this is not surprising given the poor sensitivity of human perception to colour contrast borders.

However this view has been questioned by the recent finding of orientation selectivity for chromatic V1 cells.

Indeed recently Johnson et al. (2008) investigated the neural basis of form-colour link in macaque primary visual cortex (V1) by studying orientation selectivity of single V1 cells for pure colour patterns showing surprising results.

In their previous experiments (2001) the authors classified V1 cells population into three groups: colour-preferring, colour-luminance and luminance-preferring (depending on the stimulus that elicited the cell's higher response) on the basis of a colour sensitivity index *I* (*I* = max_response(equilum)/ max_response(lum)). The main characteristic of colour-luminance cells were their spatial frequencies tuning with equally selectivity for chromatic and achromatic patterns.

However when they studied (2004) the cones inputs to these three different classes of cells, they found that some colour-luminance cells received the same sign of input from L- and M- cones and therefore could be

colour blind. Moreover they found colour-preferring cells, that should be single-opponent neurons, were instead spatially tuned for equiluminant patterns thus resembling more colour-luminance cells than properly colour-preferring cells.

So the authors propose in their recent article (2008) a new partition of the V1 population into: single-opponent cells, double opponent cells and non opponent cells. This new classification is based on the colour sensitivity index and on the spatial tuning for colour and luminance stimuli.

Johnson et al. (2008) found that V1 single-opponent cells (similar to LGN parvocellular cells) responded both to red-green gratings and to luminance contrast and that they were not orientation selective for colour patterns. V1 Double-opponent cells had bandpass spatial-frequency tuning and were orientation selective to pure colour stimuli as well as to achromatic patterns. Both V1 complex and simple cells can be double-opponent. Finally V1 non-opponent cells did not respond to pure colour, but gave strong responses and were orientation selective for luminance patterns. Unlike double-opponent cells, the non-opponent cells could subserve high spatial acuity, be sensitive to luminance contrast, and be able to detect also coloured boundaries regardless of the configuration of the colour forming the boundary. Both double-opponent and non opponent cells were not contrast invariant: orientation selectivity increased with contrast. When stimuli were matched in average cone contrast, double-opponent cells were equally orientation selective for luminance and equiluminant colour pattern.

Double-opponent neurons are the neural basis of colour contrast. They can be distinguished between L versus M+S, forming red-cyan axes, and S versus L+M, forming blue-yellow axes.

V1 population is composed by 10% single opponent mostly in layer 2/3 and 5, 30% double-opponent in layer 2/3 and 6 and 60% non opponent in layer 4B and 6. The small percentage of double opponent cells in the cortex matched with the poor spatial resolution of colour vision.

The authors proposed that combined activities of single and double opponent cells are needed for the full repertoire of colour perception and that V1 double-opponent cells could be the neural basis of the influence of form on colour perception.

Single-opponent cells could signal colour regions, whereas double-opponent cells could signal colour boundaries. This could be possible thanks to the particular structure of receptive fields: single-opponent red-green sensitive neurons have center-surround symmetry, while this is not the case for double-opponent cells that have an oriented spatial symmetry that resembles the asymmetric or odd-symmetric spatial receptive fields of non-opponent cells.

Single-opponent red-green sensitive neurons receive inputs from L- and M-cones opposite in sign, but within each cone type signals are all of the same sign; while double-opponent simple cell spatial receptive fields are composed of subregions. Within each subregion, the L- and M-cones send signals opposite in sign and not balanced in strength.

A: single-opponent cells B: double-opponent simple cells

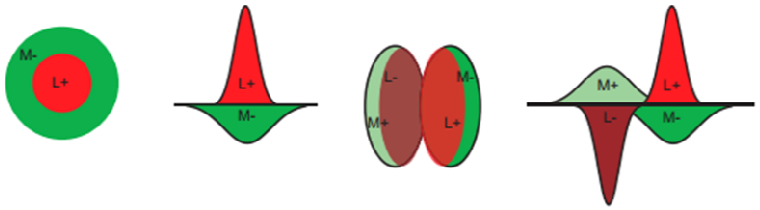

Figure 6: Models of the hypothetical spatial sensitivity profile for single-opponent and double-opponent V1 neurons. A: concentric, single-opponent red-green sensitive neuron, adequate also for LGN neurons. B: proposed sensitivity profile for an orientation selective double-opponent simple cell. From Johnson et al. 2008

The asymmetrical spatial structure of double-opponent receptive fields was already inferred by Girard and Morrone (1995) from their results on human visual evoked potential (VEPs) to gratings modulated in either luminance and red-green colour.

Varying the Fourier phase of the harmonics from 0 deg to 90 deg the authors produced a family of stimulus profiles that varied from lines to edges. As reliable VEPs were recorded in response to polarity alternation of both luminance and chromatic stimuli at all phases (with precautions that the potentials did not arise from changes in local luminance), the authors suggested that the mechanisms sensitive to chromatic contrast and those sensitive to luminance contrast have both symmetrical and asymmetrical receptive fields useful to detect edge and line. Thanks to the random jittering of the stimuli (synchronized at a higher frequency then the reversal frequency) used in this experiment it was possible to ensure that the VEPs were generated by edge or line reversal rather than by changes in local luminance or colour. The underlying logic was the same as the Burr et al (1992) study: if only one class of receptive field would exists with a specific phase spectrum, one would expect a null in the amplitude response for stimuli of certain spatial phases (equal in all except for their contrast). For example an equiluminant red-green ramp stimulus will excite all types of chromatically opponent units, independently of the symmetry of their receptive field. When the stimulus reverse in contrast all these units will change their responses, but the even-symmetrical receptive fields will give an equal and contrary response (and hence null) and the VEP response over time would not change if only even-symmetric receptive field exist. But because the author measured a strong and reliable response over time over a large range of contrast, they must conclude that chromatic receptive fields

should exist even if not even-symmetric. A similar reasoning can be applied to bar stimuli.

As the ratio between detection and discrimination threshold was the same for the luminance and the equiluminance patterns, the phase discrimination for chromatic stimuli should be as good as for luminance stimuli.

There were no appreciable differences in the contrast threshold inferred by the contrast response curve to edge or bars suggesting that the two mechanisms (respectively even- and odd-symmetric receptive fields) have the same sensitivity. Moreover this result suggests that the VEP response is directly related to the overall power of the stimuli, rather than to the local Michaelson contrast that is very different between bar and edge. Stimuli that have the same overall spatio-temporal power produce the same VEP response, irrespective of its spatial profile. One way for this to occur would be for neural mechanisms to compute a "local energy function" that probably corresponds to the activity of the complex cells type for luminance that may correspond to the V1 double opponent cells found now by Johnson et al. (2008).

V1 double opponent cells have bandpassed spatial-frequency tuning and are orientation selective to pure colour stimuli as well as to achromatic patterns. This is in agreement with Martini et al. (1996) results that suggested that similar sorts of mechanisms would operate to analyse edge and line both in luminance and in equiluminance. Indeed they reported that visual system has the same sensitivity for both luminance and chromatic (red-green) multi-harmonic pattern once scaled for the contrast. Indeed when asking subjects to perform a detection and phase discrimination task they showed that once contrast has been equated for detectability, phase discrimination was as good for chromatically modulated as for luminance-

modulated pattern. Moreover, also the thresholds for phase discrimination are similar at all spatial frequencies once taken into account the difference in contrast sensitivity between the two mechanisms. Sensitivities to multi harmonic patterns increased progressively with contrast in a similar way both in colour and in luminance. However there was an exception for the patterns of two harmonics (f + 3f) of moderate to high spatial frequencies. In this condition there was a constant advantage for luminance patterns over the colour one by a factor of about two.

But apart from these exceptions the results reported by these authors strongly suggest that phase discrimination is as good for chromatic as it is for luminance, provided they are scaled in contrast to equate for detection sensitivity. Hence as sensitivity for phase discrimination did not vary for most of pattern type, the authors inferred that both luminance and chromatic pathways should comprise detectors with both even- and odd- receptive fields. The fact that the thresholds for chromatic phase discrimination were as good as those for luminance suggests that the receptive fields of chromatic pathway are similar in shape to those of the luminance pathway.

Many studies in literature suggested that many aspects of vision are degraded at equiluminance. However most of the reduced performance can be explained by the lower cone contrast due to the partial superimposition between the cones absorption curves. After equating for cone contrast performance of the two systems are comparable (Martini, Girard et al. 1996).

All these experiments are now supported by the finding of Johnson et al (2008) that V1 double-opponent cells have bandpass spatial-frequency tuning and are orientation selective to pure colour stimuli as well as to

achromatic patterns and thus could be the neural basis that support the line and edge extraction from a complex visual scene.

2.1.3 Present Study

In the present study we aim to determine the level in the colour pathways at which the selective responses to phases emerge, and whether the colour system, as the luminance system, shows a BOLD preference for phase-congruent stimuli.

The results suggest equal sensitivity to random- and congruent- phase equiluminant patterns in V1, and a preference for equiluminant odd-symmetry features (absolute phase around ±90°) along both the ventral and dorsal pathways, implicating these pathways even in the analysis of equiluminant feature and surfaces.

2.2 Materials and Methods

2.2.1 Visual stimuli

Three different kinds of stimuli were used in the experiment. There were two experimental congruent-phase stimuli (which we label even-symmetry and odd-symmetry stimuli) and a control random phase stimulus (noise stimulus), as shown in Figure 7.

The congruent-phase stimuli were obtained by multiplying two orthogonal vertical and horizontal gratings synthesised from 19 odd sine harmonics (first, third, fifth ⋯ nth) of 1/n amplitude. For the local odd-symmetry stimuli the 19 harmonics all had the same absolute phase, equal to p/2 (square wave), while for the local even symmetry stimuli it was equal to zero. To minimise chromatic aberrations, the high spatial frequencies were

attenuated with a Gaussian filter in both the colour- and the luminance-modulated stimuli.

$$F(\omega_x, \omega_y) = e^{\frac{-(\omega_x^2 + \omega_y^2)}{2\sigma^2}}$$

where r = 5.5 cycles/° was applied to both the ωx and the ωy frequency axes.

For all 2-D stimuli, the amplitude of the fundamental harmonics (spatial frequency: 0.37 cycles/° and 45° orientation) were halved to reduce the relative contributions of the first harmonics to the response. To reliably measure phase selectivity it is important that each visual channel responds simultaneously to many harmonic components, and this would not be the case for stimuli comprising predominantly first harmonics (for details see Burr, Morrone et al. 1989).

The 2-D odd-symmetry stimulus comprises only edges, and looked like a checkerboard (Fig. 7A and D) while the 2-D even-symmetry stimulus looked like a grid of thin lines that form stars (Fig. 7B and E). The even and the odd stimuli at each point had the same local root-mean-square (RMS) contrast and by construction had the same local phase congruency between harmonics and also the same number of salient features, being 1-D edges, 1-D lines, 2-D stars or 2-D edge junctions. The local phase congruency can be estimated by computing the local energy (Morrone and Burr 1988) which for these patterns is given by the square root of the sum of the square of the even- and odd-symmetry functions (see Morrone and Burr 1997). Figure 7 shows the result of the local energy function for a particular row of the 2-D stimuli: the energy (and hence the local RMS) peaks along the major horizontal and vertical salient features. To a first approximation, the linear response of a V1 population of simple cells with even-symmetric and odd-

symmetric RFs will be maximum for the even-symmetry stimuli and odd-symmetry stimuli respectively. The response of complex cells, with ON-OFF subunits, will approximate more closely the local energy function (Hubel and Wiesel 1962; Movshon, Thompson et al. 1978; Movshon, Thompson et al. 1978; Pollen and Ronner 1981; Spitzer and Hochstein 1985).

For the random-phase stimulus, the phase of each harmonic of the 2-D congruent-phase stimuli was shifted by a random value. This procedure generated 2-D stimuli with equal global power spectra and equal global RMS contrast. Examples of the random 2-D noise stimuli of matched power spectra are shown in Figure 7C and F.

These stimuli did not have strong salient features and the local energy function was more distributed in space with small amplitude peaks.

RMS contrast, defined as SD of luminance divided by the mean, was equal for all these stimuli to within 5.16%. However, luminance Michelson contrast (defined as the difference between the maximum and minimum luminance divided by their sum) was very different, being 17% for edge stimuli and 86% for line stimuli. The noise stimuli had on average intermediate Michelson contrasts of 30%.

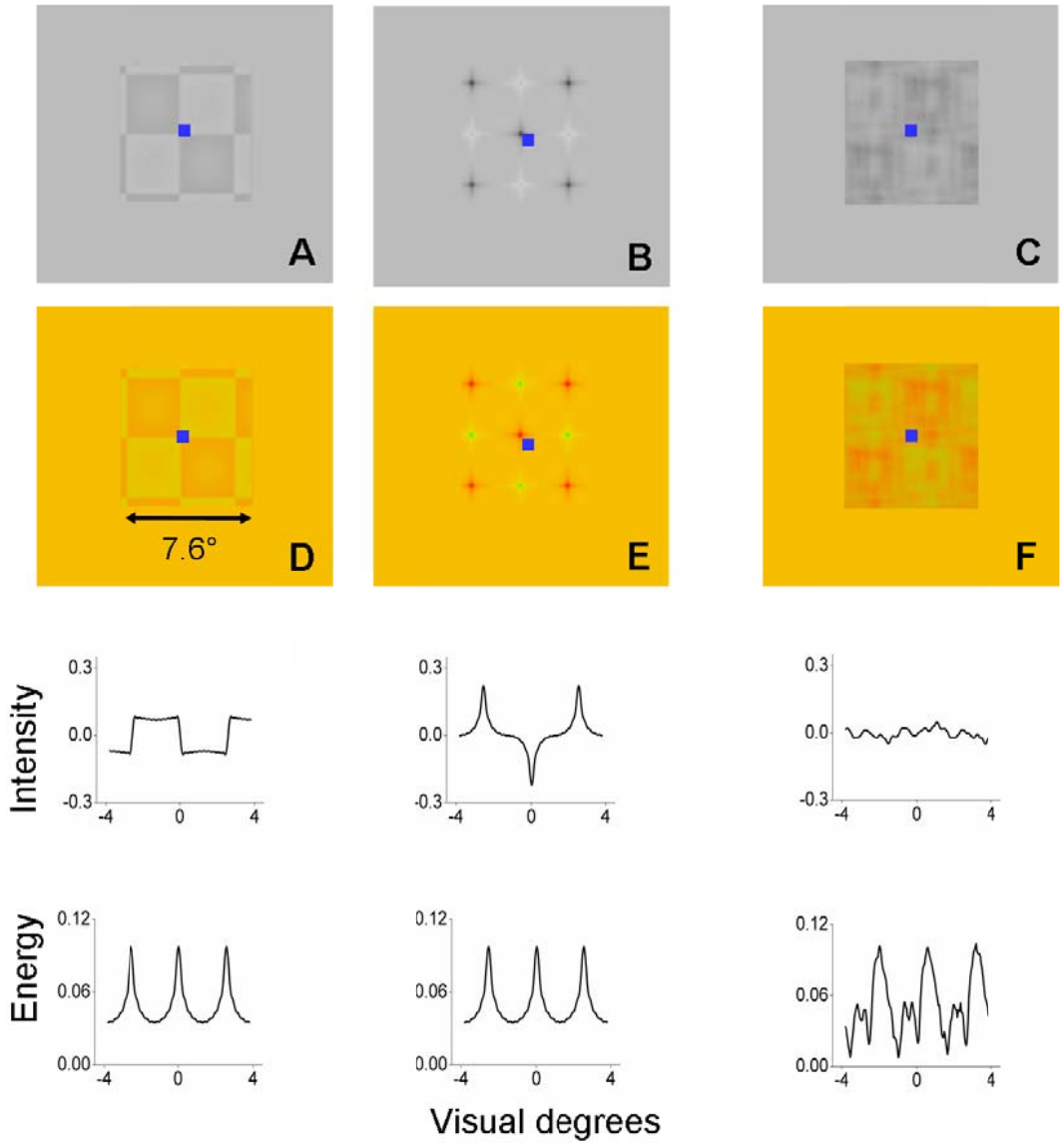

Figure 7. Visual stimuli. Examples of the visual stimuli. Red and green gun modulation were summed in phase to create luminance checkerboards (top row) and in anti-phase to create the equiluminant checkerboards (bottom row). (A and D) Example of the odd-symmetry stimulus: two-dimensional filtered checkerboard, comprising 2 X 2 periods each subtending 3.8 X 3.8° of visual angle. (B and E) Example of the even-symmetry stimulus where each individual edge (odd -symmetry feature) of A and D was transformed in a feature with local even symmetry. (C and F) Noise stimuli obtained by randomising the phase of each harmonic. All stimuli had the same RMS contrast but different Michelson contrast. The even- and odd-symmetry image pairs always had the same RMS local contrast for each individual spatial position, despite the illusory different appearance. The integral of the power over space was equal for the three stimuli, while the Michelson contrasts were equal to 17%, 86% and 30% for the odd and even symmetry and the random stimulus respectively. The lower traces show examples of the luminance modulation and the output of the local energy function for the same row of the images.

The stimuli were modulated in luminance by driving the red and the green monitor channels in phase and setting to zero the blue channel. The equiluminant (red‑green) checkerboards were obtained by modulating red and green monitor channels in anti-phase. The Commission Internationale de l'Eclairage coordinate for the red and green guns were x = 0.665, y = 0.312 and x = 0.244, y = 0.55 respectively. The CIE coordinates for the average yellow background were x = 0.374, y = 0.479. The mean luminance was 42 cd/ m². Equiluminance subjective point was assessed for each scanning session and each subject inside the scanner by flicker photometry.

At equiluminance the cone contrast (RMS of L and M cone contrast) was a factor of 1.8 lower than the luminance contrast.

2 X 2 periods (each subtending 3.8 X 3.8° of visual angle) of the 2-D stimuli were displayed. Because the equiluminance point changes with eccentricity, we restricted the stimulus to a central window of 7.6° X 7.6° (237 X 237 pixels, windowing with no edge-smoothing to display complete stimulus periods) while the remaining area (covering the 800 X 600 pixel resolution monitor) was kept at constant chromaticity and luminance. Inside this window, the coherent-phase stimuli were jittered to a random position every 2.5 s to avoid BOLD adaptation; for the non-coherent phase stimuli a new draw of random phase was selected every 2.5 s and the new stimulus synthesised and displayed. Stimuli were designed using Matlab 7.5.0 R2007b with Psychtoolbox (Brainard 1997) installed on an Asus F3JA notebook computer with an integrated 32-bit precision graphics card (ATI Radeon X1600) and 60-Hz temporal refresh rate.

The computer was connected with a Polaroid Polaview 220 DLP projector with a lens that projected from 7.5 m to a screen inside the scanner bore 45 cm from the observer’s eyes. The projected image was gamma-corrected to

ensure linearity. We verified linearity by direct luminance measurements at the screen.

In an additional experiment in which we measured the response for various red/green ratios (CIE coordinates for the red and green stimuli were x = 0.577, y = 0.357 and x = 0.292, y = 0.537 respectively and mean luminance 48 cd/m²), we displayed the stimuli through NordicNeuroLab's VisualSystem goggles (800 X 600 pixels, with a refresh rate of 85 Hz) after adequate gamma correction.

The colour‑luminance contrast ratio of the pattern was varied by changing the modulation of the green, keeping constant the modulation of the red unchanged with colour ratio. A ratio of 0.5 indicates that the luminance modulation of the green gun was half that of the red gun.

2.2.2 Retinotopic maps

Retinotopic mapping and localiser scans were performed on each subject in a separate experimental session.

To identify the representation of vertical and horizontal meridians one hundred circular dots (0.3° diameter), half black and half white, on a grey background, moved within two symmetrical sectors across the fixation point along the two principal meridians (for details see d'Avossa, Tosetti et al. 2007; Crespi, Biagi et al. 2011) . Each dot had a 20-frame lifetime at refresh rate of 60 Hz (333 ms). Locally, the dots moved along linear trajectories at a constant speed of 10°/s; globally they were perceived to follow an expanding or contracting movement along the vertical and horizontal meridians spanning 20° of visual angle. When each dot's lifetime expired, or its trajectory ended outside the sector, the dot was regenerated in another

position randomly chosen in order to maintain a uniform density of 0.44 dots per square degree of visual angle.

To activate selectively the representation of the upper, lower, right and left visual hemifields, 250 dots were displayed with radial motion (same dimensions and lifetime as for the meridian stimulus) within each circular sector of ± 40°, symmetric around the diagonal of the quadrant.

Each meridian or hemifield stimulation lasted 15 s, with motion direction inverted seven times to avoid BOLD adaptation. Each block was repeated six times. Identification of the different retinotopic maps was conducted on the basis of the upper-lower visual field representation and on the vertical-horizontal reversal, in accordance with standard retinotopy literature (Sereno, Dale et al. 1995; Engel, Glover et al. 1997; Hadjikhani, Liu et al. 1998; Press, Brewer et al. 2001; Tootell and Hadjikhani 2001; Wade, Brewer et al. 2002; Wandell, Brewer et al. 2005; Hansen, Kay et al. 2007). To localise V4 we used the criterion proposed by Hansen et al. (2007), which considers the representation of part of the lower visual field in the dorsal cortex abutting V3d, in addition to standard retinotopic localisation.

The CIP and transversal occipital sulcus (TOS) were defined as anatomically located along the specific sulci and by including all the voxels that were active for the alternation of blank against even-symmetry and odd-symmetry stimuli, modulated either in luminance or in colour. Usually the selected area was about 4 X larger than the localiser, if this did not interfere with an adjoining region of interest (RoI).

The human Middle Temporal Complex and dLO were localised using 100 moving dots (speed 10°/s, 0.5° diameter). These dots moved randomly for 21 s and then followed a coherent spiral movement (flow motion) changing from expansion to clockwise, contraction and anti-clockwise circular motion

twelve times in 21 s. These dots covered a circular patch of 15° diameter. Each dot had a limited lifetime of 10 frames (166 ms), after which it disappeared to be reborn in a new random position. Each acquisition sequence comprised six presentation for the coherent and six for the incoherent motion.

To localise the region of interest representing the stimulus patch within the retinotopic visual areas, a blank stimulus of the same chromaticity was compared against the two experimental congruent phase stimuli. The border of the activity ($P < 0.001$) elicited by this localiser was then traced in V1.

2.2.3 Subjects and Procedures

Seven healthy adults (six females, one male) with normal colour vision and normal or corrected-to-normal acuity were the subjects of the experiment. For each subject the equiluminant point was measured independently inside the scanner using flicker photometry, displaying a red-green sine wave grating flickering at 7.55 Hz and varying the modulation of the green gun keeping the red gun modulation fixed (the subjective equiluminant point defined by the ratio of the two gun modulations was always between 0.52 and 0.54).

Informed written consent was obtained for each subject prior to scanning sessions, in accordance with the guidelines of the MRI Laboratory, Fondazione CNR/Regione Toscana G.Monasterio, Pisa.

The studies were approved by the ethics committee of the Azienda Ospedaliero-Universitaria Pisana (protocol number 3255, approved on 20/01/2009) and was in accordance with the ethical standards of the 1964 Declaration of Helsinki.

fMRI measurements were performed on a 3-T scanner (GE Excite HDx) equipped with an eight-channel brain coil. Each session began by acquiring a set of anatomical images of whole brain with T1-weighted contrast. A fast spin-echo sequence was used with TR, 10.8 ms; TE, 4.9 ms; flip angle, 13°; FOV, 256 9 256 mm2; slice thickness, 1 mm; bandwidth, 15.63 Hz; and an isovoxel matrix size of 256 X 256. The data acquisition time was 5.56 min.

Echo planar imaging (EPI) sequences were used for the fMRI data acquisition (TR, 2.5 ms; TE, 40 ms; FOV, 240 X 240 mm; flip angle, 90°; matrix size, 128 X 128; and slice thickness, 4 mm; 96 volumes).

In a block design we presented four different stimuli sequences in random order. In two sequences, one modulated in luminance and the other in colour contrast, the random-phase stimuli (OFF-block) were compared against two congruent-phase stimuli (ON-block, with the even- and odd-symmetry stimuli selected randomly at equal probabilities). In the other two sequences, one modulated in luminance and the other in colour contrast, the two congruent-phase stimuli were compared with each other: even-symmetry stimulus (OFF-block) against odd-symmetry stimulus (ON-block). Each subject performed from two to four scans per condition. We also ran a localiser in which a blank equiluminant field of the same chromaticity (OFF-block) was compared with the two experimental congruent-phase stimuli (ON-block). The OFF- and ON-blocks alternated six times and each block lasted 20 s. Within each block the stimuli jumped from one random spatial position to another every 2.5 s in order to avoid BOLD-signal adaptation. For all conditions, we asked subjects to fixate on a stable target that remained in the centre of the visible screen for the entire scan duration. Head movements were minimised by padding and tape.

A non-commercial software package, 4DFP, from Neuro-Imaging Laboratory (Washington University, St Louis, MO, USA) was used for the pre-processing of the imaging data. To compensate for systematic slice-dependent differences in acquisition time, functional data were temporally interpolated and re-sampled; slice intensity differences caused by interleaved acquisition were eliminated. Overall image intensity was normalised within scans to a standard value of 1000 to compensate for inter-scan intensity differences. Functional images were corrected for three-dimensional motion by realigning data within and across scans and sessions with reference to the first realigned volume of the first scan, using a six-degrees-of-freedom rigid body transformation.

BOLD images were spatially normalised according to the atlas of Talairach & Tournoux (1988) to obtain standardised coordinates for the RoI. Lastly, data were spatially re-sampled to a cubic voxel with a resolution of 2.0 mm2.

Brain Voyager QX (version 2.2 Copyright © 2001‑2011; Rainer Goebel, Brain Innovation B.V., Maastricht, The Netherlands) was statistical analysis. To generate flat maps for each hemisphere the white‑grey boundary was traced, using an automatic segmentation algorithm supplemented by manual correction by an expert operator to correct errors generated by the automatic routine. This segmentation was also used to automatically reconstruct the surface of the outer grey-matter boundary, which was subsequently inflated and flattened by geometric projection. For all scans an initial qualitative analysis was performed applying a general linear model to a boxcar model convolved with a canonical hemodynamic response function (HRF). We marked significant variations in response with a threshold set to $P < 0.02$ and cluster size to 3 X 3 voxels. This permissive threshold was selected

to mark even noisy response variation in V1. This analysis was used to determine the Talairach coordinates of the foci activities reported in Table 1.

For primary visual areas the average response was modulated to the second harmonic, which could not be revealed by general linear model analysis without changing the boxcar model. For quantitative analysis the response was evaluated by averaging the BOLD signal in anatomical areas defined on independent criteria. For each RoI we averaged the BOLD time-courses across subjects for the entire sequence of stimuli alternations and runs. The first and second harmonic amplitudes of each average profile were then computed by projecting them along the first and second harmonic phase of a modelled response, calculated by convolving the boxcar stimulus sequence with the canonical HRF. A sign test was performed on the projected harmonic amplitude by bootstrapping (Efron 1993), resampling the individual runs (with replacement) and repeating the process 2000 times, for all the areas and all stimulus conditions.

2.3 Results

It is always difficult to compare directly the activity elicited by colour-and luminance-modulated stimuli because the contrast and the visual salience of the two sets of stimuli are different, even when equated for detectability. The best strategy is to study independently how the phase manipulation affects the BOLD response. We used stimuli that were equated for visibility so as to equate average activity to the colour- and luminance-contrast modulated stimuli.

Figure 8 illustrates qualitatively the overall effect of phase manipulation on the visual stimuli in one subject. Figure 8A and B reports the data for luminance-modulated stimuli and Figure 8C and D for equiluminant stimuli. The individual cortical areas are segmented on retinotopic criteria and are relatively large with respect to the activation elicited from our small stimuli.

When modulated in luminance (Fig. 8A), the alternation between random-phase and congruent-phase stimuli elicited a change in activation in primary visual cortex. The response was located inside the region representing the central 7° of visual field, revealed by alternating congruent-phase stimuli against blank fields of matched mean chromaticity and luminance. The dashed lines in V1 mark the peak border activities of this localiser with $P < 0.001$, while the response modulation to the luminance congruent-phase stimuli covers the foveal representation (indicated by a star in Fig. 8A). This result is consistent with the central activity measured by Perna et al. (2008) despite the more extended area of their stimuli (15°). However, the primary visual cortex modulation of activity was no longer present when the same stimuli were modulated in colour contrast (Fig. 8C), despite the low threshold used ($P < 0.02$).

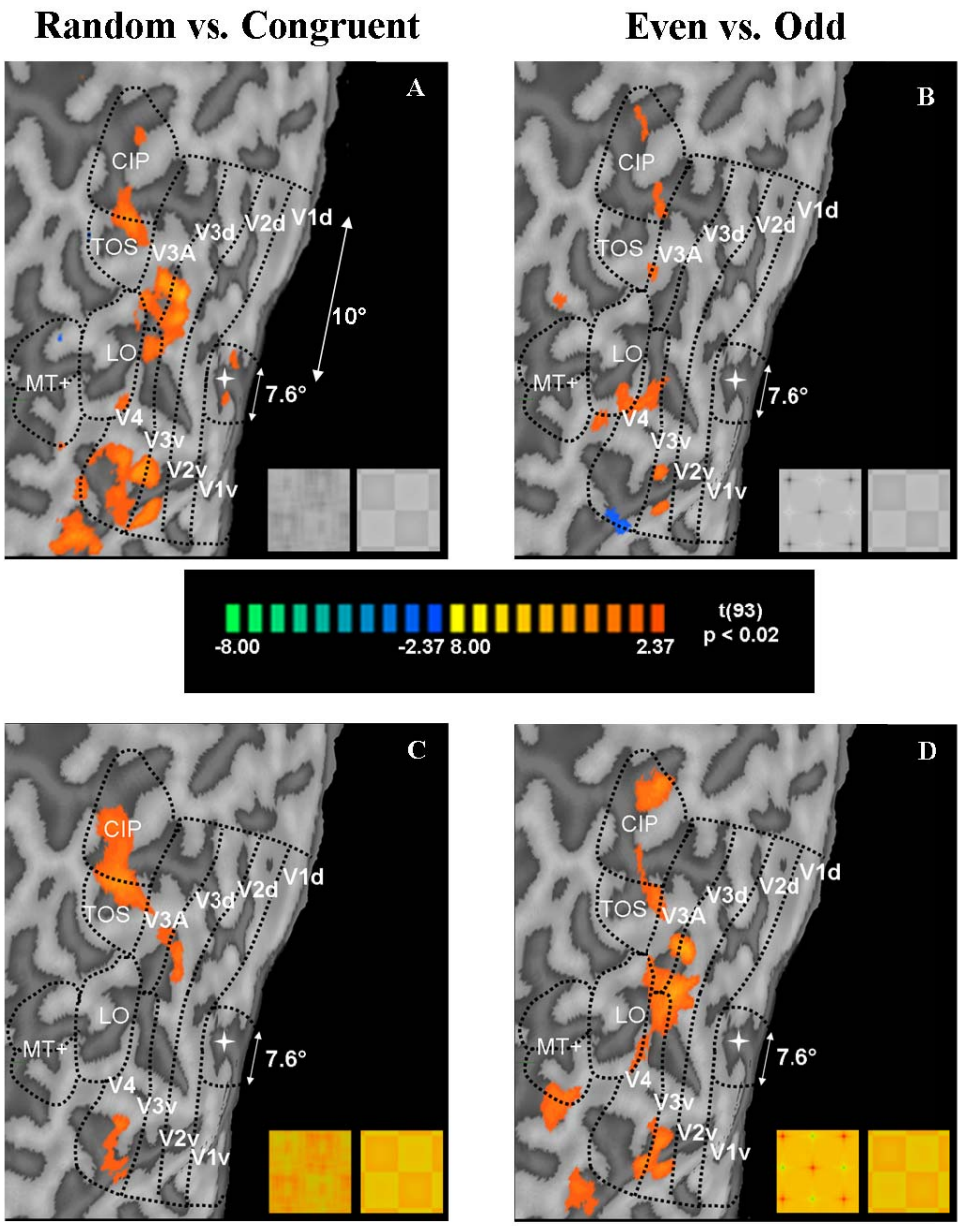

Figure 8. Activations on flat map for subject S1. Typical patterns of activation elicited by contrasting random-phase stimuli against congruent-phase stimuli (left) and contrasting even-symmetry against odd-symmetry stimuli (right) both for luminance (upper) and for colour (lower) conditions, displayed on the flat map of subject S1, with $P < 0.02$. Contrasting random against congruent-phase stimuli elicited activation changes in the primary visual cortex only for the luminance condition. The modulation was not present in any of the other stimulus conditions. Both dorsal and ventral pathway activities were strongly modulated by the equiluminant stimuli as well as by luminance-contrast stimuli.

The alternation of random- against coherent-phase stimuli, modulated both in chromaticity and in luminance, activated many associative areas along both the dorsal and the ventral pathways. We observed changes in activation along the dorsal stream, mainly in V3A, CIP, TOS and dLO. Along the ventral stream the activation changes were mainly located in V4. All these areas showed a stronger response for coherent stimuli than for random stimuli (Fig. 8A and C).

No change in response was observed in V1 when the stimuli were modulated in average phase, alternating between even- and odd-symmetry stimuli. A preference for the stimulus comprising odd-symmetry visual features (±90° average phase) was observed at higher cortical levels. Indeed, both the ventral and the dorsal pathways (V3A, CIP, TOS, dLO and V4) were more strongly activated by this stimulus both for equiluminant and luminance contrast modulation.

Table 1 shows the Talairach coordinates of the hot-spots of activation inside the RoI of these areas. The consistency between subjects is also strong for areas that could not be retinotopically segmented, such as TOS and CIP.

The average modulation of BOLD signals over the seven subjects, measured within the V1 RoI responding to the stimulus patch, is shown in Figure 9 (top row). The response to random- vs. congruent phase stimuli modulated in luminance showed a dominant first harmonic modulation of the BOLD signal, with a preference for congruent stimuli. The amplitude of modulation, projected on the direction of the presumed hemodynamic response of the boxcar model (Fig. 10, top left), was significantly ($P < 0.05$) different from zero.

		V3A			CIP			TOS			LO			V4		
		x	y	z	x	y	z	x	y	z	x	y	z	x	y	z
S1	left	-28	-89	12	-31	-68	24	-33	-76	15	-36	-79	-6	-25	-85	-11
	right	31	-86	12	30	-65	23	34	-81	11	31	-85	-4	27	-82	-8
S2	left	-22	-89	19	-21	-70	23	-30	-75	18	-33	-77	3	-34	-84	-4
	right	26	-75	17	22	-75	29	33	-80	9	40	-72	-3	36	-80	-2
S3	left	-33	-83	11	-29	-72	33	-42	-73	10	-41	-76	-3	-34	-83	-6
	right	30	-81	16	25	-64	29	37	-70	9	42	-73	-8	32	-80	-4
S4	left	-29	-79	15	-25	-71	36	-31	-66	15	-34	-78	-9	-32	-81	-6
	right	20	-74	13	25	-67	32	33	-69	19	39	-73	-6	33	-86	-9
S5	left	-23	-87	13	-24	-78	20	-31	-87	12	-26	-89	6	-32	-79	-6
	right	25	-88	16	22	-76	30	32	-76	15	35	-83	-4	30	-86	-3
S6	left	23	-81	17	-29	-68	24	-37	-82	19	-40	-80	-9	-35	-81	1
	right	28	-82	12	27	-67	29	39	-75	9	39	-72	1	30	-82	-4
S7	left	-26	-80	11	-27	-68	32	-40	-82	7	-39	-87	-3	-25	-87	-9
	right	22	-86	15	30	-65	22	37	-83	8	39	-87	-3	20	-80	-7

Table 1: Highest foci of activity (Tailarach coordinates)

The V1 first harmonic response of the same voxels group (V1F; representing 7° of eccentricity) was not evident for either colour condition, nor for even- vs. odd-symmetry stimuli modulated in luminance (Fig. 9, top row). Neither this central region of V1, nor the entire V1, showed a significant response modulated on the first harmonic. The time course showed a modulation, but it was in synchrony with the second harmonic, without a clear preference for congruent stimuli or for edges: the activity was stronger after each transition of stimuli, producing two peaks for each full period of alternation. The response on the second harmonic ($P < 0.05$ for average phase alternation) was statistically significant and present for both

the colour- and luminance-modulated stimuli (Fig. 10). This suggests that the repetition of the same stimulus in each block caused adaptation which recovered when the stimulus changed, indicating a selectivity for the phase of the stimuli in the central representation of the visual field. The pattern of the average results was different for V4 (Fig. 9, bottom row). The responses showed strong and significant BOLD modulations in all conditions (Fig. 10), both for luminance- and chromatic-modulated stimuli, with a preference for congruent-phase stimuli in the conditions over random-phase stimuli and also a preference for odd-symmetry stimulus compared with the even-symmetry stimulus.

Unexpectedly, V3A (Fig. 9, middle row; considered part of the dorsal pathway) also showed a stronger preference for congruent patterns and for odd symmetry for equiluminant stimuli. This pattern of results was not only characteristic of V3A but was also typical of the other dorsal visual areas, CIP, TOS and dLO, all showing a significant first-harmonic modulation of the BOLD signal for the chromatic as well as for the luminance stimuli (bootstrap sign-test: $P < 0.05$; Fig. 10).

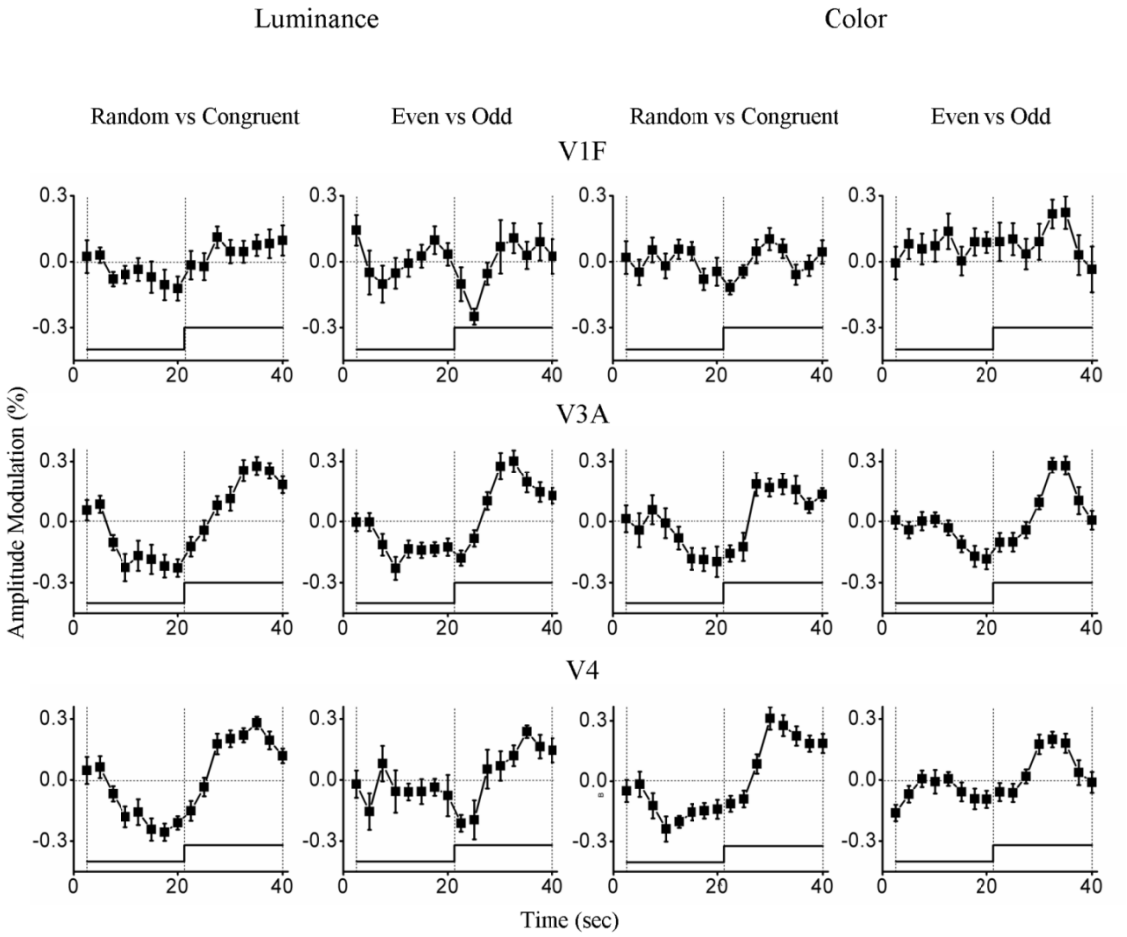

Figure 9. Averaged V1F, V3A and V4 BOLD time-courses. BOLD time-courses, averaged across all subjects' hemispheres and runs for the 7° central V1 (first row), V3A (second row) and V4 (third row) RoIs (defined by retinotopy), for the various conditions. For each subject the number of runs varied between two and four for each condition. The error bars represent 1 SEM. The transition of the stimulus is represented by the step waveform and marked by dashed lines.

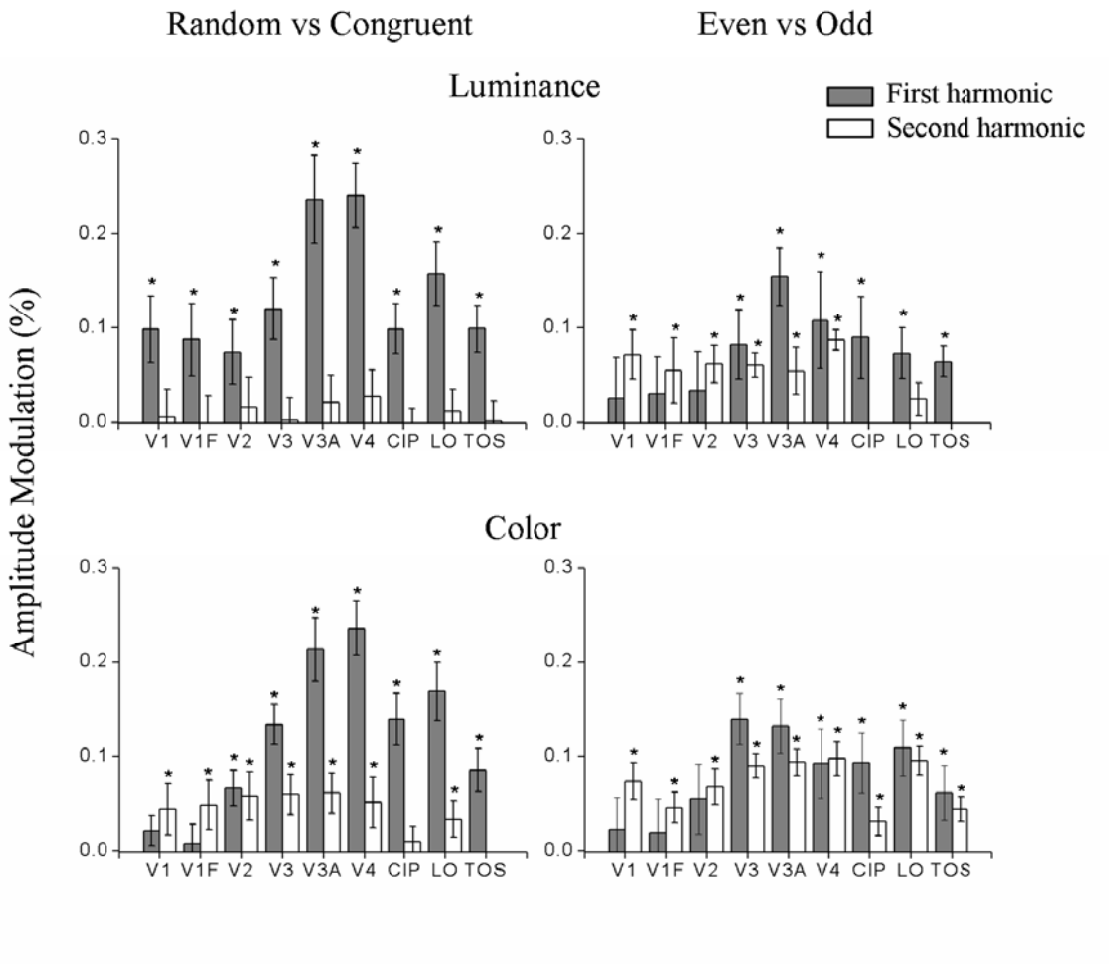

Figure 10. First and second harmonic amplitude. Bar plot of the first (gray) and second (white) harmonic amplitude projected along the phases of the corresponding hemodynamic components for all areas and conditions. The mean of the projected amplitudes was calculated across subjects' hemispheres and runs; the error bars report the error of the mean evaluated by bootstrap. A significant (*P < 0.05) first-harmonic amplitude modulation was observed in V3A, CIP, TOS, LO and V4 for all conditions. In the central 7° representation of V1 the first-harmonic modulation was significant (*P < 0.05) only for the random-phase vs. congruent-phase luminance condition, while the second-harmonic modulation was prevalent in all the other conditions.

The equiluminance point was established for each scan and subject by flicker-photometry inside the scanner, and the area of the stimuli was restricted to central vision to minimise contamination of nonequiluminant eccentric stimulation. Nevertheless, the modulation of activity to chromatic stimuli along the dorsal pathway could result from an erroneous equiluminance point or luminance contamination.

To control for this possible artefact, we measured the BOLD response change to random- against congruent-phase stimuli modulated in colour at different red : green colour ratios. Figure 11 shows how the amplitude and phase of the response modulation to random vs. congruent phase stimuli varied as a function of the red : green colour ratio. The response modulation was present and reliable in both V3A and V4 for all colour ratios.

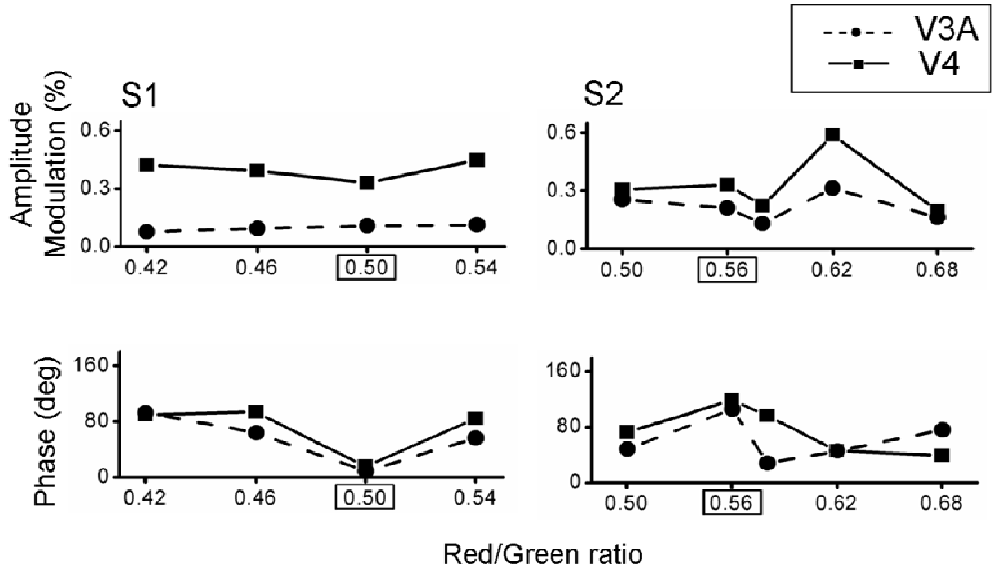

Figure 11. BOLD response amplitude as a function of the red : green colour ratio. V3A and V4 BOLD response amplitude and phase to random- vs. congruent phase stimuli modulated in colour stimuli as a function of the red : green colour ratio for two subjects.

Hence, an error in defining the subjective equiluminance point (corresponding to the outlined number on the abscissa) could not have affected the results. We conclude that the dorsal response modulation to coloured stimuli was not an artefact of an erroneous definition of the equiluminance point.

2.4 Discussion

In this experiment we measured BOLD selectivity to patterns modulated in luminance and chromatic contrast, with different phases but matched amplitude spectra. We observed strong activity modulation in associative areas, particularly in the dorsal pathways, and little or no change in activity in V1 in response to the colour-modulated stimuli.

2.4.1 Phase selectivity for V1

Confirming the previous study of Perna et al. (2008), our results show that alternation between coherent- and random-phase stimuli modulated in luminance activated several areas. The luminance BOLD response to stimuli with different phase congruencies was present in V1, even for the foveal representation. The V1 activity is strictly limited within the borders defining the central 7°, excluding the possibility that it originated from the luminance‑colour contrast variation along the external contour of the patch itself. This result reinforces the previous data by Perna et al. (2008), who used fullscreen stimuli and extend them to a lower range of spatial frequencies.

Perna et al. (2008) demonstrated that this activity was not due to a greater attentional allocation to the congruent-phase stimuli, given that performing a demanding foveal attentional task did not alter the BOLD modulation in V1, dLO or CIP. The presence of a reliable second-harmonic response in V1 also does not support the attentive explanation.

Perna et al. (2008) simulated the BOLD response to random- vs. congruent-phase alternation considering the responses of a battery of

neurons with even and odd RFs. In this simple model the pooling of all neuronal activity across space could vary from linear to highly nonlinear combinations. To simulate the stronger response to congruent-phase stimuli, the individual responses needed to be summed after a power amplification with an exponent of 3. A linear summation would produce a small response favouring the random, and an exponent of 2 a well-balanced response.

When the stimuli were modulated in colour, random- against congruent-phase stimuli did not produce any activity change in V1, suggesting that the primary visual area does not discriminate between colour-contrast stimuli of different phase congruencies. This suggests that both local and global RMS colour-contrast are good estimates of the V1 neuronal response: the response varies linearly with RMS contrast. Much evidence demonstrates that the response to equiluminant patterns is more linear (Kaplan and Shapley 1986), and that in general P-neurons have a linear response to equiluminant grating but not to luminance-modulated gratings. Similarly, VEPs to equiluminant stimuli do not show strong automatic contrast gain control mechanisms (Morrone, Burr et al. 1993), and the present data support this electrophysiological evidence. Importantly, the BOLD response to colour is nearly linear in V1 (Liu and Wandell 2005) but becomes more exponential, or saturated, in higher associative areas.

In agreement with the present results, this would indicate the presence of a response in higher associative area to congruent vs. random phase stimuli.

It is always difficult to infer a lack of response modulation from the failure to record reliable BOLD signals. However, the fact that from the same stimuli we were able to record a reliable response change at each transition (second-harmonic modulation) suggests that our technique was sensitive

enough to detect even small signals, and therefore supports the idea that V1 average activity is invariant to phase manipulation of chromatic stimuli.

We also observed a lack of a response modulation of V1 to different congruent-phase stimuli when modulated in colour. This also indicates that the critical parameter that modulates the response is not the Michelson contrast but the RMS contrast. Indeed, even- and odd-symmetry patterns differ by more than a factor five in Michelson contrast, with random-phase stimuli in between, but the V1 responses were unaffected by this variation.

Symmetry is a property encoded in the phase spectra, in particular in the relationship between the phases of the various harmonics.

Boundaries that demarcate an object are associated with changes in contrast, and have odd symmetry; these are points where harmonics of various frequencies have the same phase alignment of ± 90°.

Conversely, bars are locally even-symmetric functions and correspond to points where the phases of the harmonic components are 0° or 180° (for bright and dark polarity respectively). The responses of neurons with odd-symmetric receptive fields will be stronger to edges than to lines. Vice versa, the response of neurons with even-symmetric receptive fields will be stronger to lines than to edges.

The second-harmonic response to the alternation of even- against odd-symmetry stimuli in V1, even for the near-foveal representation, suggests the existence of different neuronal mechanisms with different RF symmetries, some that respond preferentially to lines and the others to edges. The second-harmonic response could result from the adaptation of even- or odd-symmetric RFs neurons. The two different subpopulations of neurons may be present in equal numbers in the same voxel, and the repeated presentation of the neurons' preferred stimulus (which led to

response suppression) would elicit a transient response by the unadapted neurons to the complementary stimulus.

In our experiment, the repetition within each block of the odd-symmetry stimulus could adapt the subpopulation of neurons with odd-symmetry RFs, leading to a response, in the following block, to the presentation of the even-symmetry stimulus mediated by neurons with even-symmetry RFs. The V1 BOLD second-harmonic modulation strongly suggests the existence of two subpopulations of chromatic neurons with even- and odd-symmetry RFs, with an overall balanced response.

This is in agreement with a recent electrophysiological study on monkeys (Johnson, Hawken et al. 2008) that demonstrated the existence of double-opponent cells with oriented receptive fields of various symmetries, equally selective to colour- and luminance-oriented gratings. These types of detectors could perform a 'conjunction analysis' of the chromatic and shape information.

The existence of detectors with different RF symmetries can be inferred from several previous psychophysical studies: colour and luminance phase discrimination is similar once the stimuli are equated for cone contrast (Burr, Morrone et al. 1989; Burr, Morrone et al. 1992; Girard and Morrone 1995; Martini, Girard et al. 1996). Taken together, these provide evidence to suggest that for colour, as for luminance, the brain may implement similar algorithms to segment and locate salient features.

2.4.2 Phase selectivity for dorsal and ventral associative areas

All our stimuli, modulated either in luminance or in colour, elicited widespread modulation of activity in V3A, in dorsal LO, in the more caudal part of the occipital branch of the intraparietal sulcus, in TOS and in V4. The responses in these areas showed a preference for congruent-phase stimuli and for odd-symmetry stimuli, both for luminance and for colour. These higher cortical areas are involved both in the phase-congruency computation (given the small response to noise) and also in the computation of absolute phases (given the preference for odd-symmetry), suggesting a role in contour detection. Contours are associated with the phase congruency of ±90°, and are more frequent in natural images than other features. It is possible that at higher levels of analysis the neurons that locally process contours are more numerous, eliciting a stronger BOLD response. This conclusion supports previous fMRI studies that found V3A, V4 and LO to be active during contour processing (Schira, Fahle et al. 2004) and dLO, CIP and TOS to be involved in brightness perception of surfaces (Perna, Tosetti et al. 2005; Cornelissen, Wade et al. 2006). However, we cannot dismiss the possibility that the stronger BOLD response to the checkerboard pattern may also be generated by the presence of the brightness illusion that increases the saliency of the pattern.

Unexpectedly, our results demonstrate that the dorsal pathway is involved as much as the ventral pathway in detecting salient features modulated in colour. The dorsal involvement in the analysis of colour is at odds with the traditional view that assigns colour analysis to only the ventral pathway

while shape analysis may also involve the dorsal pathway (for a review see (Grill-Spector and Malach 2004).

However, in many natural conditions, contours are associated with changes both in chromatic and in luminance contrast (Hansen and Gegenfurtner 2009). To detect the surface contour it would be more efficient to analyse the luminance and the chromatic information together. It is possible that the analysis performed by the dorsal pathway conveys the contour signals that mediate action towards equiluminant shapes, which has been demonstrated to be as reliable as for luminance-contrast stimuli.

Responses in the dorsal pathway to colour stimuli have been observed by many authors (Liu and Wandell 2005; Mullen, Thompson et al. 2010; D'Souza, Auer et al. 2011) and the responses in these areas are similar for luminance and colour stimuli, as we observed here. At present it is still unclear whether these areas form part of the cerebral network specialised for colour perception, and the strong V3A activation to chromatic stimuli observed in our study does not imply that this and other dorsal areas mediate colour perception, but that they are involved in the analysis of colour form.

As for luminance, it seems that chromatic-contrast detectors operating as early as V1 may be selective for spatial phase. However, preference for edge-like phase takes place only in areas downstream of V1, in the higher-order visual areas. The role of V4 seems not to be crucial, but only part of a network for the analysis of colour form: a stronger response to edges indicates that odd-symmetric RFs may be more common or more sensitive in V4.

In conclusion, the human visual system seems to possess the mechanisms needed to analyse colour and form together, as a single attribute of an

image, in order to structure the representation of the visual scene. This fMRI study demonstrates that both the dorsal and ventral pathways are involved in colour-form analysis, which could be based on selectivity to spatial phase.

3. Part II: Object enumeration

3.1 Introduction

3.1.1 Psychophysical studies

As mentioned in the general introduction humans share with other animals an "approximate non-verbal numerosity system" which gives them the possibility to perform an approximate estimation of the number of elements contained in a visual scene in a single glance.

By showing that the perception of numerosity is susceptible to an adaptation after-effect, psychophysical experiments have suggested that numerosity can be considered to be a primary sensory visual attribute.

Burr et al. (2008) demonstrated that after 30 seconds of prolonged exposure to two patches differing in numerosity, two subsequent patches containing the same number of dots appear to be different in numerosity. They asked subjects to judge if a test stimulus (which they varied in numerosity) was more or less numerous than a probe stimulus (of a constant numerosity). The results showed that the apparent numerosity of the probe stimulus was decreased by adaptation to high numerosities and increased by the adaptation to low numerosities. The described number adaptation effect was found to be not dependent about variations in pixel density, orientation, shape or element size but instead coupled only with the number of elements. It was found that even changing contrast of the adapting stimulus had little effect on the magnitude of the adaptation effect.

Since these initial findings it has been argued that the concept of numerosity could acutally still be explained by a simple combination of the

density within a given space and the total filled area (Durgin 1995; Durgin and Huk 1997; Durgin 2008; Dakin, Tibber et al. 2011; Tibber, Greenwood et al. 2012). Instead of being related to the number of elements, it has been suggested that the effect of numerosity adaptation was due to adaptation to the texture density of the stimulus and that the adaptation effects were happening at a very initial stage of the visual process in the primary visual area (Durgin and Huk 1997; Durgin 2008).

These authors reached their conclusions through the creation of texture stimuli in which density and spatial frequency did not co-vary and then subsequently adapting these stimuli to differentially dense textures composed of elements with different Fourier spectra. The authors then proposed that the observed density aftereffect occurs in the space domain and not in the frequency domain, and concluded that the effect therefore cannot be explained in terms of a spatial frequency shift aftereffect (Durgin and Huk 1997).

According to their view the Fourier specificity which they found instead indicates that the adaptation effects were likely due to early cortical texture analysers (presumably simple and complex cells in primary visual cortex) rather than from an abstract numeric representation.

In order to disentangle texture from number Ross and Burr (2010) conducted another experiment in which they measured numerosity judgments in response to sparse dot patterns under conditions where density and area were not informative cues to numerosity. They asked the subjects to give judgments about the apparent numerosity or density under three conditions: a constant-area condition (where area was kept constant while number and density were changing), a constant-density condition (where density was kept constant while number and area were changing)

and constant-numerosity condition (where number was kept constant while area and density were changing). Even when the trials were chosen at random from these three conditions, numerosity judgments where as precise as density judgments. Similarly, numerosity judgments under the constant area and constant density conditions were equally good, implying that there is no additional cost in their calculation (i.e. there is not a higher JND as one would have expected if density were a proxy for numerosity). Moreover numerosity judgments in the constant numerosity condition were not biased by variation in either density or area. Finally, in order to directly address the issue of texture, they replicated their experiment but this time varying the luminance both on the typical sparse dot stimulus as well as on a dense visual texture stimulus comprising black or white rectangles. As it happens for many visual properties, like apparent colour and speed, manipulating luminance affected the perceived numerosity of the dot pattern (such that lowering luminance increased perceived numeoristy), but did not affect the perceived texture density, which remained constant across all tested luminance levels.

Hence in this experiment the authors were able to demonstrate that numerosity is able to be estimated without passing through an intermediate stage of evaluating of the texture density and thus the only factor that affects the adaptation effect described is numerosity.

Support in favor of the existence of a specific number sense is also provided by a study from Stoianov and Zorzi (2012) which demonstrated that the selectivity to visual numerosity emerges naturally as a statistical property during unsupervised learning of a hierarchical generative model of perception. This model, mimicking known neural architecture, can discriminate numbers in a way similar to humans such that it both abides, by

following the Weber's law (with a Weber fraction of 0.15, similar to humans) and shows invariance to area, density and object features when judging the numerosity of visual stimuli.

Overall, although a contentious issue in the literature, the majority of recent research supports the notion that number is a primary visual attribute or qualia.

In effort to gain further experimental evidence that adaptation to number is not mediated by the adaptation to density of visual features at early visual stage of processing, the current experiment examined numerosity perception with the overall contrast energy across numbers balanced. This should help to further investigate the role of the primary visual area, since according to the local energy model of feature detection (Morrone and Burr 1988; Perna, Tosetti et al. 2008; Castaldi, Frijia et al. 2013), the energy is the fundamental property modulating V1 activity.

3.1.2 Neuroimaging studies

Many neuroimaging studies on humans have also contributed to the number sense hypothesis and demonstrated parallels between the number system as observed in monkeys and in humans.

Indeed number-specific activity in the interparietal sulcus (IPS) and in the prefrontal cortex (PFC) have been shown in humans performing number comparison and non-verbal magnitude comparison tasks (Pinel, Dehaene et al. 2001; Fias, Lammertyn et al. 2003; Piazza, Izard et al. 2004; Pinel, Piazza et al. 2004; Piazza, Pinel et al. 2007). These number-related parietal activations have been demonstrated to be already present in four-year-old children

(Temple and Posner 1998; Cantlon, Brannon et al. 2006), reinforcing the idea that the IPS is the site of the innate non-symbolic number processing system.

A particularly powerful method of searching for number-selective neurons in the human cortex is through the use of fMRI adaptation and habituation paradigms.

The basic principle of adaptation in functional imaging is that if any region of the brain contains a population of feature-selective neurons (i.e. tuned to a specific number of dots), then repeated presentations of the same feature (i.e. a certain numerosity) should reduce the activation of that region to it, while leaving its response to the other non-repeated features (i.e. to the other numerosities) unchanged (Krekelberg, Boynton et al. 2006).

Following this logic, Piazza and colleagues (2004) passively exposed subject to visual arrays of dots. Subjects habituated through repeated presentation of a standard number (either 16 or 32) and throughout this habituation occasional deviant stimuli were interleaved. The numerosity of these deviant stimuli ranged from half to double the value of the standard. Although participants were not explicitly told to discriminate the visual stimuli in any way, and instead passively viewed the displays, it was found that the recovery of the BOLD signal along the right and left IPS was proportional to the ratio of the standard and deviant stimuli. As such, the changes in brain activation observed following the presentation of a deviant numerical stimulus was found to change in exactly the way predicted by Weber's law. Changes in the shape of the elements did not alter the results. More generally, in another experiment using the same method, Piazza and colleagues were able to demonstrate that there is an abstract coding of numerical magnitude which goes beyond the notation used to present numerosity - the same pattern of results in IPS was observed with both non-

symbolic and symbolic stimuli (Piazza, Pinel et al. 2007). A further study confirmed the adaptation effect shown in the IPS and extended these finding to proportions (Jacob and Nieder 2009).

Another very powerful method which extends the capability of using functional imaging in humans to identify spatial patterns is the multivariate decoding or multi voxel pattern analysis method (Haynes and Rees 2006; Norman, Polyn et al. 2006; Mahmoudi, Takerkart et al. 2012). Multi-voxel pattern analysis (MVPA) is becoming increasingly popular in the neuroimaging community because it allows researchers to detect differences between conditions with higher sensitivity than conventional univariate analysis allows. MVPA is considered a supervised classification problem where a classifier attempts to capture the relationships between spatial patterns of fMRI activity and experimental conditions. Classification consists in defining a decision function that takes the values of various “features” (each representing a response measure of a specific voxel) to "example", i.e. exemplar trials of a certain condition, and predicts the class for that "example". To do this, the dataset (i.e., trials and the corresponding class labels) must be split into two sets: “training set” and “testing set”. The training set is used to train the classifier, a process consisting of modelling the relationship between the features and the class label by assigning a weight to each feature for each class. This weight corresponds to the relative contribution of the feature to successfully classify the classes. Then the testing set is used to evaluate the classifier performance in capturing the relationship between features and classes. This is done by utilising the calculated weights and the value of the test stimulus for each feature in order to determine the most probable class for the pattern.

This method has been used to study the organization of basic visual features, like orientation (Haynes and Rees 2005; Kamitani and Tong 2005) and motion direction (Kamitani and Tong 2006) as well as more complex stimuli such as objects (Eger, Kell et al. 2008) and faces (Kriegeskorte, Formisano et al. 2007).

Recently this method has been used also to study the perception of numerosity. In a delayed numerosity comparison paradigm, similar to the one employed previously in the study of numerosity in the monkey, Eger and colleagues (2009) asked subjects to judge if the second stimulus was smaller or larger than the previous one. In the intra-parietal sulcus it was possible to predict the numerosity of non-symbolic sets of dots in spite of changes in low-level parameters of the stimuli. Also digits were predictable and it was possible to generalise across format (digit to dots). The authors concluded that, as they were able to decode number from multivoxel pattern classifications, the layout of individual number codes must be at least sufficiently structurally organized such that individual voxels are more associated with one magnitude than another, even if there is not sufficient evidence for any columnar, numerosity-based map structure.

Another study demonstrated that in the parietal region the brain activation patterns evoked by quantities in the non symbolical (pictorial) mode were largely common across different objects (indicating the existence of some object-independent quantity representation) and that the neural representation of quantities were common across participants of, indicating some universality of the neural coding of small numbers (Damarla and Just 2013) .

The intuition that numerosity could be organized in a "numerotopic" map structure, comparable to those described in the primary sensory and motor

cortex for other primary sensory attributes, has been finally supported by Harvey and colleagues (2013) by using a method analogous to conventional population receptive field analysis in visual cortex. The results showed the existence of neural populations tuned to numerosities (from one to seven dots) in human parietal cortex. In this map the lowest numerosities were represented more medially, while the highest numerostities were represented more laterally. Cortical magnification decreased at higher numerosities, meaning that more cortical surface area was utilized to represent lower numbers than higher numbers. The over-represented parts of the topographic maps for low numbers also showed a more precise response selectivity than elsewhere in the map. In other words tuning width was smaller suggesting a better ability to discriminate low numbers from each other than high numbers from each other. Importantly this topographically organized map for numbers was robust to variations in low-level stimulus features which were carefully controlled individually in repeated experiments. This included investigations of numerosity representation in the presences of changes in area, dot size, circumference, density and in the type of features present with all manipulations found to not affect the organized representation of number. Additionally, Harvey and colleagues compared the activation of IPS and V1 with respect to the energy of the stimulus and showed that while the primary visual cortex response was highest for the stimuli with higher energy, this parameter was not determining the pattern of activity in IPS. They concluded that organizational properties described for numerosity extend topographic principles to representation of higher-order abstract features in association cortex.

3.1.3 Present Study

In the present experiment we are testing the effect of adaptation on numerosity decoding in human brain.

In order to test if classifiers could decode the different numerosities, both in V1 and in IPS, we used multivoxel pattern analysis. In case both these regions can classify numbers we would like to test if the effect of numerosity adaptation is specific to higher-order number-selective neurons of IPS or if it could be traced back to low level property detection in V1.

Importantly we manipulated the non-symbolic stimuli presented by equating the energy across the different numerosities.

If numerosity is derived only by density, the ability to classify the different numerosities (or the different densities) should be evident already in V1.

Because adaptation is expected to alter the response to numerosity of a population of number-coding neurons, we would like to characterize the impact that this change in response pattern has on the classifier's ability to discriminate numbers between before and after adaptation for the same classes.

Hence if the effect of "number adaptation" is instead just "texture adaptation" we would expect to see some changes in the classification in V1 and not, or not only, in IPS.

Moreover we would like to measure the effects of adaptation, without the response to the adaptor. Indeed in previously used habituation paradigms the activity associated with the deviant stimuli and the habituating stimuli were convoluted and only relative changes after the appearance of the deviant stimuli could be measured.

We developed a novel adaptation paradigm, wherein rapid adapting and testing stimuli were separated by more than 20 seconds and still producing psychophysical adaptation to number. This procedure is ideally suited for measuring BOLD adaptation within a rapid event related design and allows for the temporal dissociation between activity evoked by the adaptor and the test stimuli, as opposed to habituation paradigms.

3.2 Materials and Methods

3.2.1 Stimuli

The stimuli were 10 degrees diameter dot patches varying in numerosity with no possibility for the single dots to overlap. The numerosities used were 20, 30, 40, 60, 80 dots and we adapted to the 80 dots.

Half of the dots were black and half white so that their overall mean luminance matched with the one of the gray background.

Most importantly the overall energy of the stimuli was balanced across numerosities by rescaling the contrast of each numerosity pattern with the square root of the lowest numerosity divided by the numerosity considered.

In this way as the number of dots increases, the contrast decreases ensuring that the energy level across all Fourier components is kept constant.

As shown in Figure 12, the contrast of the stimuli before this correction was 90% for each numerosity pattern, hence the patterns with the highest number of dots had the highest energy. After the rescaling, the contrast varied from a maximum of 90% in the case of 20 dots to 45% in the case of 80 dots, ensuring equal energy level across the different numerosities. The

contrast of the adapting stimulus (80 dots) was 25%, since adapting to low contrast stimuli is known to produce stronger adaptation effects.

The stimuli were generated and presented under Matlab 7.10 using PsychToolbox routines (Brainard 1997). They were displayed in a dimly lit room. DELL monitor with 1920 x 1080 resolution at 60 Hz refresh rate, mean luminance 60 cd/m2, viewed binocularly from 57 cm.

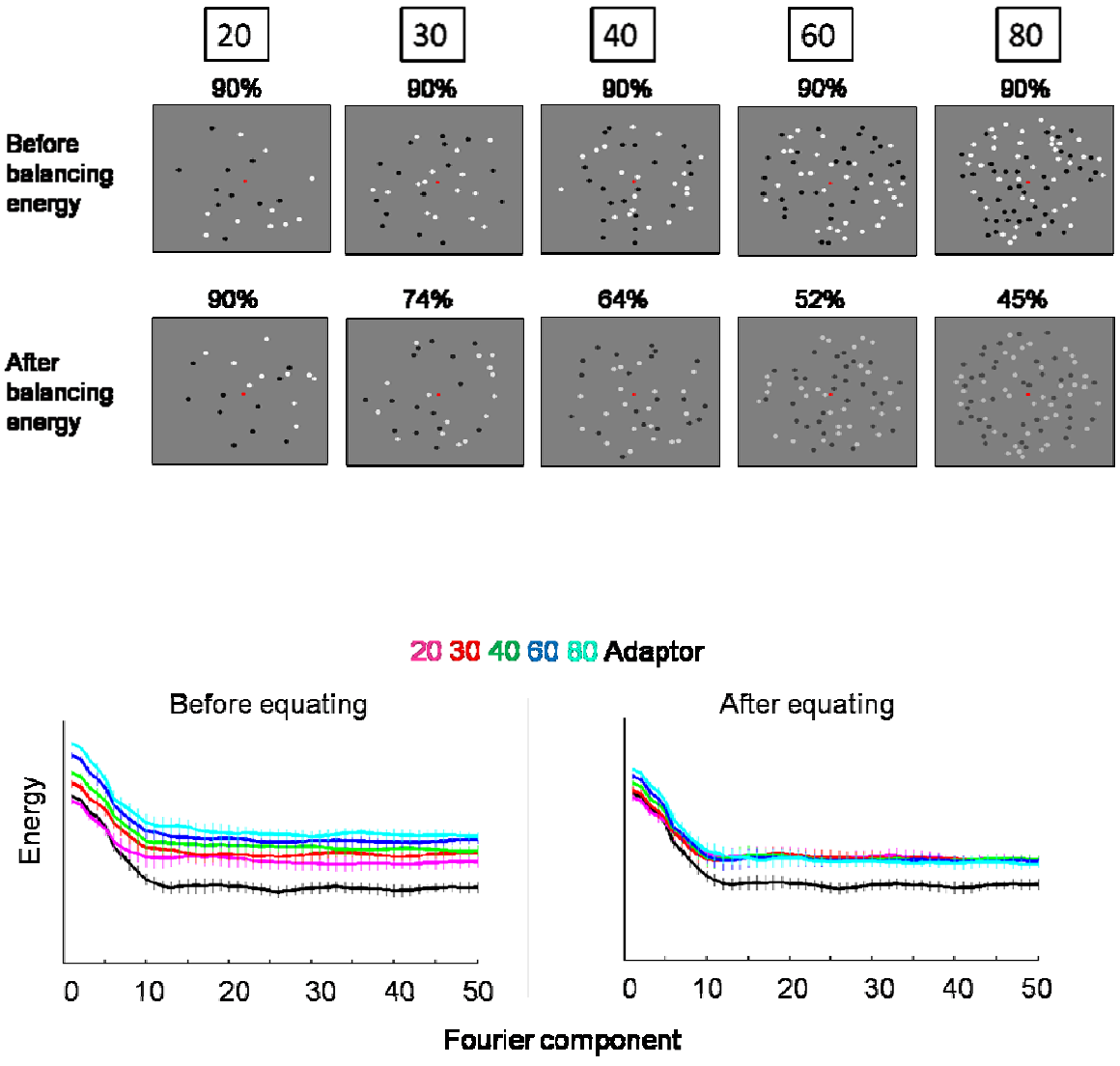

Figure 12 The top panels show examples of the stimuli used before and after the energy balancing. The graphs show the change in the energy level across the Fourier component for all the stimuli before (left) and after (right) the energy balancing. The black line (the adapter) was kept at a constant contrast of 25%.

3.2.2 Adaptation Paradigm

The experimental protocol was separated in three main parts summarised in Figure 13.

In the Pre-Adaptation part (Fig.13 panel A) subject were presented to six trials for each of the five numerosities tested. The stimuli were shown for 1 second, then after a variable delay of 3-6 seconds the subsequent one was displayed.

After the Pre-Adaptation phase, a period of rapid adaptation begun for 2 minutes and the subjects were exposed to repetitive presentations of the adapting stimuli, (i.e. the patch of 80 dots) which was displayed every 0.75 seconds for 0.75 seconds (Fig.13. panel B).

Finally in the subsequent Post-Adaptation phase (Fig.13 panel C) the five different numerosities were tested again, with the same timing used in the Pre-Adaptation phase, however now each testing period was separated by Top-Up Adaptation periods by a delay period of 24 seconds. In the Top-Up Adaptation periods the adaptation stimulus was repeated in order to boost the adaptation effect. The sequence started with one Top-up period, where the adaptor was repeated with the same timing as in Adaptation phase, but lasting only 15 seconds. After this first Top-Up period finished, a delay of 24 seconds was introduced before the testing period started.

Top-Up and testing periods were alternated for three times, so that the total number of trials per numerosities tested in the Post- Adaptation phase was the same as the Pre-Adaptation phase. We used a pseudo random order

to assign the numerosity to be tested in the three testing period of the Post-Adaptation sequence.

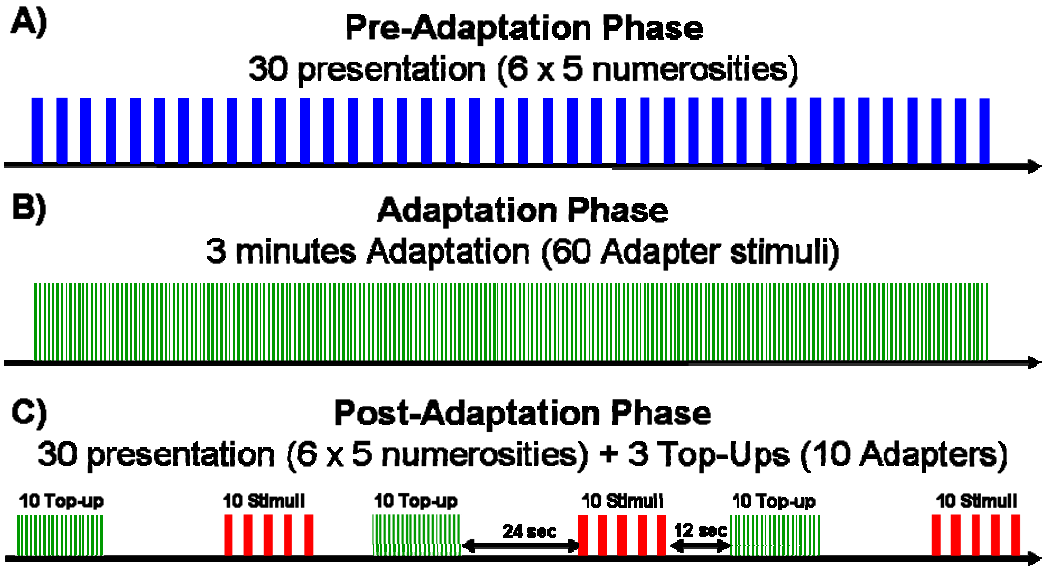

Figure 13 Schematic representation of the Adaptation paradigm divided in three phases.

A red fixation dot was always present on the centre of the screen and the subjects were required to maintain steady gaze for the entire length of the experiment.

We used this paradigm both in a psychophysical experiment as well as in an fMRI rapid event related experiment.

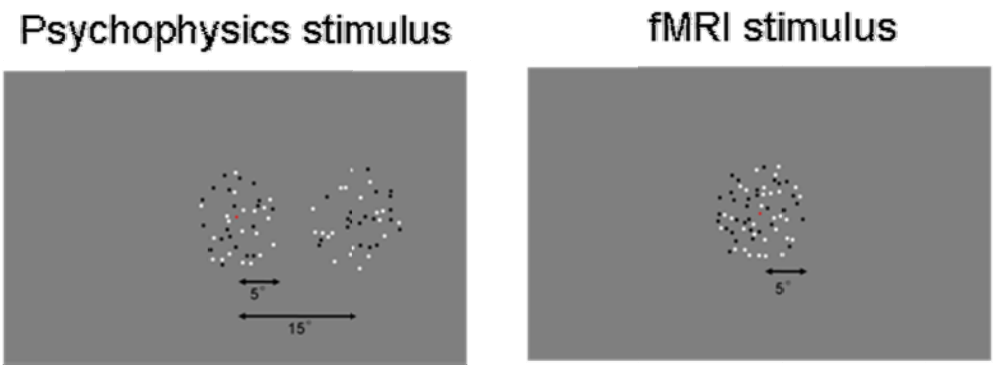

Figure 14 Left: Stimuli used in the psychophysical experiments. Right: Stimuli used in the fMRI experiments. Each patch was 10° diameter.

As shown in Fiugre 14 in the psychophysical experiment the subjects were exposed to two patches, one 15° apart from the other, one in the fovea (testing patch), at fixation, while the other in the periphery (probe patch), both during the Pre-Adaptation as well as during the Post-Adaptation phase. When the stimuli disappeared the subjects were required to judge which of the patches was the more numerous with a keyboard response. Subjects were randomly presented with either 20, 30, 40, 60 or 80 dots in the test (foveal) patch. The number of dots in the test was initially equal to the probe, then varied from trial to trial depending on subject response, with numerosity determined by the QUEST algorithm (Watson and Pelli 1983), and with parameters initial numerosity = probe numerosity, standard deviation = 0.5 log-units; beta = 3.5; epsilon = 0.01; gamma = 0. To determine the numerosity of the next trial, the algorithm estimated the point of subjective equality (PSE) after each trial, then perturbed that with a random number drawn from a Gaussian distribution of standard deviation 0.15 log-units.

Each subject performed two Pre- and two Post- Adaptation sequences. During the Adaptation period only one patch was presented in the fovea and no answer was required. At the end of each session, data were analysed separately for each subject. The proportions of trials where the test appeared more numerous than the probe was plotted against tested numerosity and fitted with cumulative Gaussian functions, yielding estimates of PSEs (median of psychometric function) and of JND (just noticeable difference, standard deviation of psychometric function, reflecting precision). The difference in PSE between Pre and Post conditions was then examined to identify the % change of PSE after adaptation.

The same timing and numerosities have been used in a rapid event related experiment where nevertheless only a single central dot patch has been used for the three different phase of the experiment (Fig.14). The choice of presenting a single central patch gives the possibility to better evaluate the contribution of the primary visual area, since the central vision is mostly sensitive to density cues.

Stimulus presentation was synchronized with the fMRI sequence at the beginning of each run.

Each subject performed four Pre- Adaptation runs and two Post-Adaptation runs. Hence, since each of the five numerosities were presented six times within each run, we recorded for each subject a total of 120 trials in the Pre-adaptation runs and 60 trials in the Post-adaptation runs.

Subjects were asked to fixate a stable target that remained in the centre of the visible screen for the entire experiment duration.

3.2.3 Subjects and Procedures

Ten healthy adults (seven females and three males between 25 and 27 years old) with normal or corrected-to-normal acuity underwent to both the psychophysical and the fMRI experiments.

This study was conducted under ethical approval from the Stella Maris Scientific Institute Ethics Committee. Subjects gave informed consent in accordance with the Declaration of Helsinki.

MRI data were acquired using a GE 1.5 THD Neuro-optimized System (General Electric Medical Systems) fitted with 40 mT/m high-speed gradients.

Each session included a whole brain set of anatomical images with T1-weighted contrast. T1-weighted scans were acquired with TR =8.4 ms, TE =3.9 ms, flip angle = 8°, FOV = 256x256 mm^2, slice thickness = 1 mm.

Echo Planar Imaging (EPI) sequences were used for the fMRI data acquisition (TR= 3000 ms, TE = 35 ms FOV= 192x192 mm, flip angle =90°, matrix size of 64x64 and slice thickness = 3 mm, 95 volumes for the Pre Adaptation scans and 125 volumes for the Post Adaptation scans).

Head movement was minimized by padding and tape.

Brain Voyager Qx (version 2.6 Copyright © 2001-2011 Rainer Goebel) was used to analyse imaging data.

Anatomical images were spatially normalized according to the Talairach and Tournox atlas (1988) to obtain standardized coordinates for the region of interest. For displaying the functional data, an averaged anatomical image has been created by combining the normalized anatomical scans of all the subjects.

Functional data were preprocessed to compensate for systematic slice-dependent differences in acquisition time (using cubic spline), three-dimensional motion correction (using Trilinear / Sync interpolation realigning data to the first volume of the first scan) and temporal filtered (High-pass filter GLM with Fourier basis set, including linear trend, with 2 cycle). No spatial smoothing has been done.

Data from the Pre- and Post- Adaptation runs have been included together in a Multi study - Multi subject RFX dummy coded GLM , in order to control for the time differences between the Pre-Adaptation runs (acquired at the beginning of the session) and the Post-Adaptation runs (acquired at the end of the session). To create the design matrix the BVA---Predictor Tool (1.5.2, J.M. Born, Maastricht, The Netherlands) has been used. The GLM was

constructed with five regressors (the five numerosities tested) , while the "Top-Up" regressor on the Post-Adaptation runs has been considered as "predictor of no interest" in order to keep the same number of regressors between Pre- and Post- Adaptation, a necessary condition to build up the design matrix. Moreover data have been z-transformed and convolved with the canonical hemodynamic response (HRF). A contrast "All five numebers>blank" was performed at the group level using a Random-effect group analysis (RFX). Moreover specific contrasts for each numerosities between Pre- and Post- Adaptation scans have been tested, balancing the contrast to take into account the different number of runs acquired before and after adaptation . The significant threshold was set to $p<0.005$, and cluster size to 10 voxels.

Additionally, a Multi study dummy coded GLM on each subject has been calculated, with a regressor per each trial presented, specifying to separate study predictors. Data were z-transformed and convolved with the canonical HRF. The beta values for each trial presented has been extracted from two anatomically defined region of interest defined on each subject's anatomy, the first along the primary visual cortex and the second including around 1000 voxels along intraparietal sulcus.

The values extracted have been used to train and test classifiers using LibSVM.

A searchlight analysis was performed by training and testing the ability of the classifier to discriminate numerosity in a sphere of three functional voxels of radius which sequentially moved across all voxels contained in the ROIs. The 50 spheres showing the highest generalization accuracy when tested were then selected.

Different models were trained and tested using leave-one-trial-out cross validation. The five left out trials (one per each numerosity) have been selected randomly and kept separated from the training dataset. This procedure was then bootstrapped 500 times, to ensure that all the trials would have been left out to be part of the test. Then the average accuracy of the classifier across bootstraps was calculated. A linear classifier was utilised and C was fixed to 1.

We trained and tested three different models. In the first model we trained a classifier using the trials from Pre- Adaptation and then tested it using first the left out trials from Pre- Adaptation and then the same number of trials from Post- Adaptation. In the second model we trained a classifier using the trials from Post-Adaptation and then tested it using first the left out trials from Post- Adaptation and then the same number of trials from Pre- Adaptation. Hence in these two models we evaluated the classifier's ability in discriminating 5 classes, i.e. the 5 different numerosities, as well as its ability to generalize to the same 5 classes when the trials from the other adaptation condition were presented as a test. However a possible inability to discriminate classes in the other adaptation condition, could be derived simply by a "sequence" effect. To exclude this potential explanation of our results, in a third model we trained and tested the classifier on 10 classes, joining the Pre- and the Post- Adaptation trials together. The statistical significance of the findings has been evaluated with t-tests comparing these results with the classifiers accuracy when trained and tested with the same data but shuffling the labels. Testing against the permutations allows us to evaluate not only whether the classification is above chance but also to control for any sporadic correlations that can occur with the relatively few number of test trials utilized in MVPA methodology.

3.3 Results

The results from the psychophysical experiment are shown in Figure 15, where the percentage change in PSE after adaptation relative to Pre-Adaptation baseline is plotted for each numerosity tested.

The shift in PSE from baseline is modest, approximately 10%, but significant for all numerosities ($p<0.05$). This result is in line with the previously described results in literature showing that adapting to high numbers (i.e. 80 dots in our case) causes an underestimation of the perceived numerosity.

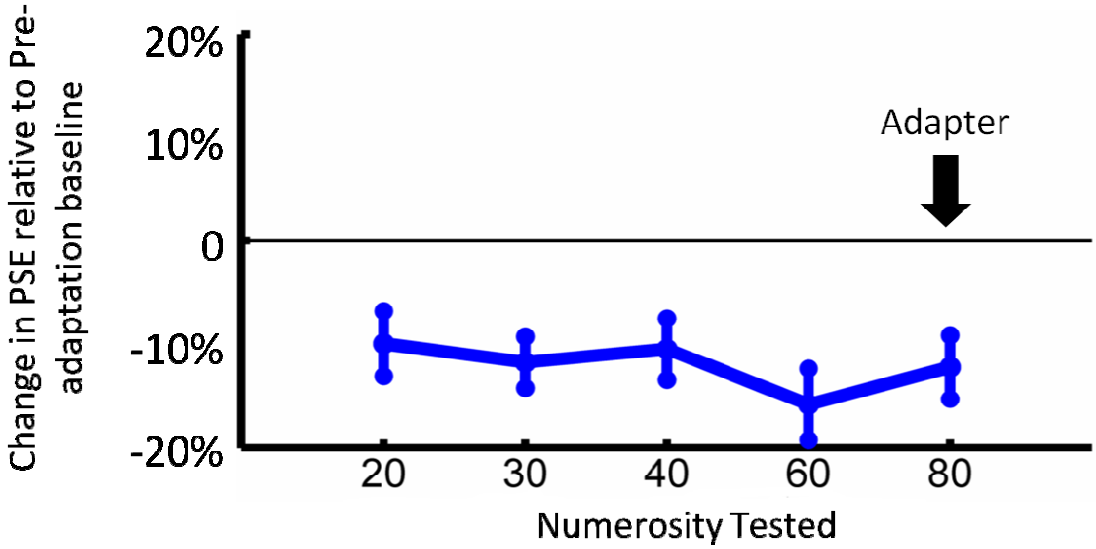

Figure 15 . Behavioural results from the psychophysical experiment consisting in foveal number adaptation to 80 dots and matching to a peripheral target. The average percentage change in PSE after adaptation across 10 subjects is shown for the five numerosities tested.

Once confirmed that the adaptation paradigm worked, the experiment has been replicated in the scanner, as described in the method section.

The results from the Multi study - Multi subject RFX dummy coded GLM analysis are showed in Figure 16.

The activity for the contrast "all the numbers higher than baseline" identifies, other than an extended activation of all the occipital areas, two bilateral regions along the intra parietal sulcus, typically descried in number experiments. In line with previous reports, the coordinates are x=27, y=-58, z=45 in the right hemisphere and x=-29, y=-57, z=48 in the left hemisphere.

A Volume of Interest (VOI) analysis has been performed for the combined left and right primary visual area and intra parietal regions.

In the primary visual area the contrasts testing the differences between Pre- and Post- Adaptation for each numerosity are not significant (20 Pre-Adaptation VS 20 Post-Adaptation: $t(9)=-1.4$, $p=0.16$; 30 Pre-Adaptation VS 30 Post-Adaptation: $t(9)=-1.9$, $p=0.054$; 40 Pre-Adaptation VS 40 Post-Adaptation: $t(9)=0.8$, $p=0.39$; 60 Pre-Adaptation VS 60 Post-Adaptation: $t(9)=0.8$, $p=0.41$; 80 Pre-Adaptation VS 80 Post-Adaptation: $t(9)=-1.1$, $p=0.25$).

The activity in the intra parietal sulcus bilaterally was instead affected from adaptation. At this level the biggest change between Pre- and Post-Adaptation was shown for 60 dots ($t(9)=2.1$, $p<0.05$) and for the numerosity 20 and 30 the results are close to significance ($t(9)=1.9$, $p=0.056$). At odd with what we would have expected the contrast was not significant for 40 dots and for 80 dots ($t(9)=1.4$, $p=0.14$; $t(9)=1.2$, $p=0.22$).

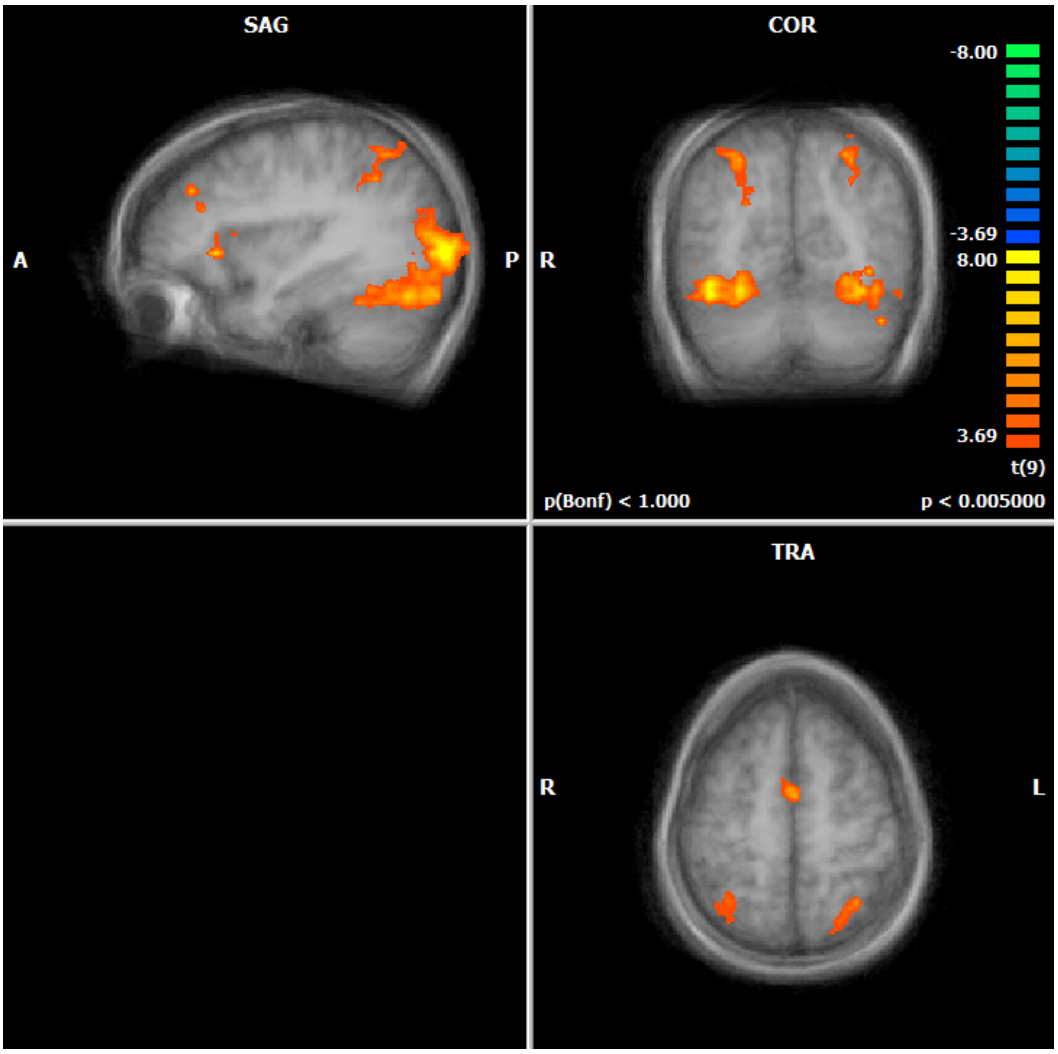

Figure 16. Brain regions activated for the contrast "All numbers>baseline" resulting from the RFX GLM analysis on 10 subjects superimposed on an average anatomical image. z=46 ; y=-57

Afterward classifiers have been trained and tested using beta values estimated for each single trial presented and extracted from two anatomically defined region of interest: the primary visual cortex and the intraparietal sulcus. These two ROIs have been defined on each subject's anatomy, by choosing a very foveal selection of V1 and around 1000 voxels for the intraparietal ROI, in line with Eger et al. (2009).

The results for the three classification models for V1 and for IPS are shown in Figure 17 and Figure 18 left column respectively. The models have been also trained and tested using the same data but with shuffled labels (Figure 17 and Figure 18 right column). As expected, when the models are trained with shuffled data the accuracy in classifying tested data is always at chance level.

In the first two rows the 5 classes models are trained and tested for V1 (Fig. 17) and IPS (Fig. 18), combining both the left and the right hemispheres together. Specifically the first rows show the averaged results from the 50 spheres with the highest generalization accuracy when tested to discriminate the 5 different numerosities with Pre-Adaptation trials (red curve) and then tested either with the five left out (one per each numerosity) Pre-Adaptation trials (green curve) and with the same number of Post-Adaptation trials (blue curve). Only the classifier in IPS, and not the one in V1, is able to generalize to trials that it has never seen before and to correctly identify the five numerosities with an average accuracy of around 30% ($p<0.05$) either when tested with Pre- Adaptation trials as well as when tested with Post-Adapatation trials. There is no significant difference in the classifier's accuracy when tested with Pre- or with Post- Adaptation trials in IPS ($p>0.05$), suggesting that the classifier is able to discriminate well between numerosities regardless the adaptation condition the tested trials are coming from. The classification accuracy in V1 is never significantly different from the accuracy of the model trained with shuffled data (which is at chance level: 20%) either when tested with Pre- Adaptation trials as well as when tested with Post-Adapatation trials.

The second rows of Figure 17 and Figure 18 show the results from classifiers trained with Post-Adaptation trials (red curve) and then tested

either with the left out Post-Adaptation trials (green curve) and with the same number of Pre-Adaptation trials (blue curve). Even in this case only in IPS and not in V1 the classifier can discriminate between numerosties, both when tested with trials from Post- and from Pre-Adaptation dataset. However respect to the first model there is an important difference in IPS. Indeed when trained on Post-Adaptation trials, the model can classify the Post-Adaptation trials even better (around 35% accuracy) than the Pre-Adaptation trials (indeed the green curve in the second row is higher than the blue curve). This difference is significant for the numbers 20, 30 and 40 ($p<0.05$), close to significance for 60 ($p=0.07$) and not significant for 80. The classification accuracy in V1 is never significant.

Finally in the last rows of Figure 17 and Figure 18 the results from the 10 classes models are shown. The models have been trained (red curve) and then tested (green curve) with 10 different classes (5 Pre- and 5 Post-Adaptation) together. Note that now the chance level is 10%. The results observed in the 5 classes models are replicated in the 10 classes one, indeed the classifier in IPS can discriminate numbers. Moreover even in this third model the accuracy of the IPS classifier in discriminating the numbers in Post-Adaptation is higher than in Pre-Adaptation. This difference is significant for the numbers 20 and 60, and close to significant for 30 and 40 and not significant for 80. This is not the case for the classifier in V1, where a significant difference is never observed.

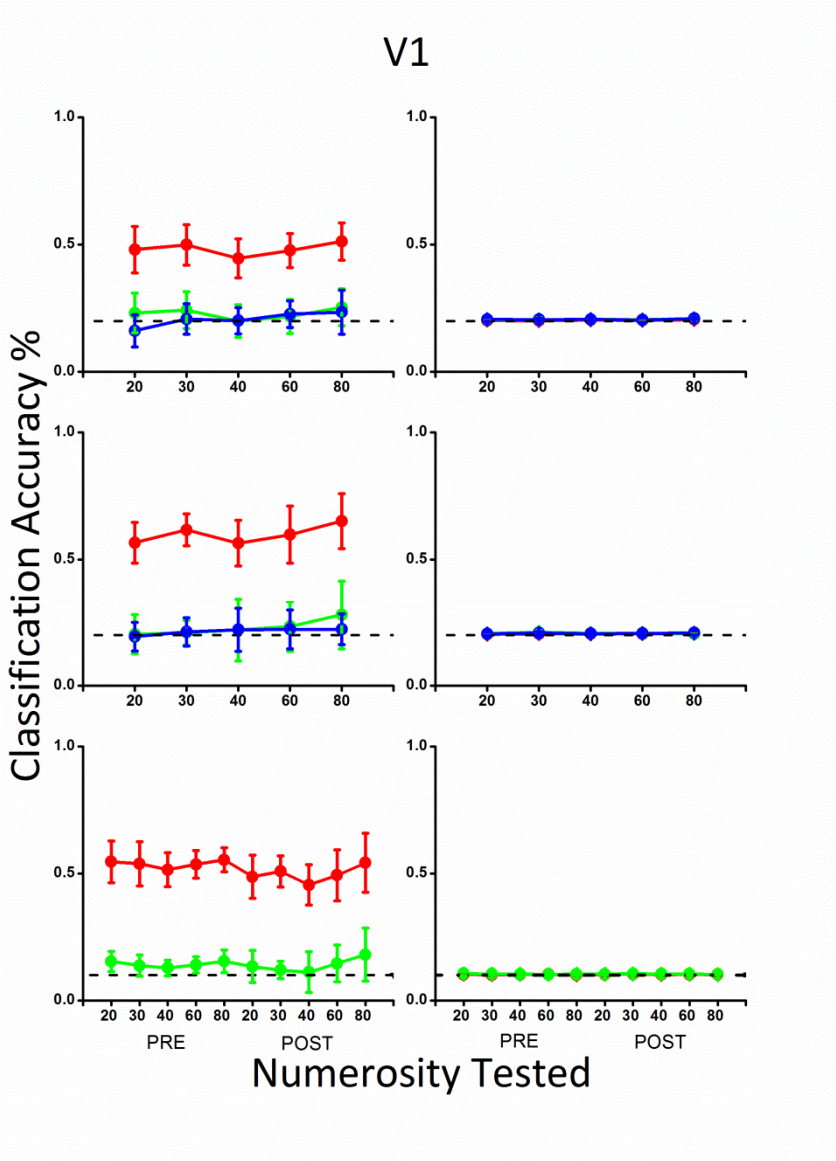

Figure 17 Results from the support vector classification for a foveal selection (not exceeding the extent of the stimulus, i.e. around 10 visual degrees) of the primary visual cortex defined on each subject's anatomy for the three models tested on dataset with correct (left column) and shuffled (right column) labels. Each curve represents the averaged results from the 50 spheres with the highest generalization accuracy when trained and tested to discriminate the different classes.The first row shows results for the 5 classes model trained on Pre-Adaptation trials (red curve) and then tested on the left-out PreAdaptation trials (green curve), and on the same numbers of Post-Adatpation trials (blue curve).

The second row shows results for the 5 classes model trained on Post-Adaptation trials (red curve) and then tested on the left-out PostAdaptation trials (green curve), and on the same numbers of Pre-Adatpation trials (blue curve).

The third row shows results for the 10 classes model trained on Pre- and Post-Adaptation trials (red curve) and then tested on the left-out Pre- and Post-Adaptation trials (green curve) together

The right column shows classification accuracy when the models were trained and tested on shuffled data.

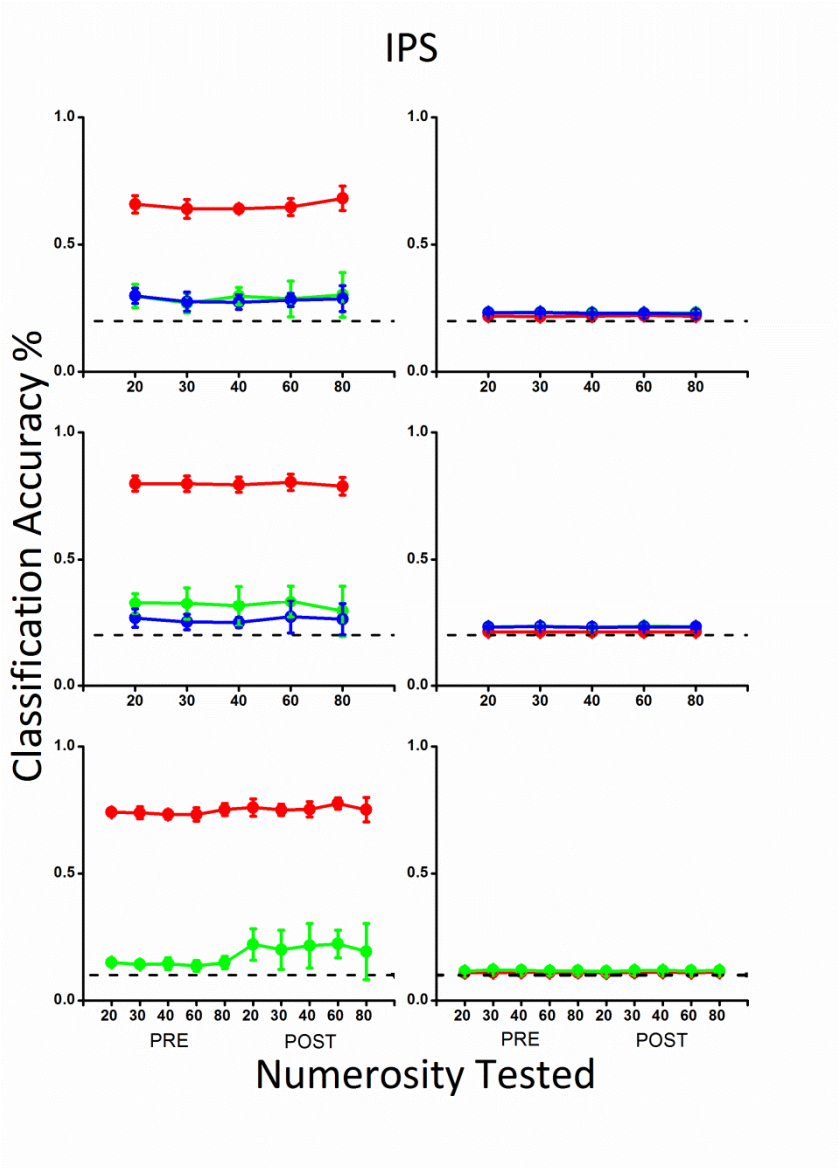

Figure 18 Results from the support vector classification for the parietal regions of interest defined on each subject's anatomy (around 1000 voxels) for the three models tested on dataset with correct (left column) and shuffled (right column) labels. Each curve represents the averaged results from the 50 spheres with the highest generalization accuracy when trained and tested to discriminate the different classes.
The first row shows results for the 5 classes model trained on Pre-Adaptation trials (red curve) and then tested on the left-out PreAdaptation trials (green curve), and on the same numbers of Post-Adatpation trials (blue curve).
The second row shows results for the 5 classes model trained on Post-Adaptation trials (red curve) and then tested on the left-out PostAdaptation trials (green curve), and on the same numbers of Pre-Adatpation trials (blue curve).
The third row shows results for the 10 classes model trained on Pre- and Post-Adaptation trials (red curve) and then tested on the left-out Pre- and Post-Adaptation trials (green curve) together
The right column shows classification accuracy when the models were trained and tested on shuffled data.

Finally we correlated the JND measured in the psychophysical experiment with the classifier's accuracy in IPS when the model was trained on Post-Adaptation trails and then tested either on the Post-Adaptation as well as on the Pre-Adaptation trials. The average JND measured after adaptation was significantly lower than the one measured before adaptation and this keeps step with the classifier significantly increased ability in discriminating Post-Adaptation respect to Pre-Adaptation trials (Figure 19).

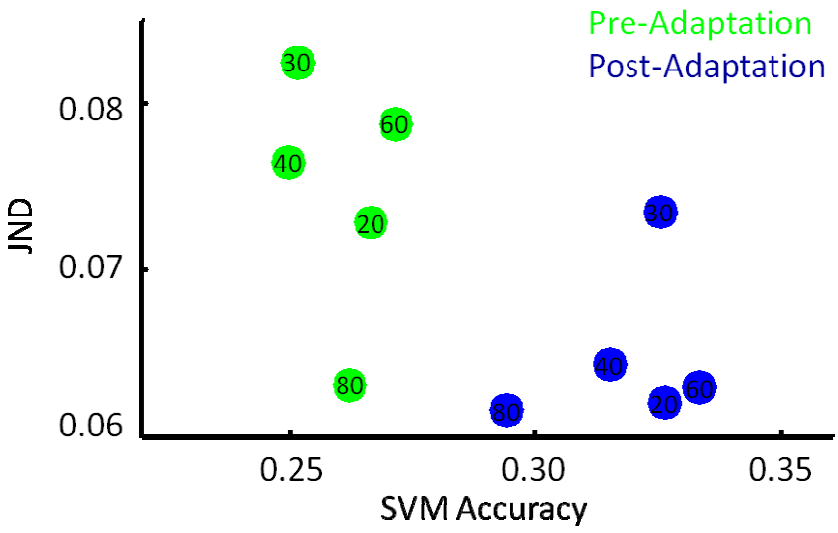

Figure 19 JND measured before (green dots) and after (blue dots) adaptation and support vector machine accuracy in discriminating the five numerosities when the classifier in IPS was trained with the Post-Adaptation dataset and then tested with the left-out Pre- (green dots) and with Post- (blue dots) Adaptation trials. The results shown are the average across 10 subjects.

3.4 Discussion

Previous experiments have demonstrated that it is possible to train a classifier to discriminate different numerosities based on the activation pattern elicited from numbers stimuli in the intraparietal cortex (Eger, Michel et al. 2009; Damarla and Just 2013). It has been also demonstrated that it is possible to record fMRI adaptation signal from this region (Piazza, Izard et al. 2004; Piazza, Pinel et al. 2007), suggesting that the same number selective neurons recoded in the monkey are present in humans as well. The results of fMRI adaptation signal studies as well as single cells recording studies (Nieder, Freedman et al. 2002; Nieder and Miller 2004) have shown that the number selective neurons follow Weber law and have recently been described to have a numerotopic organization of the intraparietal region (Harvey, Klein et al. 2013). Overall these evidence are suggesting that humans, as well as monkeys, have a number sense which is supported by a neural substrate dedicated specifically to analysing numbers, even in the presence of conflicting other visual cues. As with other primary senses, the number sense is susceptible to adaptation (Burr and Ross 2008; Ross and Burr 2010), however it is still under debate if what has been described as number adaptation could instead just be density adaptation, explaining these findings relative to numbers just in terms of area and texture computation (Durgin 1995; Durgin and Huk 1997; Durgin 2008; Dakin, Tibber et al. 2011; Tibber, Greenwood et al. 2012).

In the present experiment we tried to disentangle number and texture by keeping the energy of the stimuli constant for all the numerosities examined and evaluating the number adaptation effect on the ability of a classifier to

discriminate numbers pre- and post- adaptation in the well known intra parietal number regions as well as in the primary visual cortex.

We proposed a new adaptation paradigm placing a long pause between adapter and test stimuli that let to differentiate the activity associated with the deviant stimuli from the activity due to the habituating stimuli which, in the habituation paradigms, is usually temporally convoluted. Moreover in the present paradigm, by presenting multiple numerosities in the testing period, we removed any salience and novelty effects potentially present in previous habituation paradigm where the presentation of the deviant stimuli represented a sudden, novel stimulus.

The psychophysical results are showing that the new paradigm is able to produce number adaptation, even if the amount of effect recorded is smaller (the underestimation is around 10%) respect to the classical paradigm (reporting a 30% underestimation) where both the stimuli are presented in periphery and with a different timing.

We used this paradigm in a rapid event related experiment and we showed that BOLD signal differences between Pre- and Post- Adaptation where present only in the intra parietal sulcus bilaterally, leaving the primary visual cortex unaffected.

This result is in line with the expectations, since the long temporal gap introduced between the adapting and testing stimuli was supposed to dissociate the effect of number adaptation from the contrast one by letting the BOLD signal to return to baseline.Moreover different models have been trained using support vector machine and in all cases V1 was never able to generalize to the different numerosities.

Classifiers in IPS were instead always able to generalize when tested with unseen trials left out from the data set they have been trained with as well

as when tested with unseen trials of the other adaptation condition. However if on one hand the classifier trained on Pre-Adaptation dataset was able to discriminate the Pre-Adaptation trials as well as the Post-Adaptation one, on the other hand the classifier trained on Post adaptation dataset was able to classify better the Post-Adaptation respect to the Pre-Adaptation trials.

This means that when trained on Pre-Adaptation dataset the classifier can decode equally well Pre- and Post-Adaptation testing trials. Instead when trained on Post-Adaptation dataset the classifies can decode very well the trials coming from the adaptation condition it has been trained on (and with an higher accuracy respect to the one trained on Pre-Adapatation), but it can decode worse the testing trials coming from the Pre-Adapatation dataset.

This could be due to a increased tuning selectivity of the numerosity detectors after adaptation.

Indeed if we suppose that the tuning of these number detectors is broader before adaptation, this could result in a less clear pattern characterization of the different numerosities and in more wide boundaries defining the differences between one stimulus and another. This set of detector could generalize to equally blurred exemplars as well as to better defined and distinct pattern arising from the Post-Adaptation dataset. The latter would perfectly fit in the less strictly defined exemplars coming from the Pre-Adaptation dataset. On the other hand if we suppose and increased tuning selectivity in the numbers detectors after adaptation, a classifier trained on Post-Adaptation trials would have learned a sharper and better defined distinction between patterns associated with the different numerosities. This would let the classifier generalize very well to the left out Post-Adaptation trials, which would be as sharply distinct as the exemplars it has been trained

on, but it could turn out at the same time in a slightly worst, although still possible, generalization to the more noisy Pre-Adaptation trials.

The increase in classification accuracy in discriminating Post-Adaptation trials when the models is trained on Post-Adaptation dataset fits well with the decrease in JND after adaptation, suggesting an increased precision in numbers misclassification after adaptation which has been observed in the present experiment. A trend in JND increase can also be observed in the study of Burr and Ross (2008) using the classical number adaptation paradigm, although the authors did not explicitly reported the data. The increase in JND after adaptation is suggestive of mechanisms that dynamically adapt to improve discrimination around the adapting point. Interestingly the maximum effect is for the numerosity of 60, close to the adapter numerosity.

The increase in sensitivity after adaptation has been described also in other vision domains, this is the case of speed and direction adaptation. Kohn and Movshon (2004) showed that direction adaptation in MT neurons causes a narrowing of tuning bandwidth and a shift in preference toward the adapted direction. Krekelberg (2006) showed that brief adaptation to a moving stimulus caused either a reduction of the apparent absolute speed of the test stimulus and an increase in the sensitivity to discriminate the speeds that prevail in the current environment. In this experiment reduced magnitude of neural responses and reduced width of speed tuning curves have been recorded from MT cells after speed adaptation suggesting that this region optimizes the encoding of relative speed at the cost of incorrectly encoding absolute speed. Both these experiments reported adaptation to cause the MT tuning curve to shift toward the adaptor, rather than to be repelled from it as in V1. Such shifts can lead to a relative enhancement of

the representation of frequently occurring stimuli, a potential neural basis for the changes in likelihood function required by Bayesian explanations for perceptual effects (Clifford, Webster et al. 2007).

In conclusion in this experiment we verified that the novel number adaptation paradigm used is able to cause a decrease in PSE and an increase in JND suggesting an enhanced tuning refinement after adaptation. Moreover in line with previous neuroimaging and neurophsyisiolgical findings, this experiment support the idea of the possible existence of number selective neurons in the human IPS that are susceptible to adaptation. The adaptation effect is specifically affecting these parietal neurons and cannot be simply explained by adaptation to other low level visual properties of the stimulus. Indeed V1 is not even able to distinguish reliably between energy matched stimuli of different numerosities.

4. General Discussion

In this thesis I have analysed the neural mechanisms underlying visual object segmentation and enumeration in humans using fMRI.

Object segmentation is based on form identification, a task which has been traditionally linked to achromatic detectors. However increasing evidence are questioning the modular view of a clear segregation between colour and form being analysed by parallel streams coming from the retina. Instead it seems that visual cortex's response to colour is strongly linked to other visual properties like lightness contrast, shape, texture, orientation, showing reciprocal interactions between form and colour in human vision.

The visual system can identify contours of an object and segregate it from the background thanks to its selectivity to spatial phase.

The earlier psychophysical evidence (Burr, Morrone et al. 1989; Burr, Morrone et al. 1992; Girard and Morrone 1995; Martini, Girard et al. 1996) suggesting that a spatial phase analysis is performed also for purely colour information are now supported by neurophysiological data (Johnson, Hawken et al. 2008) showing the existence of chromatic neurons with phase sensitive oriented receptive fields in primary visual area which are ideally suited for colour shape analysis.

To test this alternative view the first experiment presented investigated the link between form and colour perception by studying the selectivity to spatial phase of chromatic cortical mechanisms.

The current study is showing that possibly an equal number of chromatic even- and odd- symmetry receptive fields are already present in human primary visual cortex and that colour defined shape analysis based on spatial phase coding is carried on either along the dorsal (V3A, dorsal LO, the caudal

part of the occipital branch of the intraparietal sulcus, TOS) and the ventral (V4) stream. Like it happens for luminance, even if phase analysis starts already from V1, only at higher level it is possible to discriminate absolute phase, that is to differentiate between edges and lines, suggesting a role of these hierarchically superior areas in contour detection.

In our experiment V4 has been found not to have a crucial role in phase detection of chromatic features, but to be only part of a network for the colour form analysis.

The modular approach assigned to this areas the fame of being the "colour center" mainly because of patients with achromatopsia (Zeki 1990). Cerebral achromatopsia is the name given to the condition caused by lesions in ventral occipital cortex that cause a loss of the ability to recognize colours without the loss of the perception of form and motion.

However it has been demonstrated that this area have also important role in shape perception. For example Schein et al. (1990) showed that most V4 cells responded to shape cues as much as, if not more than to colour. Moreover, they reported that most V4 cells were colour-luminance cells, responding almost equally well to brightness as to colour. Moreover lesions in macaque V4 cause deficits in shape discrimination (Walsh, Butler et al. 1992).

Hence the functional role of V4 is possibly to integrate colour and form information for perception, other than contribute to colour perception per se.

On the other hand our experiment shows an important involvement of dorsal areas in detecting salient features modulated both in colour and in luminance. Having mechanisms able to detect contours on the basis of both luminance and colour information at the same time would be more efficient

in natural conditions where these two attribute are used to vary together (Hansen and Gegenfurtner 2009).

Previous studies have already confirmed a role in dorsal areas in both contour detection (Schira, Fahle et al. 2004), brightness perception (Perna, Tosetti et al. 2005; Cornelissen, Wade et al. 2006) and colour processing (Liu and Wandell 2005; Mullen, Thompson et al. 2010; D'Souza, Auer et al. 2011).

We speculated that possibly dorsal areas are involved in the analysis of colour form to mediate action toward shapes, being them defined either on the base of luminance- or chromatic contrast.

The parietal region is known to be constituted by a mosaic of distinct specialized regions which are activated following a posterior to anterior gradient by many visuospatial tasks, including saccades, grasping, pointing, hand reaching, as well as other more "abstract" tasks like calculation and phoneme detection (Culham and Kanwisher 2001; Simon, Mangin et al. 2002; Simon, Kherif et al. 2004).

However along the intraparietal sulcus there are not only highly specialized subregions, but also areas of inter-dimensional overlap. For example Pinel et al. (2004) investigated the brain activities elicited by comparative judgements on three different dimensions: number size, physical size and luminance. The authors found an interaction between numbers (Arabic digits) and size and between size and luminance, even if not between numbers and luminance. This evidence support the existence of either partially specialized and partially overlapping representation of numbers, size and luminance in a wider network linking the bilateral intraparietal sulci, precentral and occipitotemporal regions. This experiment showed more generally that the intraparietal sulcus is not only activated for

number processing but it is also involved in judgment of magnitude regardless if a numerical or a physical comparison is required.

An interesting theory postulates the existence of a single magnitude system responsible for the approximate computation of quantity, be it number, space or time, whose neural substrate relays presumably on the posterior parietal cortex (Walsh 2003).

The intraparietal sulcus is known to be involved in the preparation of eye movement and in the maintenance of perceptual stability (Duhamel, Colby et al. 1992).

Supporting this view, two recent experiments showed that numerical processing (both for symbolic and non symbolic stimuli) underwent to similar misperception at the time of saccade, as those previously observed in the space and time domains. Indeed visual stimuli briefly flashed just before a saccade are systematically compressed in both space and time (Morrone, Ross et al. 1997; Ross, Morrone et al. 1997; Morrone, Ross et al. 2005; Binda, Cicchini et al. 2009).The receptive fields of visual neurons at the intra-parietal level shift in anticipation of saccades, displacing visual spatial representations (Duhamel, Colby et al. 1992).

Confirming the link between space, time and number, the experiment from Binda et al. (2011) show that large numerosities are constantly underestimated during saccades. Hence space, time and number undergo to similar distortion at the time of saccade causing a compression of stimulus magnitude.

The link between number processing, eye position and space coding has been confirmed also by Knops et al. (2009) . The authors trained a classifier to predict the direction of eye movement, left or right, and then (without

further training!) showing that it can generalize to mental arithmetical task (subtraction or addition) which was being performed.

These studies also support the idea that the parietal lobe could contribute to the representation of space and number in the form of a mental "number line" (Dehaene 1993).

A frequent issue in studying number perception is that the number of elements contained in an image varies together with other visual properties like area and density, which per se could give enough information to infer the numerosity (Durgin 1995; Durgin and Huk 1997; Durgin 2008; Dakin, Tibber et al. 2011; Tibber, Greenwood et al. 2012).

This view also explain the effect of number adaptation in terms of texture denisty adaptation, likely due to early cortical texture analysers (presumably simple and complex cells in primary visual cortex) rather than abstract numeric representation (Durgin and Huk 1997).

The first experiment presented in this thesis, as well as other studies in literature (Perna, Tosetti et al. 2005; Perna, Tosetti et al. 2008) show how shape perception relies on object segmentation which is based on the analysis of the energy of the image, both for luminance and for coloured stimuli. In the second experiment described in this thesis I have shown that controlling for energy let us to study another important property of the image, the numerosity, overcoming the texture issue.

Indeed results suggest that the perception of numerosity takes place in IPS, and that the primary visual area is not even able to discriminate between different numerosities, once the energy of the stimuli is balanced.

Moreover the effect of adaptation to numerosity is able to alter activation patterns in the intraparietal sulcus, presumably changing the tuning of the number selective neurons.

A recent experiment (Harvey, Klein et al. 2013) described topographic representation of numbers in the intra parietal sulcus demonstrating how this organization, common in sensory and motor areas, can be applyed also to abstract features like numerosity, suggesting that the distinction between primary, topographic and abstract representation of higher cognitive functions must be reconsidered.

The results of the experiment presented in this thesis let to conclude that numerosity is a primary property of the image, and that numerosity adaptation does not rely on low-level stimulus parameters but that it alters higher order representation of magnitude.

5. Acknowledgement

I would like to thank Prof.ssa Morrone Maria Concetta and Prof. David Burr for their invaluable support and for guiding me through the course of my PhD and this thesis. I want to thank them also for letting me having a very useful experience working for the last 6 months of my PhD in Paris in the Cognitive Neuroimaging Unit, directed by Prof. Dehaene.

I am really thankful also for the colleagues who worked with me during these years, taking part to the experiments presented in this thesis.

In particular I want to thank Dott. Cicchini Guido Marco, Dott.ssa Frija Francesca, Dott. Montanaro Domenico, Dott.ssa Tosetti Michela for helping and teaching.

I don't have enough words to thank Dott. Murphy David who have been a great friend other than an exceptionally good colleague. I think we managed to create a solid team, which is going to hold for life.

I am really thankful to all my colleagues of the PisaVision Lab for the friendly atmosphere they all created while working.

I would like to thank my family and my friends without whom I would never be the person I am now and I would never get up to here.

Finally a very special thanks goes to Luca who has always been there for me, supporting and encouraging me with great love throughout these, the previous and, hopefully, the future years.

6. References

Binda, P., G. M. Cicchini, et al. (2009). "Spatiotemporal distortions of visual perception at the time of saccades." J Neurosci **29**(42): 13147-13157.

Binda, P., M. C. Morrone, et al. (2011). "Underestimation of perceived number at the time of saccades." Vision Res **51**(1): 34-42.

Brainard, D. H. (1997). "The Psychophysics Toolbox." Spat Vis **10**(4): 433-436.

Brannon, E. M., M. E. Libertus, et al. (2008). "Electrophysiological measures of time processing in infant and adult brains: Weber's Law holds." J Cogn Neurosci **20**(2): 193-203.

Burr, D. and J. Ross (2008). "A visual sense of number." Curr Biol **18**(6): 425-428.

Burr, D. C., M. C. Morrone, et al. (1992). "Electro-physiological investigation of edge-selective mechanisms of human vision." Vision Res **32**(2): 239-247.

Burr, D. C., M. C. Morrone, et al. (1989). "Evidence for edge and bar detectors in human vision." Vision Res **29**(4): 419-431.

Cantlon, J. F., E. M. Brannon, et al. (2006). "Functional imaging of numerical processing in adults and 4-y-old children." PLoS Biol **4**(5): e125.

Cantlon, J. F., M. E. Libertus, et al. (2009). "The neural development of an abstract concept of number." J Cogn Neurosci **21**(11): 2217-2229.

Castaldi, E., F. Frijia, et al. (2013). "BOLD human responses to chromatic spatial features." Eur J Neurosci **38**(2): 2290-2299.

Clifford, C. W., M. A. Webster, et al. (2007). "Visual adaptation: neural, psychological and computational aspects." Vision Res **47**(25): 3125-3131.

Cornelissen, F. W., A. R. Wade, et al. (2006). "No functional magnetic resonance imaging evidence for brightness and color filling-in in early human visual cortex." J Neurosci **26**(14): 3634-3641.

Crespi, S., L. Biagi, et al. (2011). "Spatiotopic coding of BOLD signal in human visual cortex depends on spatial attention." PLoS One **6**(7): e21661.

Culham, J. C. and N. G. Kanwisher (2001). "Neuroimaging of cognitive functions in human parietal cortex." Curr Opin Neurobiol **11**(2): 157-163.

d'Avossa, G., M. Tosetti, et al. (2007). "Spatiotopic selectivity of BOLD responses to visual motion in human area MT." Nat Neurosci **10**(2): 249-255.

D'Souza, D. V., T. Auer, et al. (2011). "Temporal frequency and chromatic processing in humans: an fMRI study of the cortical visual areas." J Vis **11**(8).

Dakin, S. C., M. S. Tibber, et al. (2011). "A common visual metric for approximate number and density." Proc Natl Acad Sci U S A **108**(49): 19552-19557.

Damarla, S. R. and M. A. Just (2013). "Decoding the representation of numerical values from brain activation patterns." Hum Brain Mapp **34**(10): 2624-2634.

Dehaene, S., Bossini, S., & Giraux, P. (1993). "The mental representation of parity and numerical magnitude. ." Journal of Experimental Psychology: General **122**: 371–396.

Dehaene, S. and L. Cohen (1997). "Cerebral pathways for calculation: double dissociation between rote verbal and quantitative knowledge of arithmetic." Cortex **33**(2): 219-250.

Duhamel, J. R., C. L. Colby, et al. (1992). "The updating of the representation of visual space in parietal cortex by intended eye movements." Science **255**(5040): 90-92.

Durgin, F. H. (1995). "Texture density adaptation and the perceived numerosity and distribution of texture." Journal of Experimental Psychology Human Perception & Performance **21**(1): 149-169.

Durgin, F. H. (2008). "Texture density adaptation and visual number revisited." Curr Biol **18**(18): R855-856; author reply R857-858.

Durgin, F. H. and A. C. Huk (1997). "Texture density aftereffects in the perception of artificial and natural textures." Vision Res **37**(23): 3273-3282.

Efron, B. T., and Tibshirani, R. J. (1993). An Introduction to the Bootstrap, Chapman and Hall.
Eger, E., C. A. Kell, et al. (2008). "Graded size sensitivity of object-exemplar-evoked activity patterns within human LOC subregions." J Neurophysiol **100**(4): 2038-2047.
Eger, E., V. Michel, et al. (2009). "Deciphering cortical number coding from human brain activity patterns." Curr Biol **19**(19): 1608-1615.
Engel, S. A., G. H. Glover, et al. (1997). "Retinotopic organization in human visual cortex and the spatial precision of functional MRI." Cereb Cortex **7**(2): 181-192.
Feigenson, L., S. Dehaene, et al. (2004). "Core systems of number." Trends Cogn Sci **8**(7): 307-314.
Felsen, G., J. Touryan, et al. (2005). "Cortical sensitivity to visual features in natural scenes." PLoS Biol **3**(10): e342.
Fias, W., J. Lammertyn, et al. (2003). "Parietal representation of symbolic and nonsymbolic magnitude." J Cogn Neurosci **15**(1): 47-56.
Frank, M. C., D. L. Everett, et al. (2008). "Number as a cognitive technology: evidence from Piraha language and cognition." Cognition **108**(3): 819-824.
Gallistel, C. R. and I. I. Gelman (2000). "Non-verbal numerical cognition: from reals to integers." Trends Cogn Sci **4**(2): 59-65.
Girard, P. and M. C. Morrone (1995). "Spatial structure of chromatically opponent receptive fields in the human visual system." Vis Neurosci **12**(1): 103-116.
Gordon, P. (2004). "Numerical cognition without words: evidence from Amazonia." Science **306**(5695): 496-499.
Grill-Spector, K. and R. Malach (2004). "The human visual cortex." Annu Rev Neurosci **27**: 649-677.
Hadjikhani, N., A. K. Liu, et al. (1998). "Retinotopy and color sensitivity in human visual cortical area V8." Nat Neurosci **1**(3): 235-241.
Hansen, K. A., K. N. Kay, et al. (2007). "Topographic organization in and near human visual area V4." J Neurosci **27**(44): 11896-11911.
Hansen, T. and K. R. Gegenfurtner (2009). "Independence of color and luminance edges in natural scenes." Vis Neurosci **26**(1): 35-49.
Harvey, B. M., B. P. Klein, et al. (2013). "Topographic representation of numerosity in the human parietal cortex." Science **341**(6150): 1123-1126.
Haynes, J. D. and G. Rees (2005). "Predicting the orientation of invisible stimuli from activity in human primary visual cortex." Nat Neurosci **8**(5): 686-691.
Haynes, J. D. and G. Rees (2006). "Decoding mental states from brain activity in humans." Nat Rev Neurosci **7**(7): 523-534.
Henriksson, L., A. Hyvarinen, et al. (2009). "Representation of cross-frequency spatial phase relationships in human visual cortex." J Neurosci **29**(45): 14342-14351.
Hubel, D. H. and T. N. Wiesel (1962). "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex." J Physiol **160**: 106-154.
Huntley-Fenner, G. (2001). "Children's understanding of number is similar to adults' and rats': numerical estimation by 5--7-year-olds." Cognition **78**(3): B27-40.
Hyde, D. C. and E. S. Spelke (2011). "Neural signatures of number processing in human infants: evidence for two core systems underlying numerical cognition." Dev Sci **14**(2): 360-371.
Izard, V., G. Dehaene-Lambertz, et al. (2008). "Distinct cerebral pathways for object identity and number in human infants." PLoS Biol **6**(2): e11.
Jacob, S. N. and A. Nieder (2009). "Tuning to non-symbolic proportions in the human frontoparietal cortex." European Journal of Neuroscience **30**(7): 1432-1442.
Johnson, E. N., M. J. Hawken, et al. (2001). "The spatial transformation of color in the primary visual cortex of the macaque monkey." Nat Neurosci **4**(4): 409-416.
Johnson, E. N., M. J. Hawken, et al. (2004). "Cone inputs in macaque primary visual cortex." J Neurophysiol **91**(6): 2501-2514.
Johnson, E. N., M. J. Hawken, et al. (2008). "The orientation selectivity of color-responsive neurons in macaque V1." J Neurosci **28**(32): 8096-8106.

Jordan, K. E. and E. M. Brannon (2006). "A common representational system governed by Weber's law: nonverbal numerical similarity judgments in 6-year-olds and rhesus macaques." J Exp Child Psychol **95**(3): 215-229.
Jordan, K. E. and E. M. Brannon (2006). "Weber's Law influences numerical representations in rhesus macaques (Macaca mulatta)." Anim Cogn **9**(3): 159-172.
Kamitani, Y. and F. Tong (2005). "Decoding the visual and subjective contents of the human brain." Nat Neurosci **8**(5): 679-685.
Kamitani, Y. and F. Tong (2006). "Decoding seen and attended motion directions from activity in the human visual cortex." Curr Biol **16**(11): 1096-1102.
Kaplan, E. and R. M. Shapley (1986). "The primate retina contains two types of ganglion cells, with high and low contrast sensitivity." Proc Natl Acad Sci U S A **83**(8): 2755-2757.
Knops, A., B. Thirion, et al. (2009). "Recruitment of an area involved in eye movements during mental arithmetic." Science **324**(5934): 1583-1585.
Kohn, A. and J. A. Movshon (2004). "Adaptation changes the direction tuning of macaque MT neurons." Nat Neurosci **7**(7): 764-772.
Krekelberg, B., G. M. Boynton, et al. (2006). "Adaptation: from single cells to BOLD signals." Trends in Neurosciences **29**(5): 250-256.
Krekelberg, B., R. J. van Wezel, et al. (2006). "Adaptation in macaque MT reduces perceived speed and improves speed discrimination." J Neurophysiol **95**(1): 255-270.
Kriegeskorte, N., E. Formisano, et al. (2007). "Individual faces elicit distinct response patterns in human anterior temporal cortex." Proc Natl Acad Sci U S A **104**(51): 20600-20605.
Levine, D., J. Warach, et al. (1985). "Two visual systems in mental imagery: dissociation of "what" and "where" in imagery disorders due to bilateral posterior cerebral lesions." Neurology **35**(7): 1010-1018.
Lipton, J. S. and E. S. Spelke (2003). "Origins of number sense. Large-number discrimination in human infants." Psychol Sci **14**(5): 396-401.
Liu, J. and B. A. Wandell (2005). "Specializations for chromatic and temporal signals in human visual cortex." J Neurosci **25**(13): 3459-3468.
Mahmoudi, A., S. Takerkart, et al. (2012). "Multivoxel pattern analysis for FMRI data: a review." Comput Math Methods Med **2012**: 961257.
Marr, D. (1976). "Early processing of visual information." Phil. Trans. R. Soc. Lond.: 485-526.
Martini, P., P. Girard, et al. (1996). "Sensitivity to spatial phase at equiluminance." Vision Res **36**(8): 1153-1162.
Mechler, F., I. E. Ohiorhenuan, et al. (2007). "Speed dependence of tuning to one-dimensional features in V1." J Neurophysiol **97**(3): 2423-2438.
Mechler, F., D. S. Reich, et al. (2002). "Detection and discrimination of relative spatial phase by V1 neurons." J Neurosci **22**(14): 6129-6157.
Morrone, M. C. and D. C. Burr (1988). "Feature detection in human vision: a phase-dependent energy model." Proc R Soc Lond B Biol Sci **235**(1280): 221-245.
Morrone, M. C. and D. C. Burr (1997). "Capture and transparency in coarse quantized images." Vision Res **37**(18): 2609-2629.
Morrone, M. C., D. C. Burr, et al. (1993). "Development of infant contrast sensitivity to chromatic stimuli." Vision Res **33**(17): 2535-2552.
Morrone, M. C., J. Ross, et al. (2005). "Saccadic eye movements cause compression of time as well as space." Nat Neurosci **8**(7): 950-954.
Morrone, M. C., J. Ross, et al. (1997). "Apparent position of visual targets during real and simulated saccadic eye movements." J Neurosci **17**(20): 7941-7953.
Movshon, J. A., I. D. Thompson, et al. (1978). "Receptive field organization of complex cells in the cat's striate cortex." J Physiol **283**: 79-99.
Movshon, J. A., I. D. Thompson, et al. (1978). "Spatial summation in the receptive fields of simple cells in the cat's striate cortex." J Physiol **283**: 53-77.

Mullen, K. T., B. Thompson, et al. (2010). "Responses of the human visual cortex and LGN to achromatic and chromatic temporal modulations: an fMRI study." J Vis **10**(13): 13.
Nieder, A. (2013). "Coding of abstract quantity by 'number neurons' of the primate brain." J Comp Physiol A Neuroethol Sens Neural Behav Physiol **199**(1): 1-16.
Nieder, A. and S. Dehaene (2009). "Representation of number in the brain." Annu Rev Neurosci **32**: 185-208.
Nieder, A., D. J. Freedman, et al. (2002). "Representation of the quantity of visual items in the primate prefrontal cortex." Science **297**(5587): 1708-1711.
Nieder, A. and K. Merten (2007). "A labeled-line code for small and large numerosities in the monkey prefrontal cortex." J Neurosci **27**(22): 5986-5993.
Nieder, A. and E. K. Miller (2004). "Analog numerical representations in rhesus monkeys: evidence for parallel processing." J Cogn Neurosci **16**(5): 889-901.
Nieder, A. and E. K. Miller (2004). "A parieto-frontal network for visual numerical information in the monkey." Proc Natl Acad Sci U S A **101**(19): 7457-7462.
Norman, K. A., S. M. Polyn, et al. (2006). "Beyond mind-reading: multi-voxel pattern analysis of fMRI data." Trends Cogn Sci **10**(9): 424-430.
Olman, C. A., K. Ugurbil, et al. (2004). "BOLD fMRI and psychophysical measurements of contrast response to broadband images." Vision Res **44**(7): 669-683.
Perna, A. and M. C. Morrone (2007). "The lowest spatial frequency channel determines brightness perception." Vision Res **47**(10): 1282-1291.
Perna, A., M. Tosetti, et al. (2005). "Neuronal mechanisms for illusory brightness perception in humans." Neuron **47**(5): 645-651.
Perna, A., M. Tosetti, et al. (2008). "BOLD response to spatial phase congruency in human brain." J Vis **8**(10): 15 11-15.
Perona, P. and J. Malik (1990). "Detecting and localizing edges composed of steps, peaks and roofs."
Piazza, M., V. Izard, et al. (2004). "Tuning curves for approximate numerosity in the human intraparietal sulcus." Neuron **44**(3): 547-555.
Piazza, M., P. Pinel, et al. (2007). "A magnitude code common to numerosities and number symbols in human intraparietal cortex." Neuron **53**(2): 293-305.
Pica, P., C. Lemer, et al. (2004). "Exact and approximate arithmetic in an Amazonian indigene group." Science **306**(5695): 499-503.
Pinel, P., S. Dehaene, et al. (2001). "Modulation of parietal activation by semantic distance in a number comparison task." Neuroimage **14**(5): 1013-1026.
Pinel, P., M. Piazza, et al. (2004). "Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments." Neuron **41**(6): 983-993.
Pollen, D. A. and S. F. Ronner (1981). "Phase relationships between adjacent simple cells in the visual cortex." Science **212**(4501): 1409-1411.
Press, W. A., A. A. Brewer, et al. (2001). "Visual areas and spatial summation in human visual cortex." Vision Res **41**(10-11): 1321-1332.
Rainer, G., M. Augath, et al. (2001). "Nonmonotonic noise tuning of BOLD fMRI signal to natural images in the visual cortex of the anesthetized monkey." Curr Biol **11**(11): 846-854.
Rainer, G., M. Augath, et al. (2002). "The effect of image scrambling on visual cortical BOLD activity in the anesthetized monkey." Neuroimage **16**(3 Pt 1): 607-616.
Ross, J. and D. Burr (2012). "Number, texture and crowding." Trends Cogn Sci **16**(4): 196-197.
Ross, J. and D. C. Burr (2010). "Vision senses number directly." J Vis **10**(2): 10 11-18.
Ross, J., M. C. Morrone, et al. (1997). "Compression of visual space before saccades." Nature **386**(6625): 598-601.
Schein, S. J. and R. Desimone (1990). "Spectral properties of V4 neurons in the macaque." J Neurosci **10**(10): 3369-3389.
Schira, M. M., M. Fahle, et al. (2004). "Differential contribution of early visual areas to the perceptual process of contour processing." J Neurophysiol **91**(4): 1716-1721.

Sereno, M. I., A. M. Dale, et al. (1995). "Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging." Science **268**(5212): 889-893.
Simon, O., F. Kherif, et al. (2004). "Automatized clustering and functional geometry of human parietofrontal networks for language, space, and number." Neuroimage **23**(3): 1192-1202.
Simon, O., J. F. Mangin, et al. (2002). "Topographical layout of hand, eye, calculation, and language-related areas in the human parietal lobe." Neuron **33**(3): 475-487.
Spitzer, H. and S. Hochstein (1985). "A complex-cell receptive-field model." J Neurophysiol **53**(5): 1266-1286.
Starkey, P., E. S. Spelke, et al. (1990). "Numerical abstraction by human infants." Cognition **36**(2): 97-127.
Stoianov, I. and M. Zorzi (2012). "Emergence of a 'visual number sense' in hierarchical generative models." Nat Neurosci **15**(2): 194-196.
Talairach, J., & Tournoux, P. (1988). Co-Planar Stereotaxic Atlas of the Human Brain. . New York, Thieme Medical Publishers.
Temple, E. and M. I. Posner (1998). "Brain mechanisms of quantity are similar in 5-year-old children and adults." Proc Natl Acad Sci U S A **95**(13): 7836-7841.
Tibber, M. S., J. A. Greenwood, et al. (2012). "Number and density discrimination rely on a common metric: Similar psychophysical effects of size, contrast, and divided attention." J Vis **12**(6): 8.
Tjan, B. S., V. Lestou, et al. (2006). "Uncertainty and invariance in the human visual cortex." J Neurophysiol **96**(3): 1556-1568.
Tootell, R. B. and N. Hadjikhani (2001). "Where is 'dorsal V4' in human visual cortex? Retinotopic, topographic and functional evidence." Cereb Cortex **11**(4): 298-311.
Wade, A. R., A. A. Brewer, et al. (2002). "Functional measurements of human ventral occipital cortex: retinotopy and colour." Philos Trans R Soc Lond B Biol Sci **357**(1424): 963-973.
Walsh, V. (2003). "A theory of magnitude: common cortical metrics of time, space and quantity." Trends Cogn Sci **7**(11): 483-488.
Walsh, V., S. R. Butler, et al. (1992). "The effects of V4 lesions on the visual abilities of macaques: shape discrimination." Behav Brain Res **50**(1-2): 115-126.
Wandell, B. A., A. A. Brewer, et al. (2005). "Visual field map clusters in human cortex." Philos Trans R Soc Lond B Biol Sci **360**(1456): 693-707.
Watson, A. B. and D. G. Pelli (1983). "QUEST: a Bayesian adaptive psychometric method." Percept Psychophys **33**(2): 113-120.
Xu, F. and E. S. Spelke (2000). "Large number discrimination in 6-month-old infants." Cognition **74**(1): B1-B11.
Zeki, S. (1990). "A century of cerebral achromatopsia." Brain **113 (Pt 6)**: 1721-1777.