

The Neural Correlates of Perceptual Awareness

Abstract

The study of the neural correlates of awareness is nowadays an active research field in Neuroscience. This has been basically boosted by the study of neural correlates of conscious perception with single cell recordings in monkeys and voxel activities with human fMRI experiments. In this review, we discuss the main experiments with recording of single neurons and related evidence about the neural events underling visual perceptual awareness.

Alberto Capurroⁱ and Rodrigo Quian Quirogaⁱⁱ
*University of Leicester
Department of Engineering
NeuroEngineering Labⁱⁱⁱ*

Several researchers are now experimentally addressing the study of the neural correlates of the different aspects of consciousness. This issue was considered to be an exclusive matter of philosophy until a short time ago (Crick, 1994; Crick and Koch, 1995; Koch, 2004). Consciousness or awareness can, in principle, be defined as the occurrence of perceptions, thoughts and feelings (Sutherland, 1996). A precise definition, however, remains elusive, because covering every aspect of the conscious experience is difficult. For instance, speaking about the perception of something that we see and recognize is not the same as speaking about our self consciousness, or about the different levels of awareness that we experience during the sleep-wake cycle. To begin with, we must distinguish between the factors that enable consciousness to occur (such as the ascending activation system of the brainstem and basal forebrain, or the glutamatergic neurotransmission) from the so called content of consciousness, that is, the formation of specific percepts. In this review we will address this last issue through evidence of neural correlates underling visual perceptual awareness.

ⁱ ac331@leicester.ac.uk

ⁱⁱ rqqg1@leicester.ac.uk

ⁱⁱⁱ LE1 7RH Leicester, United Kingdom
Tel / Fax: +44 (0)116 252 2314 / 2619
www.le.ac.uk/neuroengineering

Perception and Neural Coding in the Visual Pathways

The processing of visual information in the cerebral cortex takes place through two hierarchically organized neural streams with extensive feedback. The dorsal or action pathway, going from the V1 area in the striate occipital cortex to the parietal and frontal cortices, processes information related to the position of an object as required for performing an action, such as reaching for a given target location. The ventral pathway, going from the V1 area to the infero-temporal (IT) cortex, is related to the recognition of objects, particularly faces and places (Ison and Quian Quiroga, 2008; Logothetis and Sheinberg, 1996; Mishkin et al., 1983; Tanaka, 1996).

Converging evidence from monkey recordings supports the notion that visual awareness is linked to the neural activity of the ventral visual stream (Logothetis, 1998; Tanaka, 1997). The preferred stimulus of the neurons increases in complexity and selectivity in successive hierarchical levels of the pathway. In V1, neurons respond to contours or edges with specific orientation. Cells having the same stimulus preferences are grouped into columns that occupy the entire thickness of the cortex. These columns are in turn grouped in functional units that examine contours in all orientations of their receptive fields with both eyes, covering the contra-lateral half of the space (Hubel, 1988). From V1 the information is conveyed to V2 and then to V4, where neurons have larger receptive fields and respond to angles and curves pointing in particular directions (Tanaka, 1996). At the top of the visual hierarchy, the IT cortex has even larger receptive fields that are not retino-topically organized but include the center of the gaze with the strongest response to stimuli presented in the projection of the fovea (Gross, 2008). In the IT cortex most neurons respond to complex stimuli, such as faces and hands, in a way that is often relatively invariant to changes in stimulus size, contrast, color and position within the receptive field (Gross, 2008). In the human brain, the ventral surface extending from the occipito-temporal border to the middle part of the temporal cortex is considered to be the homologue of the monkey IT cortex (Tanaka, 1997).

The IT cortex provides massive projections to the medial temporal lobe (MTL), which also receives input from multiple sensory modalities and is essential for declarative memory (Squire, Stark and Clark, 2004). The MTL includes the hippocampal region (CA fields, dentate gyrus, and subicular complex) and the adjacent perirhinal, entorhinal, and parahippocampal cortices. The main cortical input to the hippocampus is the entorhinal cortex, which receives projections from the peri-rhinal and parahippocampal cortices, which in turn receive projections from unimodal and polymodal areas in the frontal, temporal, and parietal lobes. Thus, the hippocampus lies at the end of a cortical processing hierarchy (Squire et al., 2004).

Neural Responses in MTL

Patients suffering from MTL epilepsy refractory to medication may undergo the implantation of intracranial electrodes to determine the seizure focus and to evaluate the feasibility and potential outcome of a curative resection of the focus (e.g., Engel, 1996). Extra-cellular single unit recordings conducted in this type of patient led to the discovery of MTL neurons that responded selectively to visual stimuli from different categories, such as faces, landmarks and animals (Kreiman, Koch and Fried, 2000). A few years later, neurons that responded selectively to the identity of a given person were found (Quian Quiroga et al., 2005). For example, one neuron in the left hippocampus responded only to a variety of images of the actress Jennifer Aniston, and another neuron in the right hippocampus responded only to photos of Halle Berry and even to her written name. These results suggest that hippocampal neurons are encoding an abstract representation of a given individual known to the patient.

The remarkable degree of selectivity and invariance found in MTL neurons has led to the idea of a sparse and abstract representation of concepts (Ison and Quian Quiroga, 2008; Quian Quiroga, Kreiman, Koch and Fried, 2008a; Quian Quiroga, Reddy, Kreiman, Koch and Fried, 2005; Waydo, Kraskov, Quian Quiroga, Fried and Koch, 2006). These findings are in line with the classic concepts summarized in Barlow's five dogmas of perceptual physiology (Barlow, 1972), and particularly with the fourth dogma, which states that perception corresponds to the activity of a small selection of the very numerous high level neurons: *“Just as physical stimuli directly cause receptors to initiate neural activity, so the active high-level neurons directly and simply cause the elements of our perception”*. (380)

This does not mean that any particular object is represented by only one specific cell, often referred as the “grandmother cell” (e.g., Barlow, 1994), but it seems conceivable that a relatively small set of highly specialized neurons that remain silent most of the time could be tuned to codify specific objects that are highly significant for the species, such as faces and homes in the case of humans, or flying bugs in the case of frogs (Quian Quiroga et al., 2008a). Indeed, the original idea of Jerry Lettvin (see appendix in Barlow, 1994) was that a cluster of “18000” units would encode a concept. Often used to remark on the unlikeliness of this concept, the single grandmother cell has been proposed as the theoretical extreme of a sparse representation. One of the main criticisms to the grandmother cell theory, however, is that there would not be enough cells to encode every possible percept, because the number of stimuli that can be encoded increases linearly with the number of cells. In addition, the death of only one particular cell would imply the loss of the encoded concept. So, although the brain may encode concepts using very sparse representations (Ison and Quian Quiroga, 2008; Quian Quiroga et al., 2008a; Quian Quiroga et al., 2005), it is clear that a one to one mapping is impossible.

Neural Correlates of Visual Perception

A particularly interesting methodology that has been used to study the function of the ventral visual pathway is the presentation of visual stimuli that can be perceived in more than one way, as the case of binocular rivalry (Leopold and Logothetis, 1996; Logothetis, 1998; Logothetis, Leopold and Sheinberg, 1996; Logothetis and Sheinberg, 1996) and ambiguous pictures (Logothetis, 1998). These bi-stable (or multi-stable) stimulation paradigms constitute an invaluable tool for the study of visual perception, because they permit the dissociation of the neural responses that underlie what is perceived from the straight forward sensory representation of the visual pattern.

In the binocular rivalry paradigm, two different images are shown for each eye and the resulting percept does not correspond to a fusion of both but rather to an alternating perception of each of them (Logothetis, 1998). The perceived image changes spontaneously between both figures in a random manner, but a switch can be induced when a new image is presented, because the novel picture perceptually suppresses the other for a certain time. This effect, known as flash suppression (Wolfe, 1984), is extremely useful. It allows one to manipulate experimentally which image of the pair (that creates the bi-stable pattern) will capture the perception in a given moment.

Using binocular rivalry and flash suppression in monkeys trained to report their perception, Logothetis (1998) determined that some cells are activated by the actual visual pattern, while others are affected by the way in which the stimulus is perceived. The proportion of the latter type of cells increased in the higher processing stages of the visual system. For example they increased from 20% in V1, to 90% in IT neurons. Experiments with flash suppression in epileptic patients showed a similar picture in human MTL: two out of three neurons modulated their activity according to the perceived stimulus (Kreiman, Fried and Koch, 2002).

Results obtained with fMRI techniques in humans usually parallel the results of studies using single cell recordings in monkeys (for a comprehensive review see Kanwisher, 2001). For example, regions of the human visual cortex, such as the fusiform face area and the parahippocampal place area, were activated specifically by the perception of faces (Kanwisher, McDermott and Chun, 1997) and images of places (Epstein and Kanwisher, 1998), respectively. Studies with binocular rivalry have shown that the neural activation was correlated with the content of the subject's awareness and not with the retinal input (Kanwisher, 2001).

Using single cell recordings in patients, a different experiment was recently designed to assess if a neuron was activated just by a given image or by its conscious recognition (Quiñones Quiroga et al., 2008b). To this end, the presentation time of the image was controlled through stimulus masks, and photos eliciting responses were shown very

briefly, such that they were sometimes recognized and sometimes not. Interestingly, although the stimulus was exactly the same, the neurons fired only when a photo was recognized by the patient and remained mostly silent when it was not. Furthermore, the duration of the responses did not increase with the stimulus duration. Thus, we can conclude that the firing of these neurons strongly correlates with the conscious recognition of the photos. These results parallel the observations of Grill-Spector using fMRI (Grill-Spector, Kushnir, Hendler, and Malach, 2000), and support the idea that the higher visual areas respond mainly to the image that is actually perceived.

Given clinical and lesion studies about the function of the hippocampus and surrounding areas (e.g., Squire et al., 2004) and the long latency of the MTL responses (300 ms) (Mormann et al., 2008) compared to the latencies reported for picture recognition (150 ms) (Thorpe, Fize and Marlot, 1996), it is likely that MTL neurons are not performing the recognition process itself. The fact that they are mainly modulated by conscious perception is in line with the interpretation that these cells may be underlying the link between consciously perceived inputs and long-term memory (Quian Quiroga et al., 2008a; Quian Quiroga, Mukamel, Isham, Malach and Fried, 2008b). Indeed, to form an abstract representation that could be memorized, neurons must encode the identity of the individual in an abstract manner, and this is achieved through an invariant response. Based on these findings, it seems plausible that MTL neurons play a key role in the transformation of visual percepts into long term abstract memories, and in the recollection of memory traces related to a concept (Ison and Quian Quiroga, 2008; Quian Quiroga et al., 2008a). In line with these ideas, a recent report has shown that neurons in the hippocampus and entorhinal cortex that were specifically activated during the display of short films also responded during the free recall of the same sequences by the patients (Gelbard-Sagiv, Mukamel, Harel, Malach and Fried, 2008).

Causal Relations Between Neural Activity and Awareness

The evidence that neural activity in the ventral stream is correlated with visual perception of faces or places leads to the issue of the underlying causal structure (Kanwisher, 2001). In other words, is there a causal relationship between neuronal firing and perception, or just mere correlation? Evidence for a causal connection was provided by experiments using cortical micro-stimulation in a cluster of face neurons of monkeys trained to report the presence of a face in images degraded by visual noise (Afraz, Kiani and Esteky, 2006). When the cluster of face neurons was stimulated, the monkeys were more likely to report faces, even when only visual noise was presented. This effect was restricted to the area in which face neurons were located and to the time window in which they responded to the face. In line with these findings, previous studies in epileptic patients reported that cortical micro-stimulation in fusiform electrode sites with face selective responses produced perceptual experiences of a face (Puce, Allison and McCarthy, 1999).

These experimental results imply that the activation of brain areas where the perceptual representation occurs can bias perception. The activation of such areas, however, may not always be sufficient or be the cause of conscious perception. As discussed by Kanwisher (2001), subjects could be unaware of the low contrast grid stimulation that evoked clear fMRI responses in V1-V4. Similarly, people did not report having perceived briefly presented angry or happy faces (masked with neutral expression faces) although the angry faces evoked a significantly larger activation of the amygdala (Whalen, Rauch, Etcoff, McInerney, Lee and Jenike, 1998). There also exists considerable evidence in psychological research indicating that stimulus information can affect measures of forced-choice discrimination even when it is not consciously perceived (Merikle, Smilek, and Eastwood, 2001). What else is needed for conscious awareness? Of many proposed suggestions, Baars has offered the simpler idea that certain strength of activation may be required (1998). In this context, non-conscious representations could be explained in terms of sub-threshold activations. The threshold does not need to be a clear cut edge but a range of values that may also depend on attention levels. In this threshold range, the perception might not be conscious but the information contained in the stimulus could still be influential in a forced choice, or could condition perceptions related to another supra-threshold stimulus. In addition to this strength activation hypothesis, it has been suggested that a focal representation may also have to be accessible to other parts of the brain in order to reach awareness (Kanwisher, 2001). Indeed, a local activation pattern can rapidly reach distant areas of the neocortex through long range connections. According to Buzáki (2007), in order to enter in the conscious experience, environmental inputs would need to reach a long enough duration and intensity to extend beyond local circuits and perturb the ongoing rhythmic activity of the brain. In the specific case of visual awareness, connections within the IT cortex and executive parts of the brain, as prefrontal and anterior cingulate cortices, are likely to be involved (Koch, 2004). However, the specific locus that would be needed for awareness remains an open discussion because perceptual representations are present in multiple cortical areas. This issue is related to the so called binding problem outlined in the next section.

The Binding Problem and Neural Oscillations

The problem of how a visual scene composed of different objects is combined to form a unified perceptual experience is known as the binding problem (for example, Roelfsema, 2006).

The idea of binding can be applied to many levels, from grouping parts of the visual field into particular objects, to the combination of these objects into a complete visual scene, and the integration of stimuli coming from different sensory modalities in a unified perceptual experience.

In the visual pathways, different attributes of the image, such as color, orientation, size, distance and movement are mapped in parallel pathways. In a first step of binding,

known as pre-attentive process, selected aspects of these feature maps converge in a bottom-up manner into a coherent map that contains the distinction of objects from the background. After that, the focus of attention can highlight selected parts of this unified map by referring back to the feature maps in a serial and top-down manner (Treisman, 1986).

The detailed neuronal mechanisms underlying the different steps of perceptual binding are an open issue at present. It has been suggested that this binding could be constructed through the correlated firing of the neurons that respond to the properties of the different elements that compose an object. This correlated firing could be achieved through coherent oscillations of neural assemblies in the gamma frequency (35-75 Hz) (Eckhorn et al., 1988; Gray and Singer, 1989). This idea can be extended to the binding of a whole perceptive scene. Then, gamma oscillations may constitute a neural correlate for visual awareness through the activation of a network able to join together relevant aspects of the existing information into a coherent percept (Crick and Koch, 1990). Alternatively, such binding could be just represented by the activity of a cell population, with coherent oscillations playing little or no roll at all (Shadlen and Movshon, 1999). The relevance of gamma oscillations for perceptual binding has been a topic of active research and controversy in neuroscience (for example, Gray, 1999; Shadlen and Movshon, 1999).

Independent of the fact that oscillations or single cell firing are the basis of perception, another question presents itself. Where does this binding actually takes place? In other words, is there a “reader” of this activity? A simple answer is that the neurons of the following step in the visual hierarchy could be performing this reading. This is in line with Barlow’s idea of a relatively small population of “cardinal” neurons, which when activated, would cause the perceptual experience (Barlow, 1972). As an alternative proposal, it has been suggested that perception is not linked to single cells but to field oscillations in an extended mass of tissue. In this conception, coherent gamma oscillations could be organized as spatiotemporal fields that may represent the contents of consciousness (Freeman, 2007). Since the field oscillations are sustained by networks of cells, both hypotheses do not necessarily exclude each other completely. After all, the idea of a cell or process to which consciousness would be attached is not new (for example, James, 1890). Although nowadays we understand many more details about the physiological processes involved in neural function, the basic quest outlined in the prophetic writings of James seems to remain essentially the same. As remarked by Crick (1994), the key difference is that we are now able to address this question experimentally.

References

- Afraz, S-R., Kiani, R., & Esteky, H. (2006). Microstimulation of inferotemporal cortex influences face categorization. *Nature*, 442, 692-695.
- Baars, B. (1988). *A cognitive theory of consciousness*. Cambridge: Cambridge University Press.
- Barlow, H.B. (1972). Single units and sensation: A neuron doctrine for conceptual psychology? *Perception*, 1, 371-394.
- Barlow, H.B. (1994). The neuron doctrine in perception. In M. Gazzaniga (Ed.), *The Cognitive Neurosciences*. (pp. 415-435). Boston: MIT press.
- Buzáki, G. (2007). The structure of consciousness. *Nature*, 446, 267.
- Crick, F. (1994). *The astonishing hypothesis*. NY: Touchstone.
- Crick, F., & Koch, C. (1990). Towards a neurobiological theory of consciousness. *Seminars in the Neurosciences*, 2, 263-275.
- Crick, F., & Koch, C. (1995). Why neuroscience may be able to explain consciousness? *Scientific American*, 273, 84-85.
- Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M., & Reitboeck, H.J. (1988). Coherent oscillations: A mechanism of feature linking in the visual cortex? Multiple electrode and correlation analyses in the cat. *Biological Cybernetics*, 60, 121-130.
- Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local environment. *Nature*, 392, 598-601.
- Engel, J. (1996). Surgery for seizures. *New Eng J Med*, 334, 647-652.
- Freeman, W.J. (2007). 2007 Special Issue: Indirect biological measures of consciousness from field studies of brains as dynamical systems. *Neural Networks*, 20, 1021-1031.
- Gelbard-Sagiv, H., Mukamel, R., Harel, M., Malach, R., & Fried, I. (2008). Internally generated reactivation of single neurons in human hippocampus during free recall. *Science*, 322, 96-101.
- Gray, C.M., & Singer, W. (1989). Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *PNAS USA*, 86, 1698-1702.
- Gray, C.M. (1999). The Temporal Correlation Hypothesis of Visual Feature Integration: Still Alive and Well. *Neuron*, 24, 31-47.
- Grill-Spector, K., Kushnir, T., Hendler, T., & Malach, R. (2000). The dynamics of object-selective activation correlate with recognition performance in humans. *Nature Neuroscience*, 3, 837-893.
- Gross, C.G. (2008). Single neuron studies of inferior temporal cortex. *Neuropsychologia*, 46, 841-852.

- Hubel, D.H. (1988). *Eye, brain, and vision*. New York: Scientific American Books.
- Ison, M.J., & Quian Quiroga, R. (2008). Selectivity and invariance for visual objects perception. *Frontiers in Bioscience*, 4889-4903.
- James, W. (1890). *The Principles of Psychology*. New York: Dover Publications.
- Kanwisher, N., McDermott, J., & Chun, M.M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, 17, 4302-4311.
- Kanwisher, N. (2001). Neural events and perceptual awareness. *Cognition*, 79, 89-113.
- Koch, C. (2004). *The Quest for Consciousness: A Neurobiological Approach*. Englewood: Roberts & Company.
- Kreiman, G., Koch, C., & Fried, I. (2000). Category-specific visual responses of single neurons in the human medial temporal lobe. *Nature Neuroscience*, 3, 946-953.
- Kreiman, G., Fried, I., & Koch, C. (2002). Single-neuron correlates of subjective vision in the human medial temporal lobe. *PNAS*, 99, 8378-8383.
- Leopold, D.A., & Logothetis, N.K. (1996). Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature*, 379, 549-553.
- Logothetis, N.K., Leopold, D.A., & Sheinberg, D.L. (1996). What is rivaling during binocular rivalry? *Nature*, 380, 621-624.
- Logothetis, N.K., & Sheinberg, D.L. (1996). Visual object recognition. *Annual Review of Neuroscience*, 19, 577-621.
- Logothetis, N.K. (1998). Single units and conscious vision. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 353, 1801-1818.
- Merikle, P.M., Smilek, D., & Eastwood, J.D. (2001). Perception without awareness: perspectives from cognitive psychology. *Cognition*, 79, 115-134.
- Mishkin, M., Ungerleider, L.G., & Macko, K.A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences*, 6, 414-417.
- Mormann, F., Kornblith, S., Quian Quiroga, R., Kraskov, A., Cerf, M., Fried, I., & Koch, C. (2008). Latency and selectivity of single neurons indicate hierarchical processing in the human medial temporal lobe. *Journal of Neuroscience*, 28, 8865-8872.
- Puce, A., Allison, T., & McCarthy, G. (1999). Electrophysiological studies of human face perception. III. Effects of top-down processing on face-specific potentials. *Cerebral Cortex*, 9, 445-458.
- Quian Quiroga, R., Reddy, L., Kreiman, G., Koch, C., & Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature*, 435, 1102-1107.
- Quian Quiroga, R., Kreiman, G., Koch, C., & Fried, I. (2008a). Sparse but not "Grandmother-cell" coding in the medial temporal lobe. *Trends in Cognitive Sciences*, 12, 87-91.

- Quian Quiroga, R., Mukamel, R., Isham, E., Malach, R., & Fried, I. (2008b). Human single neuron responses at the threshold of conscious recognition. *PNAS*, 105, 3599-3604.
- Roelfsema, P.R. (2006). Cortical algorithms for perceptual binding. *Annu. Rev. Neurosci.*, 29, 203-227.
- Shadlen, M.N., & Movshon, J.A. (1999). Synchrony unbound: A critical evaluation of the temporal binding hypothesis. *Neuron*, 24, 67-77.
- Squire, L.R., Stark C.E.L., & Clark, R.E. (2004). The medial temporal lobe. *Annual Review of Neuroscience*, 27, 279-306.
- Sutherland, S. (1996). *The International Dictionary of Psychology*. US: Crossroad Publishing Co.
- Tanaka, K. (1996). Representation of visual features of objects in the inferotemporal cortex. *Neural Networks*, 9, 1459-1475.
- Tanaka, K. (1997). Mechanisms of visual object recognition: monkey and human studies. *Current Opinion in Neurobiology*, 7, 523-529.
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, 381, 520-522.
- Treisman, A. (1986). Features and objects in visual processing. *Scientific American*, 255, 114-125.
- Waydo, S., Kraskov, A., Quian Quiroga, R., Fried, I., & Koch, C. (2006). Sparse representation in the human medial temporal lobe. *Journal of Neuroscience*, 26, 10232-10234.
- Whalen, P.J., Rauch, S.L., Etcoff, N.L., McInerney, S.C., Lee, M.B., & Jenike, M.A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *Journal of Neuroscience*, 18, 411-418.
- Wolfe, J.M. (1984). Reversing ocular dominance and suppression in a single flash. *Vision Research*, 24, 471-478.