Seeing faces and objects with the “mind’s eye”

ALUMIT ISHAI

Institute of Neuroradiology, University of Zurich, Switzerland

ABSTRACT

With the advent of functional brain imaging techniques and recent developments in the analysis of cortical connectivity, the focus of mental imagery studies has shifted from a semi-modular approach to a more realistic, integrated, cortical networks perspective. Recent studies of visual imagery of faces and objects suggest that activation of content-specific representations stored in the ventral visual stream is top-down modulated by parietal and frontal regions. The relation of these findings to other cognitive functions is discussed, as well as their clinical implications for patients with impaired states of conscious awareness.

Key words

Connectivity • fMRI • Imagery • Objects • Perception

Introduction

Visual imagery is the ability to generate percept-like images in the absence of retinal input and therefore is a vivid demonstration of retrieving pictorial information from memory. Psychophysical and brain imaging studies have demonstrated functional similarities between visual perception and visual imagery, to the extent that common mechanisms appear to be activated by both (Roland et al., 1987; Farah et al., 1988; Goldenberg et al., 1989; Ishai and Sagi, 1995). Numerous neuroimaging studies have shown that visual imagery, like visual perception, evokes activation in occipito-parietal and occipito-temporal visual association areas (Mellet et al., 1996; D’Esposito et al., 1997). In some studies, the primary visual cortex (Le Bihan et al., 1993; Kosslyn et al., 1993, 1999) and the lateral geniculate nucleus (Chen et al., 1998) were activated during imagery, suggesting that the generation of mental images may involve sensory representations at the earlier processing stages in the visual pathway. Studies of patients with brain damage have demonstrated a dissociation of visual-object and visual-spatial imagery (Levine et al., 1985), indicating that different parts of the visual system mediate “where” and “what” imagery, a dissociation that parallels the two anatomically distinct visual systems proposed for visual perception (Ungerleider and Mishkin, 1982).

Although many studies have focused on the overlap and similarities between perception and imagery, the subjective experience of imagining and seeing is clearly different. The intriguing case study of CK, a patient with severe visual agnosia who cannot recognize objects but can draw them with considerable details from memory, suggests that visual imagery can be dissociated from visual perception (Behrmann et al., 1992). It has been shown that during visual imagery, deactivation in auditory cortex is negatively correlates with activation in visual cortex and with the score of subjective vividness of visual imagery. Thus, in order to generate vivid mental images, the brain needs to filter out irrelevant sensory information (Amedi et al., 2005). Intriguingly,
recent studies in vegetative state patients have discovered normal patterns of activation in their brain during mental imagery (Owen et al., 2006; Boly et al., 2007), suggesting that imagery can be used to assess the degree of consciousness in non-communicative brain-damaged patients, and perhaps predict their recovery (Di et al., 2008). A recent study has shown that the precuneus, which mediates memory-related imagery (Fletcher et al., 1995), is also activated during hypnosis, suggesting that such a state of enhanced self-monitoring is achieved by control of motor responses by internal representations (Cojan et al., 2009). It therefore seems that understanding the neural mechanisms of mental imagery could have far reaching implications for understanding conscious awareness and its various disorders. The main issue addressed in this review is what happens in the brain when we see faces and objects with our “mind’s eye”? Converging empirical evidence suggests that content-specific activation in the ventral stream is mediated by top-down activation in parietal and frontal regions.

Faces, objects and the imagery network

Functional MRI studies have reported that within the ventral pathway faces and other objects, such as outdoor scenes, houses, chairs, animals, and tools, have distinct representations (Kanwisher et al., 1997; Epstein and Kanwisher, 1998; Aguirre et al., 1998; Chao et al., 1999). In particular, it has been shown that faces, houses and chairs evoke maximal responses in distinct occipital and ventral temporal regions with a highly consistent topological arrangement across subjects (Ishai et al., 1999, 2000a). As each category was associated with its own differential pattern of responses across a broad expanse of cortex, it has been proposed that the representation of objects in the ventral stream is not restricted to discrete, highly selective patches of cortex, but, rather, is comprised of distributed representation of information about object form (Ishai et al., 1999, 2000a; Haxby et al., 2001). Interestingly, similar activation in occipito-temporal cortex is also observed during visual and tactile recognition of faces and man-made objects in sighted subjects, and during tactile recognition in blind subjects (Pietrini et al., 2004), supporting the idea that more abstract features of object form are represented in the visual ventral stream.

Inspired by the consistent topology of the response to faces, houses and chairs, we investigated whether visual imagery of these objects would evoke content-related activation within the same extrastriate ventral regions that are activated during perception. We further asked which brain regions, activated during imagery, provide the top-down signal to visual extrastriate cortex. To that end, an fMRI study with the following experimental conditions was designed: Perception, in which subjects viewed pictures of faces, houses, chairs; Perception-control, in which subjects viewed scrambled pictures; Imagery, in which subjects were instructed to generate vivid mental images of familiar houses, faces, and chairs from long-term memory (LTM) while viewing a gray square, and to press a button when ready with a vivid image; and Imagery-control, in which subjects were asked to press the button from time to time while viewing the gray square. Conventional Statistical Parametric Mapping (SPM) analysis revealed differential activation in medial fusiform, lateral fusiform and inferior temporal gyri during the perception of houses, faces, and chairs, respectively. During visual imagery, content-related patterns of activation were found, but this activity was restricted to small sectors of the regions that responded differentially during perception. Thus, the generation of mental images of familiar faces from LTM evoked activation within small subsets of the lateral fusiform gyrus, a face-responsive region, whereas generating mental images of houses and chairs evoked activation within subsets of the medial fusiform and inferior temporal gyri, respectively. Visual perception and visual imagery evoked activity with opposite patterns of hemispheric asymmetry in ventral temporal cortex, with stronger responses in the right hemisphere during perception, and stronger responses in the left hemisphere during imagery. Such hemispheric asymmetry during visual imagery was reported by Farah and colleagues, who found that the event-related potentials associated with the generation of mental images were greater on the left posterior temporal site than on the right (Farah et al., 1989); and by D’Esposito and colleagues who found, using fMRI, that when subjects generated mental images of the referents of concrete nouns,
the left inferior temporal lobe (Brodmann’s area 37) was the most reliably and robustly area activated (D’Esposito et al., 1997). A special role for the left hemisphere in visual imagery is also supported by numerous case studies of patients with image generation impairment due to damage in left temporoparietal areas (Farah, 1995). Other studies, however, have concluded that both cerebral hemispheres contribute equally to the generation of mental images (Sergent, 1989), or that both hemispheres can generate mental images, but in different ways (Kosslyn et al., 1995b). It therefore seems that both hemispheres are activated during visual imagery, albeit asymmetrically. Contrasting responses evoked by visual imagery and responses evoked during the imagery control condition revealed activation within a network of parietal and frontal regions. This imagery network included the precuneus, intraparietal sulcus (IPS) and inferior frontal gyrus (IFG), regions that were implicated in various attention and retrieval from episodic memory tasks (e.g., Fletcher et al., 1995; Buckner et al., 1996; Mellet et al., 1998). We interpreted these findings to suggest that retrieval of content-specific memory traces, stored in the ventral pathway, is top-down controlled by a parieto-frontal network that mediates the generation and maintenance of mental images (Ishai et al., 2000b).

Bottom-up, top-down and effective connectivity

To investigate the neuronal interactions and effective connectivity that mediate the category-specific responses in the visual cortex, we used Dynamic Causal Modeling (DCM), a relatively new analytic approach that allows the assessment of effective connectivity within cortical networks (Friston et al., 2003). The aim of DCM is to estimate and make inferences about the coupling among brain areas, and how that coupling is influenced by changes in experimental context (e.g., stimuli or tasks). Consistent with a previous DCM study, which showed that category effects in occipito-temporal cortex were mediated by forward connections from early visual areas (Mechelli et al., 2003), we hypothesized that the category-specific patterns of activation observed in occipito-temporal cortex during visual perception of faces and objects could be explained by a selective enabling of forward connectivity from early visual areas. We also predicted that content-related activation observed during visual imagery would be associated with category-dependent changes in backward connectivity from parietal and frontal areas. Finally, we examined whether the DCM analysis would reveal different patterns of effective connectivity during imagery of faces, houses and chairs in parietal and frontal cortices, as our original SPM analysis did not show category-specific imagery activation within these regions (Ishai et al., 2000b). We found that during visual perception, when subjects viewed gray-scale pictures of faces, houses, and chairs, the category-selective effects in occipito-temporal cortex were mediated by forward connections from early visual areas. In contrast, during visual imagery, when subjects generated mental images of faces, houses, and chairs from LTM, the category-selective effects in the visual cortex were mediated by backward connections from prefrontal cortex. Interestingly, the backward connections from prefrontal, but not from parietal cortex, to occipito-temporal cortex were category-selective (Fig. 1). Thus, the DCM analysis revealed that dynamic neuronal interactions between occipito-temporal, parietal and frontal regions are task- and stimulus-dependent. Sensory representations of faces and objects in ventral extrastriate cortex are mediated by bottom-up mechanisms arising in early visual areas during perception, and top-down mechanisms originating in prefrontal cortex during imagery. Additionally, non-selective, top-down processes, originating in superior parietal areas, contribute to the generation of mental images and their maintenance in the “mind’s eye” (Mechelli et al., 2004).

Our findings propose a new perspective on the neural basis of visual imagery. We found evidence for content-related activation in ventral temporal cortex: small sectors of extrastriate regions that participate in visual perception of faces and objects are also involved in representing perceptual information retrieved from LTM during visual imagery of faces and objects (Ishai et al., 2000b, 2002; Mechelli et al., 2004). As imagery evoked activity in small portions of the regions that participate in perception, it is possible that stored information evoked by imagery is simply weaker than equivalent representations evoked by actual visual input. Alternatively, only a
specific subset of cortical regions may be dedicated to mental imagery, allowing perception and imagery to operate simultaneously. These results suggest that sensory representations of faces and objects stored in ventral temporal cortex are reactivated during the generation of visual images, consistent with electrophysiological and other fMRI studies, which have reported content-specific activation in sensory areas during imagery and memory retrieval (Buckner and Wheeler, 2001; Kreiman et al., 2000; Wheeler et al., 2000). Interestingly, electric stimulation of regions in the temporal lobe of humans results in imagery recall, suggesting that memory traces are localized in these regions (Penfield and Perot, 1963). Similarly, studies in non-human primates indicate that the temporal lobe is the memory storehouse for visual representations of complex stimuli (Miyashita and Chang, 1988; Miyashita, 1988).

**Short- and long-term memory**

Behavioral studies have reported differential effects of visual imagery on the performance of a perceptual task: visual recall from short-term memory (STM) facilitated task performance, whereas visual recall from LTM interfered with performance, suggesting that imagery-induced facilitation and interference are memory-dependent (Ishai and Sagi, 1995, 1997a, 1997b). Moreover, the type of memory
(short- or long-term) required for the generation of mental images seems to determine the extent to which the primary visual cortex (V1) is activated during visual imagery. In the vast majority of studies reporting activation during visual imagery in V1, the imagery tasks were based on recall from STM (e.g., Le Bihan et al., 1993; Kosslyn et al., 1993; Chen et al., 1998; Kosslyn et al., 1999; O’Craven and Kanwisher, 2000). To investigate whether similar or different cortical regions are activated during visual imagery generated from STM and from LTM, a special class of stimuli was used, namely famous faces (Ishai et al., 2002). We chose faces of contemporary Hollywood celebrities, assuming that as a result of exposure to these faces in everyday life, subjects would have pictorial representations in memory that could be retrieved by way of visual imagery. Moreover, with the same visual cue, a famous name, an image could be generated from either LTM or STM. For example, one could imagine Marilyn Monroe without seeing her picture before the imagery task (LTM), or one could memorize a specific picture of Marilyn Monroe and shortly after generate a mental image of that picture (STM). We also tested the effect of focal attention during visual imagery. Numerous fMRI studies have shown that selective attention to particular attributes of visual stimuli, such as color or motion, enhanced the activity in regions of extrastriate cortex that process these attributes (e.g., Corbetta et al., 1990). We therefore hypothesized that focusing attention on features of a mental image, rather than on the global configuration of that image, might also result in increased activation. We also speculated that focal attention to mental images would evoke activation in primary visual cortex, based on a model suggested by Sakai and Miyashita (1994), according to which visual imagery is implemented by the interactions between memory retrieval of representations stored in higher visual association areas, and the effect of focal attention on early visual areas. To test all these predictions, the experimental design included the following imagery conditions: Imagery from STM, in which subjects were presented with names of famous faces they had seen and memorized shortly before, and were instructed to generate vivid images of the exact same faces; Imagery from LTM, in which subjects were presented with names of famous faces they had not seen during the experiment, and were instructed to generate any vivid images of these faces; Imagery from STM + Attention, in which subjects were presented with names of famous faces they had seen and memorized shortly before, and were instructed to generate vivid images of these faces and then to answer questions about some facial feature (e.g., “thick lips?”); and Imagery from LTM + Attention, in which subjects were presented with names of famous faces they had not seen during the experiment, and were instructed to generate vivid images of these faces and then answer questions about some facial feature (e.g., “big nose?”). We found that visual perception of famous faces activated the inferior occipital gyrus, lateral fusiform gyrus, the superior temporal sulcus (STS), and the amygdala, regions of the distributed network that mediates face perception (Haxby et al., 2000; Ishai et al., 2005; Fairhall and Ishai, 2007), whereas visual imagery of famous faces activated small subsets of these face-responsive regions. Visual imagery of famous faces activated a network of regions that included the calcarine, precuneus, hippocampus, IPS and IFG. Within these regions, imagery generated from STM evoked more activation than imagery generated from LTM (Ishai et al., 2002). Furthermore, during imagery generated from both STM and LTM, focusing attention on features of the imagined faces resulted in increased activation in the right IPS and right IFG, consistent with reports on activation in these regions during sustained attention (e.g., Pardo et al., 1990).

The parietal and frontal regions activated during visual imagery of faces and objects also mediate other cognitive functions, such as the retrieval of object representations from LTM, their maintenance in a working memory ‘buffer’, and the attention required to generate those mental images. A similar network of prefrontal areas was activated during motion imagery (Goebel et al., 1998). On the basis of our effective connectivity analysis, we suggest that this imagery network is composed of a general attentional mechanism arising in parietal cortex, and a content-sensitive mechanism originated in prefrontal cortex. Numerous studies of spatial and non-spatial attention tasks have shown activation in parietal cortex (e.g., Corbetta et al., 1998; Kastner et al., 1999; Wojciulik and Kanwisher, 1999). Moreover, parietal activation has been reported in a variety of mental imagery (Mellet et al., 2000; Ishai et al., 2002), and
recognition memory (Wagner et al., 2005; Yago and Ishai, 2006; Wiesmann and Ishai, 2008) tasks. It is therefore reasonable to assume that regions in parietal cortex mediate the attentional processes required to perform the imagery task, irrespective of stimulus-content. Electrophysiological and lesion studies have shown that the prefrontal cortex is not only crucial for object recognition (Bechevalier and Mishkin, 1986), but importantly, that category-selective responses exist in the monkey prefrontal cortex (Freedman et al., 2001). Models of visual working memory (Fuster and Bauer, 1974; Miller et al., 1996) posit that visual working memory is mediated by neuronal interactions between prefrontal and occipito-temporal cortices. Evidence for the existence of similar mechanism in the human brain comes from working memory studies showing that the retrieval of visual information is mediated by a top-down flow of information from prefrontal cortex to category-selective regions in the ventral stream (e.g., Druzgal and D’Esposito, 2003). Our visual imagery studies further show that content-specific imagery effects in the ventral stream are mediated by top-down mechanisms arising in prefrontal cortex (Ishai et al., 2000b, 2002; Mechelli et al., 2004).

Summary

Visual imagery of faces and objects is a multi-component cognitive process that requires re-activation of specific representations stored in the visual cortex and their maintenance in the “mind’s eye”. It is therefore not surprising that the imagery network comprise of parietal and frontal regions that mediate attention and memory retrieval. Recent development in deciphering the mental chronometry (Formisano and Goebel, 2003) and functional and effective connectivity during spatial imagery (Sack et al., 2008), suggest hierarchical temporal dynamics within the imagery network. Understanding the neuronal interactions among visual, parietal and prefrontal regions has far reaching clinical implications, as vegetative state patients show healthy patterns of brain activation during mental imagery, a surprising finding that reflects will and intention, the hallmark of conscious awareness. Finally, the use of classifiers for decoding mental states would enable additional insights into the neural mechanisms of mental imagery (Reddy et al., 2010). Hopefully in the near future combining these cutting-edge analytic techniques would enable making accurate predictions about the recovery of non-communicative brain damaged patients.

Acknowledgments

The author is supported by the Swiss National Science Foundation grant 3200B0-105278 and by the Swiss National Center for Competence in Research: Neural Plasticity and Repair.

References


