

Multisensory spatial interactions: a window onto functional integration in the human brain

Emiliano Macaluso¹ and Jon Driver²

¹Neuroimaging Laboratory, Fondazione Santa Lucia, IRCCS, Via Ardeatina, 306-00179 Rome, Italy

²Institute of Cognitive Neuroscience and Department of Psychology, University College London, 17 Queen Square, London WC1N 3AR, UK

Incoming signals from different sensory modalities are initially processed in separate brain regions. But because these different signals can arise from common events or objects in the external world, integration between them can be useful. Such integration is subject to spatial and temporal constraints, presumably because a common source is more likely for information arising from around the same place and time. This review focuses on recent neuroimaging data concerning spatial aspects of multisensory integration in the human brain. These findings indicate not only that multisensory integration involves anatomical convergence from sensory-specific ('unimodal') cortices into multisensory ('heteromodal') brain areas, but also that multisensory spatial interactions can affect even so-called 'unimodal' brain regions. Such findings call for a revision of traditional assumptions about multisensory processing in the brain.

Introduction

A common theme throughout much research in neuroscience and neuropsychology has been 'functional specialization': that distinct cognitive, sensory and motor functions can be localized in distinct areas of the brain [1]. Functional neuroimaging now allows confirmatory and exploratory mapping of many such functions in the human brain. A simple example is identification of sensory-specific areas, responding to stimulation in one but not another sensory modality. Sensory-specific cortices can readily be localized using functional magnetic resonance imaging (fMRI) in humans, to regions in occipital cortex for vision, to regions in and around the superior temporal gyrus for audition, and to regions in post-central cortex for touch [2] (Figure 1a,i).

Psychological studies have shown for many years that stimuli in different sensory modalities can powerfully interact under some circumstances, to determine perception or behaviour. Textbook examples include crossmodal illusions such as the ventriloquist effect (perceived location of a sound can shift towards its apparent visual source, as when watching a movie at the cinema) and the

McGurk effect (whereby a seen lip-movement can change how a concurrent speech-sound is heard) [3]. Although such illusions can arise with anomalous or incongruent multisensory combinations, many further crossmodal effects indicate the impact of multisensory congruence. For instance, visual detection can be enhanced at the location of a sound [4,5], and there are many further examples of crossmodal links in spatial attention [6]. In the real world, signals in different modalities from a common external event or object will often be spatially and temporally aligned, and multisensory integration appears subject to corresponding spatial and temporal constraints [7].

The neural basis of multisensory interactions has been studied using both intra-cranial recordings (mainly in animals) and non-invasive electrophysiological or haemodynamic measures in humans. Pioneering animal studies focused on multisensory interactions that can arise owing to converging feedforward projections from sensory-specific structures to heteromodal areas. Neurons in the latter areas respond to stimuli in more than one modality, enabling multisensory interactions to occur at the single-cell level. Such multisensory neurons have now been discovered using single-cell recording in many cortical and sub-cortical regions. Cortical multisensory regions include numerous areas in parietal cortex (e.g. the ventral intraparietal area, VIP), temporal cortex (e.g. the caudal superior temporal polysensory region, cSTP) and frontal cortex (e.g. ventral premotor cortex, vPM) [8]. Subcortical regions include the superior colliculus (thoroughly investigated as a model case of multisensory interactions) [9], the basal ganglia and the putamen [9]. Such 'functional specialization' (here, localized neural selectivity for stimulation in multiple modalities or in just one modality) might provide one mechanism for 'functional integration' of information from different modalities (here, arising in multisensory areas, via feedforward convergence from unimodal areas). Indeed, this has been the standard assumption in research on multisensory integration for many years. But recent fMRI evidence from the human brain suggests that although feedforward convergence from unimodal to heteromodal regions is part of the story, it is by no means the whole story.

Corresponding author: Macaluso, E. (e.macaluso@hsantalucia.it).

Available online 1 April 2005

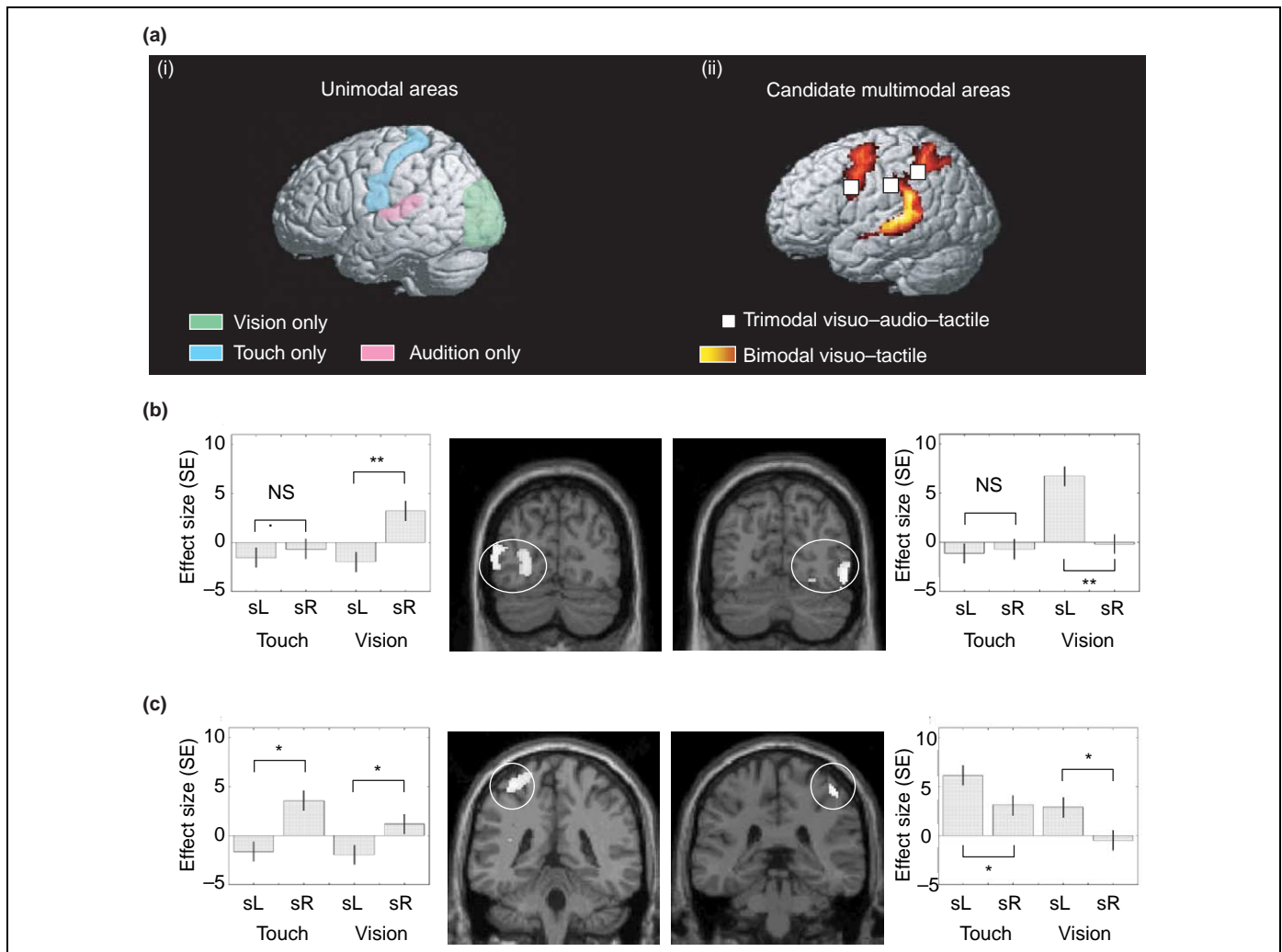


Figure 1. Unimodal and multimodal brain areas. **(a)** (i) Schematic representation of locations of unimodal, sensory-specific cortical regions for vision, touch and audition. (ii) Activation of frontal, parietal and posterior temporal regions in common for visual and tactile stimuli (adapted, with permission, from Ref. [10]), plus peak-locations of trimodal activations (in common for vision, touch and audition; white squares), as reported by Bremmer and colleagues [2]. **(b,c)** Spatially-specific unimodal (visual) and bimodal (visual–tactile) activations (adapted, with permission, from Ref. [10]). Subjects passively received tactile stimuli on either the left or the right hand, or brief flashes of light in close spatial correspondence to either hand. Left and right sides of the panels show responses on the left or the right side of the brain, respectively. The experiment tested for brain areas responding differentially depending on the stimulated hemifield (left minus right, and vice versa), and whether any such spatial effects were specific for stimulation of one modality only (unimodal responses) or common for both modalities (multimodal responses). **(b)** Occipital areas showed spatially-specific activation in the hemisphere contralateral to visual stimulation, but no response during tactile stimulation, as expected. **(c)** The anterior part of the intraparietal sulcus showed spatially-specific contralateral responses, independently of the stimulated modality (vision or touch). Abbreviations: NS, $P > 0.1$; SE, standard error; sL, stimulation of the left-side; sR, stimulation of the right side. ** indicates $P < 0.001$; * indicates $P < 0.05$.

Anatomical convergence as one mechanism for multisensory processing in humans

Perhaps the simplest approach for uncovering candidate multisensory areas in the human brain using neuroimaging is to measure brain activity during stimulation of particular modalities, and then determine whether any regions respond to stimulation of more than one modality. Using visual, auditory and tactile stimuli, Bremmer and colleagues [2] (Figure 1a,ii, white squares) identified multisensory responses in this way for the intraparietal sulcus (IPS), inferior parietal lobule (IPL) and vPM. Parietal, temporal and premotor responses similarly generalized across vision and touch [10] (Figure 1a,ii). This accords with single-cell recordings uncovering cortical heteromodal neurons in such areas for other primates [7,11,12].

An important constraint on multisensory processing is stimulus location [6,7] (although see also Box 1). One

study used fMRI to investigate location-dependent activations for touch and vision, with tactile stimulation of the left or right hand, or visual flashes near either hand [10]. As expected, activations for contralateral versus ipsilateral stimulation were found in contralateral occipital cortex only for vision (Figure 1b) and in post-central regions only for touch. More importantly, a region in the IPS showed activation for contralateral stimulation irrespective of modality (i.e. side-specific responses in common for vision or touch; Figure 1c). In addition, regions of frontal, parietal and posterior-temporal cortex also responded multimodally, but irrespective of stimulus position [10] (Figure 1a,ii). Such results can be readily explained by feedforward convergence of information from sensory-specific cortices (e.g. visual occipital and somatosensory post-central) to heteromodal cortical regions (e.g. the IPS), as also observed in non-human primates [8,12].

Box 1. Temporal aspects of multisensory integration in the brain

Owing to length constraints, this review focuses mainly on spatial multisensory interactions, but it should be noted that temporal factors also constrain multisensory integration. In the real world, not only do multisensory objects and events often produce signals in different modalities that are spatially coincident (Box 3), but also related signals in different senses can tend to occur at around the same time. Single-cell recordings in animals and some functional-imaging studies in humans have investigated the neural consequence of temporal synchronization (or desynchronization) of multisensory signals. Multisensory neurons in the superior colliculus typically exhibit more firing if sensory signals from different modalities reach the cell at around the same time [9], within a window of ~500 ms. Using fMRI in humans, Calvert *et al.* [49] examined auditory–visual interactions during speech perception. Subjects listened to a story (auditory condition), viewed lip-movements of the same story being read (visual condition), were presented with a bimodal version of the story (auditory and visual

signals in usual synchrony) or received auditory and visual signals concerning different stories (which were thus asynchronous, although note that phonemic and semantic incongruence would also arise). Activity in the posterior superior temporal sulcus (a region that can respond to both visual and auditory stimuli) increased for bimodal auditory–visual stimulation, but only when the visual and auditory inputs were related (and hence synchronous). This study also found that unimodal sensory areas of visual and auditory cortices (i.e. responding to stimulation in one modality but not the other) were also affected by auditory–visual relationships in this way [50], suggesting that temporal crossmodal factors might also affect apparently ‘unimodal’ cortices, as for the spatial effects discussed in the main text. Finally, Fujisaki *et al.* [51] have shown in elegant behavioural work that perception of auditory–visual synchrony and asynchrony can be adaptable rather than fixed, depending on experience.

Multisensory interactions between concurrent stimuli in different modalities

Work such as that outlined in the preceding sections can provide initial evidence for candidate multisensory areas in humans, as established in other species. But human neuroimaging studies of the type described so far stimulated only one modality at a time. Two further crucial steps are to examine how the brain responds when several modalities are stimulated in combination, and how this might depend on the spatial and temporal relationships between the stimuli [6,7,13].

Behavioural and single-cell data indicate that during multisensory stimulation, the relative locations and timings of stimuli in different modalities can determine how they interact in the brain [6,7]. Pioneering single-cell studies in superior colliculus developed a basic methodology for investigating this [9]. Responses to bimodal (or trimodal) stimulation are compared with those for the same neuron when each unimodal stimulus is presented alone. Typical results show that response to multisensory stimuli depends on their spatial and temporal relationships. For spatially and temporally aligned stimuli, the response to multimodal stimuli can exceed unimodal responses, and in some cases can even exceed the sum of the unimodal responses (i.e. can be super-additive), although this might be less common than once thought [14].

Beyond convergence to multimodal brain regions: multisensory interactions can affect ‘unimodal’ brain regions

In several recent fMRI studies (e.g. Refs [15,16]), the relative location of concurrent visual and tactile stimuli has been varied. On a trial-by-trial basis, subjects saw a brief flash in the left or right hemifield near one hand. Unpredictably, on half the trials this visual stimulus was combined with a synchronous unseen vibration to one hand (left or right hand in separate experiments). Thus, visual and tactile stimuli could be presented in either multisensory spatial congruence or incongruence. The results, since replicated and extended, showed that adding spatially-congruent touch increased occipital activity contralateral to the flash [15] (Figure 2a). This is noteworthy because occipital cortex is traditionally

thought to process only visual stimuli in the normal brain (Figure 1b), but touch was found to have spatially-specific crossmodal influences on the visual occipital responses. More recently, contralateral tactile responses in the parietal operculum have been analogously enhanced by adding a visual stimulus at the same location as a touch (Figure 2b). Data in Figure 2(b) are from an experiment in which subjects passively received bimodal stimuli in vision and touch either at the same or at opposite locations, while maintaining central fixation. Maximal activity in parietal operculum was observed for contralateral tactile stimulation when this was coupled with visual stimulation at the same, contralateral position (Figure 2b, dark bars). The parietal operculum includes secondary somatosensory cortex that responds to tactile stimuli but not to visual stimuli [16].

Such results indicate that multisensory effects do not in fact impact solely on heteromodal regions (e.g. the IPS; Figure 1c). Seemingly unimodal areas (e.g. visual cortex and parietal operculum) can also be affected by crossmodal interactions (in this case, spatial interactions), as in Figure 2. Functional neuroimaging has uncovered this surprising fact relatively quickly. By contrast, most previous single-cell studies of multisensory interactions could not have done so (although see Refs [17,18]), because such studies had understandably directed their electrodes primarily to candidate heteromodal structures. Functional neuroimaging, by contrast, can readily assess the whole brain (albeit at the level of haemodynamic activity relating to many neurons).

At first glance, an influence of multisensory interactions on modality-specific areas (e.g. on visual cortex or somatosensory cortex) might seem paradoxical [19]. But in fact this general pattern could accord well with the behavioural and phenomenal evidence that analogous crossmodal interactions can influence ‘unimodal’ percepts, such as heard locations or heard phonemes (as in the Ventriloquist or McGurk illusions), or even visual sensitivity for threshold detection of a light at the same or different locations as a task-irrelevant sound [4,5]. In the latter spatial situation, it might be functionally useful for the modality-specific ‘expert systems’ of one sense (i.e. detectors for modality-specific features, such as colour) to prioritize processing of signals originating

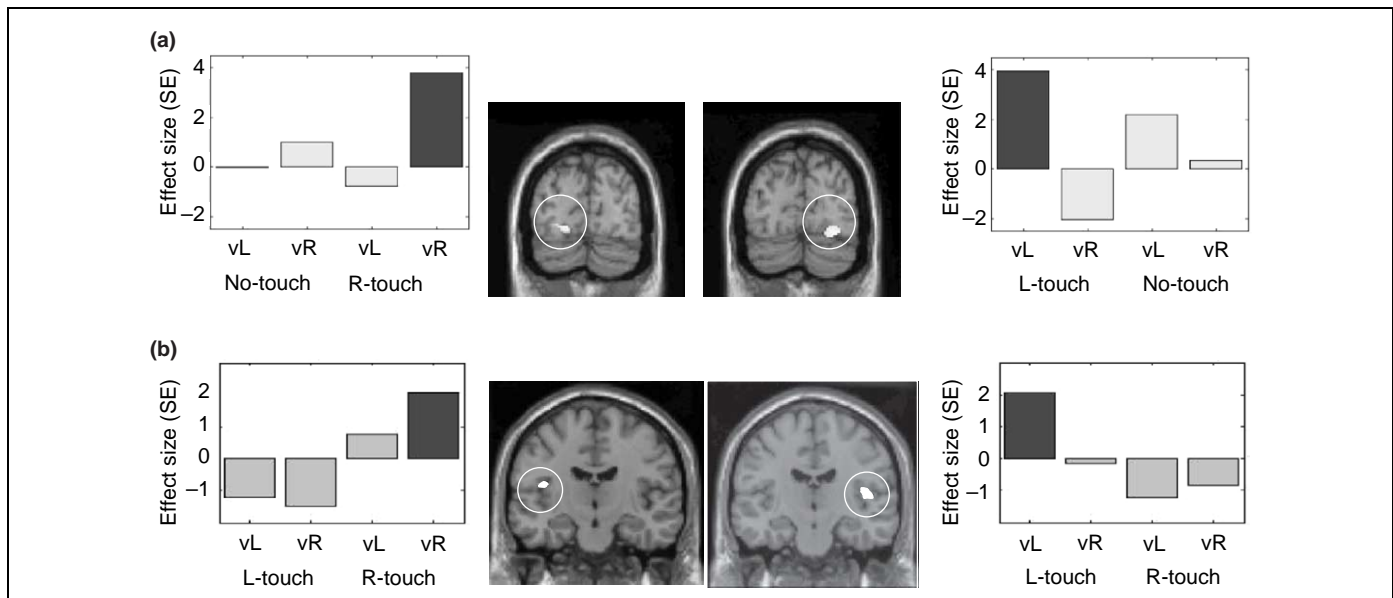


Figure 2. Spatially-specific stimulus-driven crossmodal influences on unimodal cortices. **(a)** Modulation of contralateral visual responses by spatially congruent tactile stimulation (adapted, with permission, from Ref. [10]). Subjects received either left (vL) or right (vR) visual targets. On half of the trials, touch was simultaneously delivered either to the right or to the left hand (R-touch and L-touch, in two different groups of subjects). These task-irrelevant tactile stimulations modulated visual responses in contralateral occipital areas, with higher activity when the visual event was presented together with a spatially congruent tactile stimulus on the same side (dark bars), even though these areas did not respond to touch alone. **(b)** Modulation of contralateral tactile responses by spatially congruent visual stimuli in the parietal operculum (secondary somatosensory cortex; adapted, with permission, from Ref. [47]). Subjects passively received bimodal visual–tactile stimuli, either at the same location or on opposite sides. Activity in the parietal operculum increased for contralateral tactile stimuli when presented with a visual stimulus at the same location (dark bars).

from a location that becomes salient in another sense. For instance, this could be useful when a sound gives away the presence of an animal whose visual properties then need to be urgently extracted. Evidence that crossmodal spatial interactions can affect relatively early ‘unimodal’ levels of visual processing in the human brain has now also been found using methodologies other than fMRI, such as event-related potentials (ERPs) [20] (Box 2).

Putatively crossmodal influences on visual occipital cortex have also been reported for Braille-reading in people with congenital or early-onset blindness [21], for some tactile orientation tasks in sighted subjects [22] and for haptic object recognition [23]. Activation of visual areas during these tactile tasks did not require bimodal tactile plus visual inputs (in contrast to the studies shown in Figure 2) but instead depended on the task and features judged. One possibility is that visual imagery might contribute to those situations [22] and that plasticity might contribute in the case of Braille reading in those with congenital or early-onset blindness. Neither of these specific possibilities seems likely to apply to the findings of specifically spatial crossmodal interactions between combined stimulation in different senses [15,16]. But, as the next section discusses, all of these types of crossmodal influence on visual cortex might in principle involve back-projections from higher-level brain regions that receive inputs from several sensory modalities, and/or that represent task-relevant target properties in a supramodal way.

Possible mechanisms for crossmodal effects on ‘unimodal’ sensory-specific cortices

The findings of crossmodal effects on sensory-specific cortices or sensory-specific ERP components (Figure 2

and Box 2) raise the question of how information concerning one modality could reach brain regions dedicated primarily to processing a different modality. Two contrasting proposals will now be considered, although these need not be mutually exclusive. One involves newly discovered direct connections between sensory-specific areas for different modalities; the other concerns ‘top-down’ influences from multisensory regions on sensory-specific areas, via back-projections.

Using retrograde tracing in monkeys, Falchier *et al.* [24] found direct connections from primary auditory cortex (A1) to regions in primary visual cortex (V1). Moreover, monkey electrophysiology has now demonstrated some responses to somatosensory stimuli in a region closely adjacent to A1 [18]. The spatiotemporal profile of these along the grey-matter laminae indicated that tactile information can enter this region as feedforward input, rather than as modulatory feedback [25]. In humans, evidence for relatively direct influences between unimodal regions comes from ERPs. Giard and Peronnet reported modulation of ERP components arising over visual cortex for bimodal auditory–visual stimuli, as early as 40 ms post-stimulus [26]. More recently, Foxe and colleagues demonstrated apparent modulation of early ERP components for concurrent audition and touch, ~50 ms post-stimulus-onset [27,28]. The groups used analogous analytical approaches, comparing ERPs in response to bimodal stimuli with those calculated as the algebraic sum of the responses to the two unimodal stimuli when presented alone. A possible caveat on this approach is that slow-wave preparatory potentials might affect baseline activity, potentially resulting in artefactual differences in early phases of the ERPs (see Ref. [29] for full discussion of this issue). However, such early crossmodal effects have

Box 2. ERP measurements of crossmodal interactions in the human brain

In addition to fMRI, other neural measurements such as event-related potentials (ERPs) or magnetoencephalography (MEG) also provide evidence on neural multisensory interactions, and they do so with much higher temporal resolution (in the millisecond range). Using ERPs, Kennett *et al.* [33] investigated crossmodal spatial interactions between vision and touch (analogous to those studied using fMRI [15,16]; Figure 2). The results showed that the position of task-irrelevant touch modulated visual ERPs, with larger amplitudes of visual N1 (and Nd) components of ERPs for a light at the same location as a shortly preceding unseen touch (Box 3). These spatially-specific crossmodal effects influenced brain activity as early as ~140 ms after onset of the visual stimulus, for components usually thought to reflect sensory-specific visual processing, again emphasizing crossmodal spatial effects at a relatively early stage of sensory processing.

McDonald and colleagues [31] showed that the position of auditory events can modulate ERP responses to visual stimuli, with enhanced

negativities when auditory and visual stimuli were in the same hemifield ('ipsilaterally-cued') versus opposite sides ('contralaterally-cued') (Figure 1 of this box). Reconstruction of sources at two different time windows (120–140 ms and 150–170 ms) suggested that crossmodal spatial correspondence might have first affected multisensory regions in the superior temporal sulcus, and then unimodal ventral occipital cortex, as if multisensory regions influenced activity in unimodal visual cortex crossmodally, via back-projections.

Finally, Eimer and colleagues have shown in a long series of studies (reviewed in Ref. [20]) that endogenously directing covert spatial attention for a task in one modality (e.g. audition or touch) can modulate sensory responses to incoming stimuli not only within that modality but also (in a spatially-corresponding way) for sensory responses in other modalities (e.g. for visual P1 and N1 components of ERPs), consistent with the neuroimaging results in Figure 3.

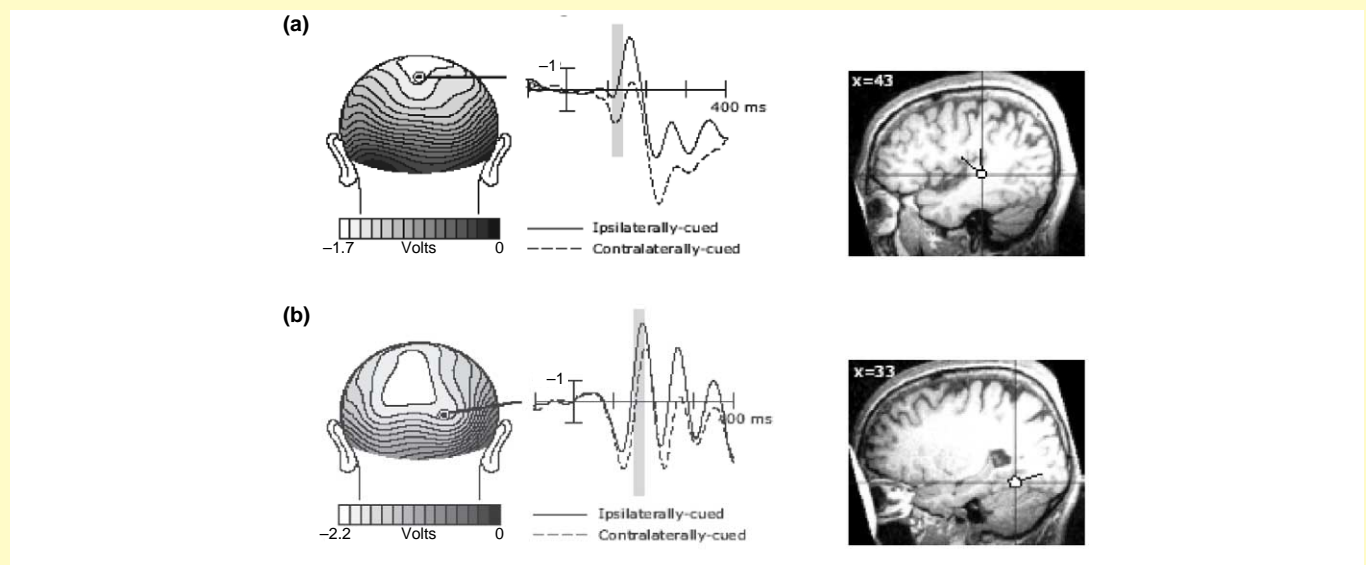


Figure 1. Spatiotemporal dynamics of crossmodal effects, as measured using ERPs. (a) Enhanced negativities for auditory and visual stimuli in the same side ('ipsilaterally-cued') versus the opposite side ('contralaterally-cued'). Scalp distributions (left) and dipolar sources (right) calculated in a 120–140 ms time-window (grey bar superimposed on the ERP traces, centre) after the visual target localized this crossmodal modulation in the superior temporal sulcus. (b) Scalp distributions (left) and dipolar sources (right) calculated in a later 150–170 ms time-window (grey bar) localized this second crossmodal effect in ventral occipital cortex. Adapted, with permission, from Ref. [31].

recently been confirmed [30]. Effects at such short latencies apparently rule out modulatory feedback via longer pathways through multimodal areas in association cortex (see following discussion), suggesting that some direct connections between sensory-specific areas might also exist in humans. But unlike the spatial crossmodal influences on visual ERP components (Box 2), which typically arise after ~100 ms [31,32], the earlier (~50 ms) ERP interactions studied by Foxe and colleagues do not depend on the relative position of bimodal stimuli [30]. Direct connections between sensory-specific areas might thus serve a non-spatial function (possibly related to arousal), and appear unlikely candidates for the spatially-specific crossmodal links now found in many behavioural, fMRI and ERP studies [15,20,31,33] (Figure 2 and Box 2). Moreover, these spatially-specific crossmodal effects can depend on postural factors (e.g. current hand or eye-in-orbit position [16]), 'remapping' across such changes in posture to keep the different senses spatially aligned (Box 3) in a manner that could require cortical control.

Several groups have proposed that these relatively sophisticated crossmodal spatial influences on 'unimodal' sensory cortices could involve back-projections from heteromodal regions, such as those in parietal and frontal cortices [6,7,10,31]. Interestingly, a functional role for modulatory back-projections from higher-order (heteromodal) frontal-parietal regions to posterior sensory areas has also often been suggested in a seemingly very different context – namely, for control of spatial attention – in the purely visual literature [34,35]. In unimodal studies where subjects voluntarily direct covert spatial attention to one location or another, to judge visual targets, three types of fMRI result are commonly observed. First, occipital visual cortex typically shows increased activity in sectors corresponding to (e.g. contralateral to) the attended location [36–38]. In addition, activity in a network of frontal-parietal-temporal regions is also commonly found in association with attentional control [39]. Finally, preparatory activations can also be observed in anticipation of any visual target, consistent with signals

Box 3. Influences of posture on crossmodal spatial effects

Different sensory modalities initially code stimulus locations according to different coordinate systems. For instance, vision is initially retinotopic, whereas touch is somatotopic and audition is head-centred. Because the eyes and other body-parts (e.g. hands, arms and head) can move independently, as they frequently do in daily life, these coordinate systems will be continuously re-aligned as posture changes. This raises important questions for spatially-specific cross-modal effects observed in 'unimodal' brain areas (e.g. as for the effect of spatially-congruent touch on visual responses; Figure 2). If touch to the right hand usually boosts responses to visual stimuli in the right visual hemifield, what happens if the same right-hand tactile stimulation is located in the left visual hemifield? This can arise either when the eye looks towards the right of that right-hand (Figure 1a of

this box) or when the right-arm is placed across the body-midline while the person still looks straight-ahead (Figure 1b of this box). Studies of crossmodal spatial interactions are now addressing this issue [16,33], showing that in many situations the shared location in external space determines crossmodal spatial effects. Hence, although right-hand touch can boost the right visual hemifield when located there, it can boost the left visual hemifield instead if posture relocates the hand there. Such effects imply that one modality communicates spatially with another via intermediate multisensory brain structures than can represent current posture [6]. But some exceptions to this might apply. For instance, some crossmodal effects on early visual ERPs have recently been found not to take change of posture into account [52].

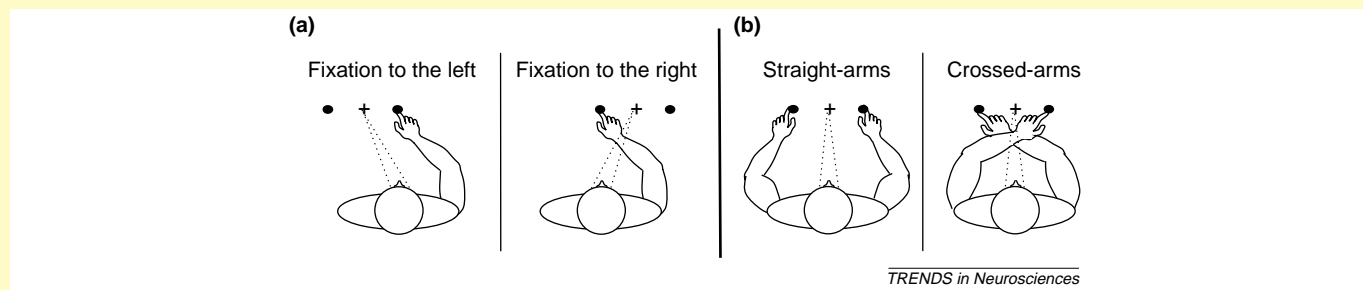


Figure 1. Posture changes. (a) Change of gaze direction. (b) Change of arm posture.

from frontal–parietal regions producing top-down spatial 'biases' (or 'baseline-shifts') in occipital cortex [39,40].

Analogous effects of spatial attention have now been found in situations involving multiple sensory modalities [41]. Subjects again endogenously direct covert spatial attention to perform a judgement at one or other location, in one modality only, but stimuli might also be presented in another task-irrelevant modality, at currently relevant or irrelevant locations (see Ref. [20] for ERP analogues). In such situations, modulation of contralateral visual cortex can be observed not only when subjects attend to one hemifield for a visual task but also when they attend to that side for a purely tactile task [14] (Figure 3). Analogous crossmodal effects of endogenous tactile spatial attention upon vision can also be found for P1 and N1 components of visual ERPs [20,33]. Moreover, in fMRI

studies, crossmodal effects of endogenous spatial attention are preceded by preparatory activations that include frontal–parietal regions (Figure 4a,b) within the so-called 'superior' attention network [35] (see Ref. [42] for cross-modal aspects of more inferior attentional-control regions, in the temporal–parietal junction and inferior frontal cortex). All these putative attention-control areas can be activated equivalently regardless of the particular sensory modality that is the target for spatial attention. The apparently 'supramodal' nature of attentional-control activations in such structures seems consistent with these being heteromodal brain regions, as also suggested from Figure 1(a,ii). Moreover, Eimer *et al.* [43] have observed preparatory ERP components over frontal and parietal sites that likewise generalize across the modality being spatially prepared for. Finally, using fMRI [44],

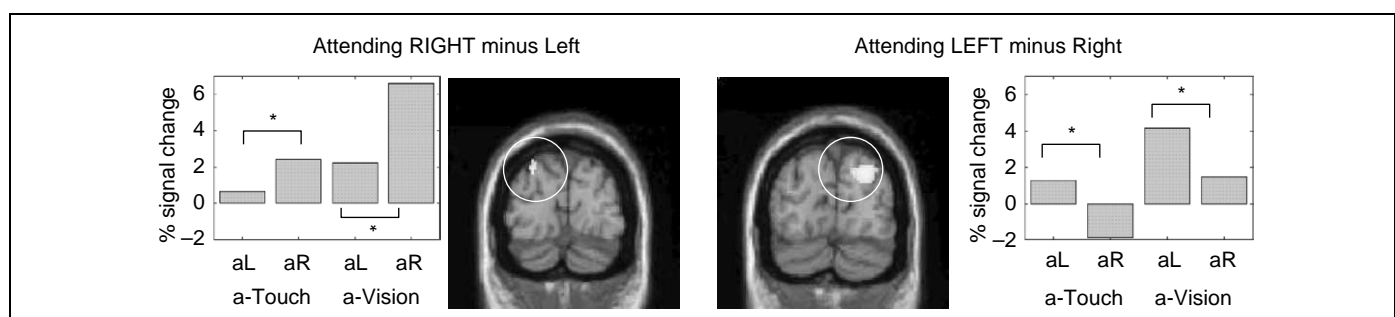


Figure 3. Spatially-specific endogenous crossmodal influences on unimodal visual cortex. Crossmodal attentional effects in visual areas during endogenous attention to one or other side for either a purely visual or a purely tactile task on that side (adapted, with permission, from Ref. [10]). Subjects were presented with concurrent bimodal and bilateral stimulation. During each block they attended covertly to one side (aL or aR) for a task in only one modality (a-Vision or a-Touch). The occipital cortex showed increased activity during attention to the contralateral hemifield. Crucially, these spatially-specific attentional modulations in visual cortex were observed not only when vision was task-relevant (rightmost two bars in each plot) but also when only touch was attended as task relevant (compare the two leftmost bars within each plot), even though this visual area was more active overall when attending vision. * indicates $P < 0.05$.

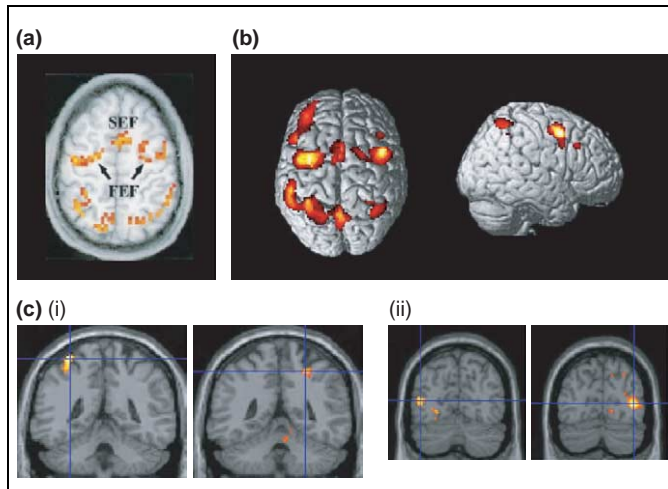


Figure 4. Attentional preparation and control. fMRI evidence indicates that high-level regions in frontal–parietal cortex are involved in attentional control, showing increased activation following attention-directing cues, before any target stimulus is presented. (a) Preparatory activation of frontal–parietal regions in a purely visual study, when subjects diverted covert spatial attention to a particular visual quadrant, preparing for a visual discrimination there (adapted, with permission, from Ref. [48]). Abbreviations: FEF, frontal eye-fields; SEF, supplementary eye-fields. (b) A similar pattern of activation was observed when subjects shifted attention to one or the other hemifield, preparing to discriminate either visual or tactile targets there (adapted, with permission, from Ref. [44] © 2003 Springer-Verlag). (c) Spatially-specific preparatory activations were also observed, both in multimodal intraparietal areas (i) and in unimodal occipital cortices (ii). These preparatory activations were observed in the hemisphere contralateral to the currently attended hemifield, but irrespective of the attended modality. These findings suggest that the allocation of spatial attention occurs supramodally, and that this might lead to spatially-specific crossmodal influences upon unimodal occipital visual areas, possibly via back-projections.

preparatory attention to one or the other side of space has been found to produce spatially-specific ‘baseline shifts’ (cf. Ref. [39]), not only in contralateral heteromodal areas such as IPS (Figure 4c,i; cf. Figure 1c) but also in contralateral visual cortex (Figure 4c,ii).

Given all these results, we suggest that attention is directed spatially at the level of supramodal control structures. This then leads to spatially-specific back-

projection influences on each of the modalities projecting to those control structures, hence leading in turn to the crossmodal influences of spatial attention on unimodal brain structures and components (Figure 2, Figure 3 and Box 2). Such top-down, back-projection modulatory influences are likely to be important for crossmodal effects involving spatial attention. It remains to be determined whether there are any modality-specific levels to such control of spatial attention [45,46] (Box 4), in addition to the supramodal aspects discussed here.

Concluding remarks

Although signals in different sensory modalities are initially processed in different brain regions (functional segregation), these signals can also be combined and can interact crossmodally (providing examples of functional integration). Here, we have reviewed recent neuroimaging findings on spatial crossmodal interactions in the human brain. These indicate that feedforward convergence from lower-level, sensory-specific areas to higher-order, heteromodal areas might not be the only mechanism for such functional integration, although it is one part of the story. The recent evidence indicates that stimulation (and the distribution of attention in space) for one modality can also have spatial influences on brain areas responsible for processing signals in a different modality. We suggest that this could reflect back-projection influences from heteromodal brain regions to multiple unimodal brain regions, at least for the spatial effects reviewed here. In this way, remote regions of the brain can come to have a common spatial focus.

Acknowledgements

Our thanks to Chris Frith and Martin Eimer for the past and ongoing collaboration on the work reported here. E.M. is supported by the Italian Ministry of Health and Telecom Italia Mobile. J.D. is supported by the Wellcome Trust, and was previously supported by the Medical Research Council (UK).

Box 4. Outstanding issues and future directions

(i) We have proposed that modulatory back-projections from multisensory to unimodal brain areas could mediate crossmodal spatial effects. But this still needs direct confirmation. It is broadly consistent with the effects of posture (Box 3) and of top-down attention (Figure 4), with some analyses of ‘effective connectivity’ in fMRI [15], and with the timing of certain ERP dipole-sources (Box 2). But a more direct causal demonstration is still required. Future studies could address this by lesioning or transiently disrupting (e.g. using transcranial magnetic stimulation, TMS) the putative heteromodal sources for particular effects on specific unimodal cortices, while measuring activity in the latter, to test for ‘knockout’ of specific crossmodal effects on unimodal cortices (see Refs [45,46] for an analogous approach combining TMS with behavioural measures of crossmodal interactions).

(ii) Given that crossmodal interactions can evidently affect unimodal brain regions, how are the different senses kept apart sufficiently to prevent ‘synaesthesia-like’ phenomena in the normal brain? As stressed in the main text, we think of the crossmodal influences reviewed here as influencing modality-specific systems that specialize in extracting properties for one sense, but that can show spatial prioritization caused by another. For example, touch spatially modulates the response of visual cortex to visual stimuli in the

normal brain, rather than directly producing visual hallucinations; however, the back-projection influences we have proposed might in principle start to drive the visual cortex more directly when its normal afferent input is lost, as in the blind.

(iii) Do some crossmodal spatial effects reflect feedback to sensory systems from covert spatial motor plans? Motor plans towards a given location would of course transcend any particular sensory modality at that location, and hence might in principle produce crossmodal effects.

(iv) Here, we have emphasized the cortical effects observed; but what role might subcortical structures have in generating these? The current ERP data on spatial multisensory interactions argue against thalamic gating of the first cortical responses; but thalamic–cortical interactions might unfold over time, as might interactions with other subcortical structures.

(v) Can electroencephalogram (EEG) and MEG measurements reveal any synchronization (in specific frequency bands) between remote sensory cortices for different senses, in cases where multisensory interactions between those senses arise?

(vi) Do the same principles of multisensory integration apply with and without awareness?

References

- 1 Gazzaniga, M.S. (2004) *The New Cognitive Neurosciences* (2nd edn), MIT press
- 2 Bremmer, F. *et al.* (2001) Polymodal motion processing in posterior parietal and premotor cortex: a human fMRI study strongly implies equivalencies between humans and monkeys. *Neuron* 29, 287–296
- 3 Bertelson, P. and de Gelder, B. (2004) The psychology of multimodal perception. In *Crossmodal Space and Crossmodal Attention* (Spence, C. and Driver, J., eds), Oxford University Press, pp. 141–178
- 4 McDonald, J.J. *et al.* (2000) Involuntary orienting to sound improves visual perception. *Nature* 407, 906–908
- 5 Frassinetti, F. *et al.* (2002) Enhancement of visual perception by crossmodal visuo–auditory interaction. *Exp. Brain Res.* 147, 332–343
- 6 Spence, C. and Driver, J. (2004) *Crossmodal Space and Crossmodal Attention*, Oxford University Press
- 7 Calvert, G.A. *et al.* (2004) *The Handbook of Multisensory Processing*, MIT press
- 8 Graziano, M.S. and Gross, C.G. (1995) The representation of extrapersonal space: a possible role for bimodal, visuo–tactile neurons. In *The Cognitive Neurosciences* (Gazzaniga, M.S., ed.), MIT press, pp. 1021–1034
- 9 Stein, B.E. and Meredith, M.A. (1993) *The Merging of the Senses*, MIT Press
- 10 Macaluso, E. and Driver, J. (2001) Spatial attention and crossmodal interactions between vision and touch. *Neuropsychologia* 39, 1304–1316
- 11 Graziano, M.S. *et al.* (2002) The cortical control of movement revisited. *Neuron* 36, 349–362
- 12 Graziano, M.S. *et al.* (2004) A system of multimodal areas in the primate brain. In *Crossmodal Space and Crossmodal Attention* (Spence, C. and Driver, J., eds), Oxford University Press, pp. 51–68
- 13 Ehrsson, H.H. *et al.* (2004) That's my hand! Activity in premotor cortex reflects feeling of ownership of a limb. *Science* 305, 875–877
- 14 Stein, B.E. *et al.* (2004) Crossmodal spatial interactions in subcortical and cortical circuits. In *Crossmodal Space and Crossmodal Attention* (Spence, C. and Driver, J., eds), Oxford University Press
- 15 Macaluso, E. *et al.* (2000) Modulation of human visual cortex by crossmodal spatial attention. *Science* 289, 1206–1208
- 16 Macaluso, E. *et al.* (2002) Crossmodal spatial influences of touch on extrastriate visual areas take current gaze-direction into account. *Neuron* 34, 647–658
- 17 Zhou, Y.D. and Fuster, J.M. (2000) Visuo–tactile cross-modal associations in cortical somatosensory cells. *Proc. Natl. Acad. Sci. U. S. A.* 97, 9777–9782
- 18 Fu, K.M. *et al.* (2003) Auditory cortical neurons respond to somatosensory stimulation. *J. Neurosci.* 23, 7510–7515
- 19 Calvert, G.A. *et al.* (1997) Activation of auditory cortex during silent lipreading. *Science* 276, 593–596
- 20 Eimer, M. and Driver, J. (2001) Crossmodal links in endogenous and exogenous spatial attention: evidence from event-related brain potential studies. *Neurosci. Biobehav. Rev.* 25, 497–511
- 21 Burton, H. (2003) Visual cortex activity in early and late blind people. *J. Neurosci.* 23, 4005–4011
- 22 Sathian, K. and Zangaladze, A. (2002) Feeling with the mind's eye: contribution of visual cortex to tactile perception. *Behav. Brain Res.* 135, 127–132
- 23 Amedi, A. *et al.* (2001) Visuo–haptic object-related activation in the ventral visual pathway. *Nat. Neurosci.* 4, 324–330
- 24 Falchier, A. *et al.* (2002) Anatomical evidence of multimodal integration in primate striate cortex. *J. Neurosci.* 22, 5749–5759
- 25 Schroeder, C.E. *et al.* (2001) Somatosensory input to auditory association cortex in the macaque monkey. *J. Neurophysiol.* 85, 1322–1327
- 26 Giard, M.H. and Peronnet, F. (1999) Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *J. Cogn. Neurosci.* 11, 473–490
- 27 Foxe, J.J. *et al.* (2000) Multisensory auditory-somatosensory interactions in early cortical processing revealed by high-density electrical mapping. *Brain Res. Cogn. Brain Res.* 10, 77–83
- 28 Molholm, S. *et al.* (2002) Multisensory auditory–visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Brain Res. Cogn. Brain Res.* 14, 115–128
- 29 Teder-Salejarvi, W.A. *et al.* (2002) An analysis of audio-visual cross-modal integration by means of event-related potential (ERP) recordings. *Brain Res. Cogn. Brain Res.* 14, 106–114
- 30 Murray, M.M. *et al.* (2004) Grabbing your ear: rapid auditory–somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. *Cereb. Cortex* DOI:10.1093/cercor/bhh197 (<http://cercor.oupjournals.org/>)
- 31 McDonald, J.J. *et al.* (2003) Neural substrates of perceptual enhancement by cross-modal spatial attention. *J. Cogn. Neurosci.* 15, 10–19
- 32 Eimer, M. (2004) Electrophysiology of human crossmodal spatial attention. In *Crossmodal Space and Crossmodal Attention* (Spence, C. and Driver, J., eds), Oxford University Press
- 33 Kennett, S. *et al.* (2001) Tactile-visual links in exogenous spatial attention under different postures: convergent evidence from psychophysics and ERPs. *J. Cogn. Neurosci.* 13, 462–478
- 34 Mesulam, M.M. (1998) From sensation to cognition. *Brain* 121, 1013–1052
- 35 Corbetta, M. and Shulman, G.L. (2002) Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215
- 36 Heinze, H.J. *et al.* (1994) Combined spatial and temporal imaging of brain activity during visual selective attention in humans. *Nature* 372, 543–546
- 37 Martinez, A. *et al.* (1999) Involvement of striate and extrastriate visual cortical areas in spatial attention. *Nat. Neurosci.* 2, 364–369
- 38 Noesselt, T. *et al.* (2002) Delayed striate cortical activation during spatial attention. *Neuron* 35, 575–587
- 39 Kastner, S. and Ungerleider, L.G. (2001) The neural basis of biased competition in human visual cortex. *Neuropsychologia* 39, 1263–1276
- 40 Hopfinger, J.B. *et al.* (2000) The neural mechanisms of top-down attentional control. *Nat. Neurosci.* 3, 284–291
- 41 Macaluso, E. *et al.* (2002b) Directing attention to locations and to sensory modalities: multiple levels of selective processing revealed with PET. *Cereb. Cortex* 12, 357–368
- 42 Macaluso, E. *et al.* (2002c) Supramodal effects of covert spatial orienting triggered by visual or tactile events. *J. Cogn. Neurosci.* 14, 389–401
- 43 Eimer, M. *et al.* (2002) Cross-modal interactions between audition, touch, and vision in endogenous spatial attention: ERP evidence on preparatory states and sensory modulations. *J. Cogn. Neurosci.* 14, 254–271
- 44 Macaluso, E. *et al.* (2003) Preparatory states in crossmodal spatial attention: spatial specificity and possible control mechanisms. *Exp. Brain Res.* 149, 62–74
- 45 Chambers, C.D. *et al.* (2004) Modality-specific control of strategic spatial attention in parietal cortex. *Neuron* 44, 925–930
- 46 Chambers, C.D. *et al.* (2004) Fast and slow parietal pathways mediate spatial attention. *Nat. Neurosci.* 7, 217–218
- 47 Macaluso, E. *et al.* Multisensory stimulation with or without saccades: fMRI evidence for crossmodal effects on sensory-specific cortices that reflect multisensory location-congruence rather than task-relevance. *NeuroImage* (in press)
- 48 Kastner, S. *et al.* (1999) Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron* 22, 751–761
- 49 Calvert, G.A. *et al.* (2000) Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Curr. Biol.* 10, 649–657
- 50 Macaluso, E. *et al.* (2004) Spatial and temporal factors during processing of audiovisual speech: a PET study. *NeuroImage* 21, 725–732
- 51 Fujisaki, W. *et al.* (2004) Recalibration of audio-visual simultaneity. *Nat. Neurosci.* 7, 773–778
- 52 Macaluso, E. *et al.* (2004) Influence of gaze direction on modulation of visual ERPs by endogenous tactile spatial attention. *Cogn. Brain Res.* DOI:10.1016/j.cogbrainres.2004.11.003 (<http://www.sciencedirect.com/science/journal/09266410>)