The functional anatomy of visual-tactile integration in man: a study using positron emission tomography

R.B. Banati*a, G.W. Goerresb, C. Tjoac, J.P. Aggletonc, P. Grasbya

aImperial College, School of Medicine, MRC-Cyclotron Unit, Hammersmith Hospital, Du Cane Road, London W12 ONN, UK
bDivision de médecine nucléaire, Hôpitaux Universitaires, Rue Micheli-du-Crest, CH-1211 Genève, Switzerland
cSchool of Psychology, Cardiff University, Cardiff CF1 3YG, UK

Received 21 August 1998; received in revised form 5 May 1999; accepted 11 May 1999

Abstract

The integration of neural signals from different sensory modalities is a prerequisite for many cognitive and behavioural functions. In this study, we have mapped the functional anatomy of the integration of sensory signals across the tactile and visual modalities. Using the PET radiotracer H215O, regional cerebral blood flow (rCBF) changes were measured in eight normal volunteers performing crossmodal recognition of simultaneously presented visual and tactile stimuli using a modified version of the ‘arc-circle test’. Whilst intramodal matching within the visual modality led to relative rCBF increases in the visual association cortex, crossmodal matching (visual–tactile), when compared to intramodal matching, was accompanied by relative rCBF increases in the anterior cingulate cortex, inferior parietal lobules, the left dorsolateral prefrontal cortex (DLPFC) and the left claustrum/insular cortex. The pattern of brain activation is congruent with areas of heteromodal and supramodal cortex and indicates that activation of multimodal areas is required to solve the crossmodal problem. © 1999 Elsevier Science Ltd. All rights reserved.

Keywords: PET; Sensory integration; Crossmodal; Association cortex; Brain activation

1. Introduction

Many behaviours (e.g. examination of a held object) involve the convergence and integration of information conveyed through anatomically distinct sensory pathways [9,33]. This process of combining information from different sensory streams or ‘multimodal integration’ has been investigated in anatomical, neurophysiological, and behavioural studies [9,14,15,33,35,46]. Heteromodal cortex receives input from more than one unimodal, i.e. primary sensory, area. It contains either neurons responding to more than one modality or has closely interspersed populations of modality-specific neurons which are responsive to different modalities [9]. Supramodal (or transmodal) cortex, such as the dorsolateral prefrontal cortex (DLPFC), receives input from second order sensory association cortex [47] and has no specificity for any single modality of sensory input [33].

It has been suggested [34] that heteromodal/supramodal cortex may be the site at which modality-specific sensory inputs are bound into a multimodal representation. Candidate regions have included the prefrontal cortex, posterior parietal cortex, lateral temporal cortex, insula/claustrum, superior colliculus, parahippocampal gyrus, amygdala, rhinal cortex and specific nuclei in the thalamus [9,33,56]. Based on neuropsychological and neuroanatomical observations, impairment or failure to integrate sensory input across modalities, has been postulated to occur in schizophrenia, and Gerstmann syndrome [31,53]. Schizophrenic patients appear particularly impaired during tasks requiring performance in more than one
modality and attentional shift between them [31]. In Gerstmann syndrome a lesion in the left parietal cortex leads to a collection of symptoms, including amongst others visual–spatial abnormalities and failure to perform verbally mediated spatial functions that may reflect the damage of crossmodal connections [1].

Crossmodal recognition involving the visual and the tactile system is an example of multimodal integration [43,44]. On the basis of neuroanatomical considerations integration may occur in the insula, inferior parietal or superior temporal cortex which receives input from visual and tactile modalities. Lesion studies in monkeys suggest that anterior cingulate cortex, hippocampal areas, and amygdala and/or adjacent rhinal cortex are involved in tasks that require integration of tactile and visual information [3,40,41,55,56].

This human PET study mapped the neural correlates of the crossmodal and intramodal recognition of sensory inputs from tactile and visual modalities using a modified ‘arc-circle test’ [42–44,55,60]. Given the neuroanatomical and behavioural studies noted above we hypothesised that crossmodal recognition, when compared to intramodal, would be characterised by activation in areas of heteromodal and supramodal cortex. The arc circle test was selected because (a) it minimises verbal mediation in a crossmodal problem, (b) sensory inputs and motor outputs are closely matched in both intramodal and crossmodal tasks, (c) the simultaneous presentation of tactile and visual stimuli minimises any explicit memory load.

2. Material and methods

2.1. Subjects

We studied eight healthy, right-handed male volunteers, with a mean age of 38.6 y (range: 30–53 y). The subjects had no significant medical, neurological or psychiatric history. Seven out of the eight subjects were educated to university level. Informed written consent was obtained after the procedures had been fully explained. Ethical approval was given by the local hospital ethics committee and permission to administer radioactive H215O was given by the Administration of Radioactive Substances Advisory Committee of the Department of Health, United Kingdom (ARSAC).

2.2. Experimental design

Using the PET tracer H215O, twelve measurements (i.e. 12 scans) of regional cerebral blood flow (rCBF) were made sequentially in each subject. PET scans were performed under two different experimental conditions (crossmodal and intramodal recognition), each condition occurring six times in a randomised sequence.

2.2.1. Crossmodal matching condition (tactile–visual)

To investigate the crossmodal matching of simultaneously presented visual and tactile stimuli a modified version of the ‘arc-circle test’ was used [43,44,55]. This is a tactile–visual matching task requiring subjects to make a crossmodal match between a metal arc placed on a card (tactile presentation) and one of four circles presented simultaneously on a computer screen in front of the subject (visual presentation) (Fig. 1). Subjects did not see the metal arcs either before or during the experiments and were instructed to fix a point on the screen. The visual stimuli consisted of a row of four circles of varying diameters (2.5, 3.0, 3.5, 4.0, 4.5, 5.0 cm) presented on the screen. Beneath the circles an arc of varying length was also presented simultaneously. This visual arc was included to match the intramodal condition, but the subjects were told that it was not part of the crossmodal recognition task. The metal arcs were segments of either 80 or 120° and matched in diameter one of the four circles displayed visually. During the PET emission scan a random sequence of 10 metal arcs, hidden from sight behind a partition, had to be touched using the right index finger and matched with one of the four circles on the screen of which the arc was a matching segment. The simultaneously presented arc on the screen (that did not match any of the visual or tactile stimuli) was in this condition to be ignored. Each of the 10 metal arcs to be touched was associated with a new presentation of four circles (with an irrelevant arc beneath). A new metal arc together with the new screen display was presented every four seconds following a short signal tone which cued subjects to name the visually presented circle that matched the metal arc by identifying its position in the row, i.e. 1, 2, 3 or 4.
2.2.2. Intramodal matching condition (visual–visual)

To measure matching within the visual modality, subjects had to view the same circles and touch the same metal arcs as in the crossmodal matching condition. This time, however, none of the 10 metal arcs, which the subjects touched, matched in diameter any of the circles presented on the screen. In contrast to the crossmodal condition, the previously ignored single arc on the screen now matched one of the four circles on the screen. A new set of circles with accompanying arc (and a new metal arc) was presented every four seconds following a short signal tone which cued subjects to name the circle (by identifying its position in the row, i.e. 1, 2, 3 or 4) that matched the presented screen arc. Subjects were instructed to make regular, simple finger movements to and fro along the metal arc at a set frequency to minimise differences of finger movements between the two tasks. Subjects were practised on this aspect of the task before PET scanning. In addition, during task performance, the execution of finger movements was observed to ensure similarity of tactile input in both tasks.

2.2.3. Experimental differences and similarities of conditions

We chose a two condition subtraction design to maximise the detection of rCBF change due to the cognitive components of interest, whilst sacrificing comparisons with a resting state. The presentation and timing of the visual and tactile stimuli were identical in both tasks. The only overt difference between conditions lay in the different verbal instructions (made explicit before each scan) concerning the matching process required, namely to perform the matching either within the visual modality or across the tactile to the visual modality.

2.3. Data acquisition

Regional cerebral blood flow was measured with PET by recording the distribution of cerebral radioactivity following the intravenous bolus injection of $^{15}$O. Any increase in rCBF is associated with an increase in the amount of oxygen-15-tagged water recorded from that region [32]. PET scans were performed using a CTI 953B PET-camera (CTI Inc., Knoxville, TN), the physical characteristics of which have been described elsewhere [59].

Prior to each rCBF measurement a background scan was performed lasting for 30 s. After completion of the background scan, a 20-s bolus of $^{15}$O was injected into the patient’s ante-cubital vein, followed by a 20-s flush with normal saline solution. The infusion rate was 10 ml/min. A 90-s emission scan was started to coincide with the rising phase of radioactivity counts recorded from the head. Five seconds before this scan the subject began the task. The subjects continued with the task for 40 s. Each of the twelve scans were separated by an interval of 10 min to allow for decay of radioactivity.

Each scan, consisting of 10 crossmodal or intramodal matching tasks, lasted for 40 s during which rCBF measurements were made. Responses were registered for each subject.

All scans were performed with the subject lying supine with their eyes open in a lighted room. A stable head position was maintained by use of a moulded head rest.

2.4. Data analysis and statistics

Calculations and image manipulations were carried out on a Sparc II workstation (Sun Computers Europe Inc, Surrey, England). Analysis of images was undertaken with statistical parametric mapping by using SPM95 software from the Wellcome Dept. of Cognitive Neurology, London (UK) implemented in Matlab (Mathworks Inc., Sherborn MA, USA). The scans from each subject were realigned using the first scan as reference [18]. Following each realignment, all images were transformed into a standard space [61].

This normalising spatial transformation matches each scan to a reference template image that already confirms to the standard space [18]. As a final step before statistical analysis, the images were smoothed using an isotropic Gaussian kernel (FWHM 16 mm). The condition, subject and covariate effects were estimated according to the general linear model at each voxel [19]. Hypotheses about regionally specific condition effects were tested by comparing the estimates using linear compounds of contrasts. The resulting set of voxel values for each contrast constitutes a statistical parametric map of the t statistic, SPMt. The SPMt were transformed to the unit normal distribution (SPMZ) and thresholded at $p < 0.001$ for significance (uncorrected for multiple comparisons). This level of significance requires a hypothesis-led data analysis. Here, the hypothesis was that activations associated with crossmodal recognition would occur in heteromodal/supramodal cortex. Activated pixels were displayed on coronal, sagittal and transverse projection maps of the brain.

Differences of rCBF between tasks were investigated by (a) comparing all crossmodal scans with all intramodal scans irrespective of performance, (b) comparing only crossmodal and intramodal scans with similar levels of performance (i.e. number of correct responses). This was achieved by excluding two crossmodal (lowest scoring) and two intramodal (highest scoring) scans per subject, leaving a total of eight scans per individual.

Thirty one planes of data were obtained for each
subject. All eight subjects commonly shared an axial field of view from \( Z = -16 \) mm to \( Z = +52 \) mm in the stereotactic space of Talairach and Tournoux [61].

3. Results

3.1. Task performance

The ‘arc-circle tests’ were a sensitive measure of crossmodal and intramodal recognition with the performance of subjects always above chance (25%) and below ceiling (100%). All subjects reached higher scores in the visual–visual matching tasks than in the tactile–visual task (Fig. 2). For the intramodal matching tasks a mean of 66% (SD = 15.8%) correctly matched arcs and circles was reached whilst for the crossmodal task 46% (SD = 7.5%) correct matches were obtained.

To counteract the difference in the difficulty between the intra- and the crossmodal task, the two highest scoring (intramodal) and the two lowest scoring (crossmodal) trials of each individual subject were excluded from analysis, thus equating differences in the average performance.

Fig. 2. Average performance scores achieved by the eight subjects studied: Average scores for the intramodal tasks were higher (correct matches: 66%, SD = 15.8%) than the scores reached in the crossmodal matching tasks (correct matches: 46%, SD = 7.5%). All subjects had lower scores in the crossmodal tasks.

Fig. 3. Non-adjusted: The average performance level for the intramodal task was 66% correct matches (SD = 15.8%) and was consistently higher than the crossmodal tasks (correct matches: 46%, SD = 7.5%). Performance-adjusted: To equate the degree of difficulty between the intra- and the crossmodal tasks, the two highest scoring (intramodal) and the two lowest scoring (crossmodal) scans of each subject were excluded from the group analysis. The adjusted average performance rate was then intramodal 54% (SD = 12%) and crossmodal 52% (SD = 5%) correctly matched tasks (Fig. 4).
3.2. rCBF changes during crossmodal recognition

3.2.1. All scans

Crossmodal matching (crossmodal condition minus intramodal condition) led to relative rCBF changes in the right anterior cingulate cortex, both inferior parietal lobules (left > right), the left dorsolateral prefrontal cortex (DLPFC) and a trend level ($p < 0.01$) rCBF change in the left claustrum/insular cortex (Table 1).

3.2.2. Performance-matched scans

A similar pattern of relative rCBF changes, i.e. right anterior cingulate cortex, inferior parietal lobules bilaterally, the left dorsolateral prefrontal cortex (DLPFC) and left claustrum/insular cortex, was seen in the performance matched scans during crossmodal condition. Additional rCBF changes were seen in the middle and superior temporal gyri.

3.3. rCBF changes during intramodal recognition

3.3.1. All scans

During the intramodal recognition (intramodal condition minus crossmodal condition) relative rCBF changes occurred exclusively in the visual association cortex but spared the primary visual area (Table 2).

3.3.2. Performance-matched scans

A similar pattern of relative rCBF changes in the visual association cortex sparing the primary visual area was seen in the performance matched scans during intramodal condition (Table 2).
4. Discussion

The crossmodal recognition task (tactile–visual) revealed relative rCBF increases in the anterior cingulate, inferior parietal lobule, dorsolateral prefrontal cortex, insula and middle and superior temporal gyri. As predicted, these areas are heteromodal or supramodal cortex with converging sensory input from two or more modalities [26,38,53]. In contrast intramodal recognition (visual–visual) led to relative rCBF increases in the visual association cortex alone. The tasks were matched for sensory inputs and motor outputs and, therefore, it is unlikely that the pattern of rCBF simply reflects differences of sensory inputs or motor outputs between tasks. This is corroborated by our findings that rCBF did not change between tasks in the frontal eye fields involved in ocular [2] movement, or in the primary and supplementary motor cortex controlling simple finger movement [11,13,20,24,25].

4.1. Brain areas activated during intramodal (visual–visual) recognition

4.1.1. Visual cortex (BA 18/19)

Visual–visual recognition resulted in marked relative rCBF increases in the visual association cortex, namely the lingual and fusiform gyri (BA 18/19), but no rCBF change in primary visual cortex (BA 17). Activation of the visual association cortex is widely reported in tasks involving processing of visual stimuli such as colour, movement, form, reading and topographical mental

Table 1
Areas of relative rCBF changes, crossmodal task vs intramodal task (size = spatial extent of each activation shown in numbers of contiguous voxels above the threshold of \( p < 0.01 \))

<table>
<thead>
<tr>
<th>Brain structure</th>
<th>Left</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( x )</td>
<td>( y )</td>
</tr>
<tr>
<td>(A) Areas of increased rCBF, crossmodal task vs intramodal task. Total data set (8 ( \times ) 12 scans)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anterior cingulate cortex (BA 32/24)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Inferior parietal lobule (BA 40)</td>
<td>–48</td>
<td>–32</td>
</tr>
<tr>
<td>Dorsolateral prefrontal gyrus (BA 46)</td>
<td>–30</td>
<td>34</td>
</tr>
<tr>
<td>Insular cortex</td>
<td>–38</td>
<td>2</td>
</tr>
<tr>
<td>(B) Areas of increased rCBF, crossmodal task vs intramodal task. Performance matched (8 ( \times ) 8 scans)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anterior cingulate cortex (BA 32)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Inferior parietal lobule (BA 40)</td>
<td>–50</td>
<td>–28</td>
</tr>
<tr>
<td>Middle and superior temporal gyri (BA 21/22)</td>
<td>–42</td>
<td>–44</td>
</tr>
<tr>
<td>Dorsolateral prefrontal gyrus (BA 46)</td>
<td>–32</td>
<td>38</td>
</tr>
<tr>
<td>Insular cortex</td>
<td>–36</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 2
Areas of rCBF changes, intramodal task vs crossmodal task

<table>
<thead>
<tr>
<th>Brain structure</th>
<th>Left</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( x )</td>
<td>( y )</td>
</tr>
<tr>
<td>(A) Areas of increased rCBF, intramodal task vs crossmodal task. Total data set (8 ( \times ) 12 scans)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual cortex, lingual gyrus and fusiform gyrus (BA 18/19)</td>
<td>–6</td>
<td>–84</td>
</tr>
<tr>
<td></td>
<td>–30</td>
<td>–92</td>
</tr>
<tr>
<td>(B) Areas of increased rCBF, intramodal task vs crossmodal task. Performance matched (8 ( \times ) 8 scans)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual cortex, lingual gyrus and fusiform gyrus (BA 18/19)</td>
<td>–6</td>
<td>–84</td>
</tr>
<tr>
<td></td>
<td>–30</td>
<td>–82</td>
</tr>
</tbody>
</table>
imagery [6,12,16,27]. Our findings imply that differential involvement of the primary visual cortex is not a requirement for visual–visual recognition.

4.2. Brain areas activated during crossmodal (tactile–visual) recognition

4.2.1. Anterior cingulate cortex (BA 32 and BA 24)

The greatest statistical change of rCBF occurred in the anterior cingulate cortex. Activations of the anterior cingulate cortex are frequently reported in functional PET studies of difficult tasks with high attentional demands. For example, activity in this area increases in visual attention tasks whenever the number of targets in a set of stimuli is increased, and decreases with practice [49,50].

The difference in cingulate rCBF remained unchanged after adjustment for behavioural performance. This may suggest that rCBF increases in the anterior cingulate during crossmodal recognition may not be simply ascribed to increased ‘attention’ or ‘difficulty’ associated with the crossmodal task. The widespread connections of the anterior cingulate [26] might suggest a more fundamental role, such as the facilitation of cross-talk between modalities. However, further experiments designed to manipulate attentional load and other components that contribute to ‘task difficulty’ would be necessary to establish this point.

Indeed, lesion studies in monkeys have underlined the importance of the anterior cingulate cortex for successful crossmodal recognition: while deficits in crossmodal recognition are not associated with damage to single specific multimodal areas in the neocortex, lesioning of the anterior cingulate cortex leads to a complete loss of crossmodal recognition abilities [3,41].

4.2.2. Inferior parietal lobule (BA 40)

Discrete and symmetric activations were observed in the inferior parietal lobule in the vicinity of SII during crossmodal recognition. The inferior parietal lobules are heteromodal areas [36] serving high-level somatosensory processing. For example, somatosensory discrimination of roughness (microgeometry) has been associated with activations in the inferior part of the parietal lobe in man [28,45]. Somatosensory stimulation of the right hand (by roughness and a square pulse indentation of the skin on the index finger) activates (in addition to the anterior and posterior lip of the postcentral sulcus) areas in the parietal operculum, i.e. overlapping regions of the somatosensory association cortex (SII, B40) and the supramarginal gyrus [28]. Similarly, discrimination of length has also been reported to activate the postcentral sulcus, the supramarginal gyrus and angular gyrus bilaterally and the parietal operculum ipsilaterally to the stimulated hand [28].

In monkeys, ablation of SII leads to severe impairment in the learning of texture and shape discrimination tasks [40]. These results imply that SII is a critical station in a tactile processing pathway that proceeds from the primary somatosensory cortex (SI) to the limbic structures of the temporal lobe through links in SII and the insular cortex. In man, lesions including area 40, activated in our study, are associated with deficits of shape recognition and may cause tactile agnosia and spatial neglect [23,51,52]. Activations in this area or adjacent areas are also reported in tasks that require the use of implicit motor imagery for visual shape discrimination [48].

A major component of the arc circle test requires subjects to construct a representation of the shape of the touched metal arc so that it can be matched to the visually presented circles. As brain activations caused by touch alone were likely to have been matched across scan conditions (subjects performed similar finger movements along the metal arc in both conditions and no relative activations occurred in SI) the inferior parietal lobe activation during crossmodal recognition probably reflects higher order tactile processing, namely the extraction of shape information. Interestingly, disturbances of such higher order processing of spatial information has recently been linked to the hyperactivation of the inferior parietal cortex (together with cingulate cortices) in schizophrenic patients experiencing delusions of passivity (alien control) [58].

4.2.3. Dorsolateral prefrontal gyrus (DLPFC; BA 46)

A striking asymmetrical relative rCBF increase was seen in the left DLPFC during crossmodal recognition. Even when the statistical threshold was lowered this asymmetry remained. In contrast to other multimodal areas of the frontal cortex, DLPFC receives projections also from second-order sensory association areas, such as visual, auditory and sensory association areas [47]. The DLPFC is implicated in higher cognitive functions, such as initiation of motor behaviour [2,24,57], spatial working memory [4] and executive/planning functions [5] all of which would be expected to be cognitive components of crossmodal performance.

4.2.4. Middle and superior temporal gyri (BA 21, 22)

Activation of the middle and superior temporal gyri was seen when examining scans matched for performance. The superior temporal gyrus contains both heteromodal association cortex as well as primary sensory, and sensory association cortex. [26,54]. Middle and superior temporal gyri are part of the activation pattern in somatosensory discrimination tasks with recognition of shape and surface properties, such as roughness [28,45].
4.2.5. Insula/claustrum

Left-sided relative rCBF increases in the insula/claustrum were seen during the crossmodal recognition task (this was at a trend level before matching performance when \( Z = 2.86 \) and was significant after matching performance with \( Z = 3.77 \)). The claustrum and insula may be developmentally distinct and play different roles in cognitive processing. However, the spatial resolution of the PET scan data does not allow the activation to be localised unequivocally to one or the other structure.

The insular cortex has connections with diverse auditory, somatosensory, olfactory, limbic and paralimbic structures [37]. Given the diverse connectivity, this heteromodal area is likely to form part of the anatomical basis for crossmodal recognition. Indeed, 2-deoxy-[14C] glucose studies in monkeys, trained to perform crossmodal recognition tasks, have shown that the left insula/claustrum is consistently activated [22]. Similarly, rCBF changes in the insula/claustrum area are prominent in subjects during tactile–visual discrimination tasks [21] and in tasks combining tactile and vestibular stimuli [7]. Insular activation has also been reported for experimental tasks requiring processing of long-term and recent memory of tactile information [8].

4.2.6. Lack of activation in the amygdala and hippocampus

A structure apparently implicated in crossmodal matching on the basis of lesion studies in animals and man is the amygdala [40,55]. In our study, which included both the amygdala and the hippocampal formation in the field of view, no relative rCBF changes were found in these areas. In fact, the lack of differential activation in the amygdala is in keeping with more recent studies in animals [56] and man [29,42,60] which have failed to find evidence for the involvement of this structure to be crucial in crossmodal recognition. While these inconsistent findings from lesion studies may reflect variability in the extent of amygdala damage [30,39], it seems more likely that they reflect the involvement of the adjacent rhinal cortex or its associated fibre pathways [41]. Although there is recent evidence that the rhinal cortex is an important neural substrate for crossmodal recognition [30,39], this area was not fully imaged in our field of view for all eight subjects (however, no activations were found in a subset of subjects \( n = 4 \) who were scanned low enough to include most of the inferior temporal regions; data not shown).

Our study also failed to find any evidence of differential activation in the hippocampus. This is consistent with a clinical investigation of amnesic patients with bilateral hippocampal lesions who were found to be unimpaired on the arc-circle test [60]. It appears from this and other findings that the hippocampus is not necessary for stimulus–stimulus matching or stimulus–stimulus associative memory [40,41,60].

4.3. Activation pattern during crossmodal recognition

Hadjikhani and Roland [21] recently reported a PET activation study of crossmodal performance using a task, in which 3-dimensional ellipsoids had to be palpated and matched within or across the tactile and visual modality. Performed on eight subjects, their study included additional control tasks, such as a tactile–tactile matching task that allowed the comparison ‘tactile–visual minus tactile–tactile matching’ against ‘tactile–visual minus visual–visual matching’. With this subtraction paradigm, differential rCBF changes during crossmodal matching were found exclusively in the right insula/claustrum. Based on their results, i.e. the absence of a more widespread involvement of heteromodal areas, Hadjikhani and Roland [21] speculated that crossmodal recognition may not be based on a crossmodal transfer within heteromodal convergence areas. Instead, they suggest direct interaction of the modality specific systems with each other, possibly by ways of parallel processing in which the right claustrum/insula synchronises the activity in different modality-specific areas. In contrast, our study—a 2-dimensional extraction of shape within/ across the visual and tactile modality—suggests that crossmodal matching is associated with activations in a number of multimodal convergence areas comprising anterior cingulate, inferior parietal lobe, middle and superior temporal gyrus and dorsolateral prefrontal cortex as well as the left insula/claustrum, as found previously in experimental animals by Horster et al. [22]. Further evidence for multiple activations in heteromodal cortex during crossmodal recognition is provided by a recent functional magnetic resonance imaging (fMRI) study which involved speech perception across the auditory and visual modality and revealed activation in similar areas of heteromodal cortex, such as the inferior parietal lobule, and the middle and superior temporal gyri [10].

4.4. Limitations of present study

A possible limitation in our study is that an intramodal tactile matching task was not performed. However a tactile–tactile control task is difficult to design without introducing additional cognitive demands that are not involved in the test conditions of the visual–visual and visual–tactile matching tasks used in our study. The extent to which such a tactile–tactile condition could serve as a true control task in our study is limited because in the tactile–tactile condition there must always be a delay between the sampling of
the two sequentially presented objects to be compared (unlike the simultaneous tactile–visual matching in our experiment). This delay, caused by sequential sampling, (a) opens the possibility for verbal mediation and (b) increases the likelihood of automatic crossmodal transfer, i.e. asked to make a decision on tactually sampled shape, the subject may choose to visualise the object or employ concurrent motor imagery to help subsequent comparison. If this was to occur, subtraction of a tactile–tactile control task from a tactile–visual task would lead to an underappreciation of areas involved in crossmodal recognition.

Since a resting state condition was omitted in our study (for the sake of higher statistical power) during the matching conditions, it is possible that neural activations in other brain areas might also be necessary but not sufficient for crossmodal matching to occur (e.g. in sensorimotor and visual systems). However, the case remains that the areas identified in our study are preferentially activated during the crossmodal matching condition and, therefore, are critical for this process.

5. Conclusions

In conclusion, our results suggest that crossmodal recognition requires the concurrent activation of distinct heteromodal association (e.g. inferior parietal lobule) and supramodal (e.g. DLPFC) cortex. The described pattern of activation during crossmodal integration may imply that the binding of information into heteromodal representations takes place in the areas identified. However as emphasised by a number of other authors [9,15,33] such involvement of heteromodal areas does not necessarily imply that ‘amodal’ representations actually reside in these areas.

Whether some or all of the brain area activations, reported here, are invariably engaged during sensory integration across all sensory modalities and/or to what extent the activation pattern is task-specific remains to be investigated.

Acknowledgements

RBB is funded by the Medical Research Council and the Deutsche Forschungsgemeinschaft. GWG was supported by the Swiss National Science Foundation and the Swiss Society of Medical Radiology.

References


Streicher M, Ettinger G. Cross-modal recognition of familiar and unfamiliar objects by the monkey, the effects of ablation of polysensory neocortex or of the amygdaloid complex. Behav Brain Res 1987;23:95–107.


