Cross-Modal Transfer of Information between the Tactile and the Visual Representations in the Human Brain: A Positron Emission Tomographic Study

Nouchine Hadjikhani and Per E. Roland

1Division of Human Brain Research, Institute of Neuroscience, Karolinska Institute, 171 77 Stockholm, Sweden

Positron emission tomography in three-dimensional acquisition mode was used to identify the neural populations involved in tactile–visual cross-modal transfer of shape. Eight young male volunteers went through three runs of three different matching conditions: tactile–tactile (TT), tactile–visual (TV), and visual–visual (VV), and a motor control condition. Fifteen spherical ellipsoids were used as stimuli.

By subtracting the different matching conditions and calculating the intersections of statistically significant activations, we could identify cortical functional fields involved in the formation of visual and tactile representation of the objects alone and those involved in cross-modal transfer of the shapes of the objects.

Fields engaged in representation of visual shape, revealed in VV–control, TV–control and TV–TT, were found bilaterally in the lingual, fusiform, and middle occipital gyri and the cuneus. Fields engaged in the formation of the tactile representation of shape, appearing in TT–control, TV–control and TV–VV, were found in the left postcentral gyrus, left superior parietal lobule, and right cerebellum.

Finally, fields active in both TV–VV and TV–TT were considered as those involved in cross-modal transfer of information. One field was found, situated in the right insula–claustrum. This region has been shown to be activated in other studies involving cross-modal transfer of information. The claustrum may play an important role in cross-modal matching, because it receives and gives rise to multimodal cortical projections. We propose here that modality-specific areas can communicate, exchange information, and interact via the claustrum.

Key words: human; PET; claustrum; cross-modal transfer; visual system; somatosensory system

The relations between sight and touch have been for long a matter of debate, as in 1709 the philosopher George Berkeley in his book An Essay Towards A New Theory of Vision concluded that there were no necessary connections between a tactile world and a visual world. Even today there is theoretical and experimental support for the view that there are no cortical convergence regions, in which neuron populations integrate information from different sensory modalities and from different submodalities (Abeles, 1991; Felleman and van Essen, 1991; Young et al., 1992; Singer 1993, 1995). Information from visual submodalities seem to be processed by parallel “functional streams” (Ungerleider and Mishkin, 1982; Livingstone and Hubel, 1988; Haxby et al., 1991; Zeki et al., 1991; Gulyas et al., 1994). Because further direct anatomical connections between somatosensory and visual areas in primates are sparse, if at all existing (Selemon and Goldman-Rakic, 1988; Cavada and Goldman-Rakic 1989; Neal et al., 1990), one might expect that somatosensory and visual information is processed by segregated populations of neurons. Still everyone recognizes a key, whether it is felt in a pocket or seen on a table. How is this possible?

The research done on tactile–visual cross-modal performance has been based mainly on the assumption that there must exist amodal representations of form in so-called polysensory areas, defined as areas activated by stimuli from more than one sensory modality. Nevertheless, in a review on cross-modal abilities in nonhuman primates, Ettlinger and Wilson (1990) concluded that there is no polysensory cross-modal area, no cross-modal region “in which representations formed in one sense would reside and be accessed by another sense,” but suggested instead a system in which the senses can access each other directly from their sensory-specific systems. For the present purpose, we define cross-modal-specific areas as areas activated only when information coming from two or more different sensory modalities is compared.

We examined tactile–visual matching of the shapes of objects. In tactile–visual matching of three-dimensional objects, the tactile information is of a different nature from that of visual information. When the hand is used to palpate an object, the information is sampled in a piecemeal manner, such that only a part of the object surface is covered by the fingers during each sampling path (Roland and Mortensen, 1987), and information is integrated over time to form a truly three-dimensional shape representation. In the visual system, information about an object can be simultaneously obtained and transferred to the visual cortex, but if it is seen from a stationary angle of view, only a part of its surface is sampled. This difference speaks against any common polyomodal or amodal representation for the two modalities.

We studied tactile–tactile and visual–visual intramodal matching aiming to identify cortical fields engaged in processing and representation of tactile and visual shape. Putative polysensory areas thus should be activated by tactile as well as visual intramodal shape matching. We also studied tactile–visual cross-modal matching of shape with the purpose of identifying cross-
Figure 1. The fifteen spherical ellipsoids. The fifteen ellipsoids used in this experiment were made to have the same weight and surface, differing by their shape only. Two similar sets were used.

TT matching. The first ellipsoid in the pair was palpated for 3–4 sec, and the second was palpated for 2–3 sec. A new pair was presented approximately every eighth second. The subjects were not able to see the ellipsoids. If they thought that the two ellipsoids had identical shapes, the subjects had to raise their right thumb.

TV matching. The tactile and the visual stimuli were presented simultaneously. The subjects had ~3 sec to palpate one ellipsoid (without being able to see it) and three seconds to look at the ellipsoid that was presented on a shelf placed over their head. A new pair was presented approximately every fourth second. A mirror allowed them to see the objects at a distance of 60 cm. The subjects again extended their right thumb if they thought that the two ellipsoids were identical.

VV matching. The first ellipsoid in the pair was presented on the shelf for ~4 sec, and then the second ellipsoid was presented for ~2 sec. A new pair was presented every eighth second. The subjects answered by raising their right thumb if they thought that the two ellipsoids had identical shapes.

Control. Subjects fixated the same fixation point as in the TT experiment and were instructed to move their right hand in the same way as if they were actually palpating the objects.

The monitoring of eye movement was done by a video recording on a split screen of the subject’s face. The movement of the right hand of the subjects was recorded as well. Subjects were told not to speak during the entire procedure, and the room was kept as quiet as possible.
Psychophysics

Before the positron emission tomographic (PET) experiment, five additional subjects underwent psychophysical testing of the same conditions. The only notable exception is that we tested separately visual–tactile matching and tactile–visual matching to be sure that there was no difference to be noted depending on the modality presented first. Each subject went through four runs of 120 pairs of ellipsoids, with a priori varying probabilities of matching. Receiver operating characteristic curves and $d'$ values measuring the separation between the means of the noise and signal distribution were calculated for each subject in each modality (Green and Swets, 1966).

Scanning procedure

Each subject, lying in a supine position and equipped with a stereotactic helmet (Bergström et al., 1981), had a high-resolution MRI scan and a PET scan. The MRI scans were done using a spoiled gradient echo sequence obtained with a 1.5 T General Electric Signa scanner [echo time, 5 msec; repetition time, 21 msec; flip angle, 50°, giving rise to a three-dimensional (3D) volume of 128 $\times$ 256 $\times$ 256 in isotropic voxels of 1 mm$^3$]. Each subject had an arterial catheter inserted under local anesthesia in the left radial artery for the measurement of arterial concentration of radiotracer.

The regional cerebral blood flow (rCBF) was measured in 3D acquisition mode with a Siemens ECAT EXACT HR PET camera (for technical description, see Wienhard et al., 1994). The radiotracer used was $^{15}$O-labeled butanol, which was synthesized according to the method of Berridge et al. (1991). Fifteen microcuries of radiotracer were injected intravenously as a bolus at the beginning of each run, followed by a 20 ml flush with saline. The tasks began intravenously as a bolus at the beginning of each run, followed by a 20 ml flush with saline. The tasks began

Statistical analysis

The statistical analysis tested the hypothesis that clusters of high $t$ values occur by chance in the standard anatomical space and was described in detail previously (Roland et al., 1993). In short, individual voxel-by-voxel rCBF subtraction images were calculated for each subject and averaged. For example, for a given test repeated $t$ times in subject $k$:

$$
\Delta rCBF_{\text{TT},i,k-\text{CTRL},i,k} = rCBF_{\text{TT},i,k} - rCBF_{\text{CTRL},i,k}
$$

These images were subsequently averaged.

$$
\Delta rCBF_{\text{TT},k-\text{CTRL},k} = \frac{1}{2} \sum_{i=1}^{3} \Delta rCBF_{\text{TT},i,k-\text{CTRL},i,k}
$$

These individual mean images were used to calculate a group mean image:

$$
\Delta rCBF_{\text{TT}-\text{CTRL}}
$$

and eventually a $t$ image was calculated:

$$
\frac{\Delta rCBF_{\text{TT}-\text{CTRL}}}{\text{SEM}_{\text{TT}-\text{CTRL}}}
$$

The spatial three-dimensional autocorrelation was then determined in Student’s $t$ pictures obtained from the $\Delta rCBF_{i,k}$ images of subtracting two TT matching images to give a noise image, as described by Roland et al. (1993). The resulting noise $t$ pictures were thresholded at different $t$ values, and clusters of voxels of suprathreshold values were identified. We simulated 2000 groups of eight subjects each for each $t$ threshold. From these noise $t$ picture tables of clusters of suprathreshold $t$ values, exceeding a certain number of voxels, were produced (Roland et al., 1993). The criteria used for accepting rCBF changes in adjacent clustered voxels as activations were set so that there was an average probability of $p < 0.1$ of finding one false-positive cluster or more within the three-dimensional space of the standard anatomical brain format. Accordingly the descriptive Student’s $t$ pictures were thresholded such that the (omnibus) probability of finding one or more false-positive clusters was $p < 0.1$. The resulting cluster images thus show only the activated parts of the brain and zero elsewhere. The significance of each cluster was also assessed by nonparametric method of Holmes et al. (1996). For a $t$ threshold of 2.5 this method gave $p < 0.1$ for one or more false-positive clusters of $\approx$900 mm$^3$ in size.

These cluster images are henceforth referred to as TT–control, TV–control, VV–control, TV–TT, and TT–VV. For the convenience of the reader, the different significant clusters are listed in Tables 2 and 3 in accordance with the method of Roland et al. (1993). In Tables 2 and 3 the average $t$ value of each cluster is also calculated as the mean of the values of the voxels constituting the cluster.

The cluster images TT–control, TV–control, VV–control, TV–TT, and TT–VV were then used to form Boolean intersection images (Ledingberg et al., 1995), as, for example, TT–control $\cap$ TV–VV. This Boolean intersection carries no assumptions and shows the intersections or overlaps of clusters that correspond to cortical or subcortical regions active in both TT–control and TV–VV. For example, if a cluster in TT–control has $p < 0.05$ of being a false-positive, the probability that any cluster from TV–control by chance will overlap is $p < 0.05$, because the prerequisite is that the TT–control cluster is present.

RESULTS

Psychophysics

The psychophysical testing showed that there is a linear relationship between presented and chosen stimulus, regardless of the modalities, as shown by the linear regression curves. The direction of cross-modal information transfer (i.e., tactile to visual vs visual to tactile) had no influence on the performance (Fig. 4).

During the actual PET scanning, the probability of a matching given a matching pair, i.e., $p$(match | match), and the probability of a responding match given a nonmatching pair, i.e., $p$(match | nonmatch), was calculated for each subject. On the basis of these probabilities a measure of the performance $d'$ was calculated (Greens and Swets, 1966). The subjects had $d'$ values between 0.40 and 2.88, indicating that all of them actually performed above chance or noise level. Furthermore, this range of $d'$ values was comparable to the range of 0.65–3.0 obtained by the five subjects performing the psychophysical test outside of the PET camera, indicating no major differences between the two groups. For the subjects doing PET, there were no differences in $d'$ between conditions, indicating the same level of difficulty between the tasks. By paired comparison of $d'$ values, the subjects being their own control, the average intrasubject differences in $d'$ were TV–VV, $-0.66 \pm 0.93$ (SD); and TV–TT, $-0.08 \pm 0.61$.

To examine whether the motor activity of exploring the ellipsoids actually was balanced between the matching tasks (i.e., TV and TT), between TV and the control condition, and between TT and the control condition, we analyzed the frequencies of movements of the individual fingers of the subjects doing PET. By a paired comparison, the subjects being their own controls, the average intrasubject differences were in TT-control for the thumb ($-0.05 \pm 0.08$ Hz) and for the index ($0.02 \pm 0.12$ Hz). The corresponding differences in frequencies for TV control were for the thumb ($-0.02 \pm 0.07$ Hz) and for the index ($-0.06 \pm 0.13$ Hz). The number of thumb and index movements in TT, TV, and control did not vary very much for a given subject: the intrasubject SDs were between 0.6 and 8%. The subjects fixed the fixation point as instructed during the TT and the rest condition, and there was no significant eye movement while the subjects were presented the visual stimuli in TV and in VV.
Regional cerebral blood flow changes

Table 1 shows the rationale of the different subtractions. Somatosensory areas and perhaps polysensory areas participating in the formation of tactile representation of the stimuli were assumed to be active in tactile–tactile minus control, tactile–visual minus control, and tactile–visual minus visual–visual. In each graph, the responses of the subjects deciding that two objects were identical are shown versus the actual stimuli presented by the experimenter. Error bars indicate SD. The regression lines of chosen versus given stimuli were calculated together with the correlation coefficient. Even if the performance of the subjects was slightly better in the visual–visual matching condition, the overall performance does not depend on the mode of presentation of the stimulus. The responses are linear. Cross-modal comparison gives the same results as intramodal matching. Furthermore, there is no notable difference depending on whether the ellipsoids are first presented visually and then haptically or are presented in the reverse order.

<table>
<thead>
<tr>
<th></th>
<th>sensory tactile areas activated</th>
<th>sensory visual areas activated</th>
<th>polysensory areas activated</th>
<th>intramodal matching</th>
<th>cross-modal matching</th>
</tr>
</thead>
<tbody>
<tr>
<td>TT-ctrl</td>
<td>X</td>
<td>0</td>
<td>X</td>
<td>0</td>
<td>X</td>
</tr>
<tr>
<td>TV-ctrl</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>0</td>
<td>X</td>
</tr>
<tr>
<td>VV-ctrl</td>
<td>0</td>
<td>X</td>
<td>X</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TV-TT</td>
<td>0</td>
<td>X</td>
<td>0</td>
<td>0</td>
<td>X</td>
</tr>
<tr>
<td>TV-VV</td>
<td>X</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>X</td>
</tr>
<tr>
<td>(TV-TT) (\cap) (TV-VV)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>XX</td>
</tr>
<tr>
<td>TV-VV (\cap) TT'</td>
<td>X</td>
<td>0</td>
<td>X</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TV-TT (\cap) VV'</td>
<td>0</td>
<td>X</td>
<td>X</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TT' (\cap) TV (\cap) VV'</td>
<td>0</td>
<td>0</td>
<td>XX</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

ctrl = control; TV' = TV-control; TT' = TT-control; VV' = VV-control

Figure 4. Psychophysical experiment. This figure shows the results of the psychophysical experiment that was done before the PET study. Four different paradigms are presented: tactile–tactile, visual–visual, tactile–visual, and visual–tactile. In each graph, the responses of the subjects deciding that two objects were identical are shown versus the actual stimuli presented by the experimenter. Error bars indicate SD. The regression lines of chosen versus given stimuli were calculated together with the correlation coefficient. Even if the performance of the subjects was slightly better in the visual–visual matching condition, the overall performance does not depend on the mode of presentation of the stimulus. The responses are linear. Cross-modal comparison gives the same results as intramodal matching. Furthermore, there is no notable difference depending on whether the ellipsoids are first presented visually and then haptically or are presented in the reverse order.

Regional cerebral blood flow changes

Table 1 shows the rationale of the different subtractions. Somatosensory areas and perhaps polysensory areas participating in the formation of tactile representation of the stimuli were assumed to be active in tactile–tactile minus control, tactile–visual minus control, and tactile–visual minus visual–visual. By calculating the intersection between the two independent cluster images of TV–VV and TT–control, we expected to isolate those areas involved in the formation of the tactile representation of the stimuli. Visual areas and perhaps polysensory areas were expected to be activated during the visual matching in the following paradigms: visual–visual minus control, tactile–visual minus control, and tactile–visual minus tactile–tactile. The overlap between TV–TT and VV–control was expected to show areas specifically engaged in the formation of the visual representation of the stimuli. The overlap between the two cluster images of the cross-modal matching tasks, i.e., TV–TT and TV–VV, would isolate areas specifically engaged in the cross-modal matching procedure, whereas the overlap between all of the tasks minus control would isolate polysensory areas activated regardless of the mode of stimulation. Tables 2 and 3 show the location (i.e., center of gravity), volume, and mean t value of the fields of activation in the
different tasks. Tables 4–6 show the location and extent of the overlaps of fields engaged in the formation of tactile–visual representation of the stimuli and in the cross-modal transfer of information.

**Tactile–tactile minus control**

Several fields of activation were found in the parietal lobe. The biggest increase of rCBF was situated in the left postcentral gyrus, extending posteriorly from the posterior part of the gyrus and the cortex lining the postcentral sulcus into superior parietal lobule and the anterior part of the intraparietal sulcus. A second focus of activation was found in the right parietal lobe, situated in the supramarginal gyrus (Table 2).

Other foci of activation were found in the right thalamus, the right temporal pole, and the left anterior prefrontal cortex. The cerebellum showed several foci of activation bilaterally.

**Tactile–visual minus control**

Fields of activation were found mainly in the occipital, temporal, and parietal lobes.

The biggest activation cluster was situated on the left cuneus, extending anteroposteriorly along the calcarine sulcus. A second cluster was situated on the left superior occipital gyrus, on its posterior part. A third cluster was situated on the left middle occipital gyrus. The right occipital cortex contained one cluster of activation that was situated on the lingual gyrus and the cortex lining the collateral sulcus, extending anteriorly to the fusiform gyrus (Table 3).

In the left parietal lobe, a field of activation was found in the postcentral gyrus, extending medially onto the cortex lining the postcentral gyrus and on the posterior parietal cortex, including the cortex lining the anterior part of the intraparietal sulcus. A second cluster was situated on the left superior parietal lobule. A third cluster was found in the left precuneus and in the cortex lining the parieto-occipital sulcus. On the right parietal lobe, we found a cluster of activation on the superior parietal lobule that was situated more posteriorly than the one on the left side (Table 2).

Other foci of activation were found in the right thalamus, in the anterior prefrontal cortex bilaterally, and in the right cerebellum. Two minor clusters at the location of the right insula–cerebrum were found of 172 and 143 mm$^3$ with mean $t$ values 3.35 and 3.34, respectively. These were not statistically significant.

**Visual–visual minus control**

In the occipital lobes, several clusters of activation were found bilaterally. The most important one was situated on the left lingual gyrus, extending anteroposteriorly along the calcarine sulcus (Table 3).

The inferior part of the middle occipital gyri was bilaterally activated, and the left superior part of the middle occipital gyrus also showed a cluster of activation. The right fusiform gyrus was activated in its posterior part.

The parietal lobes showed two foci of activation on the right side. One was situated in the superior parietal lobule and extended to the cortex lining the intraparietal sulcus, and another was in the angular gyrus.

We found bilateral fields of activation in the anterior and posterior cingulate cortex and in the anterior part of the middle frontal gyrus. The left anterior insula was also activated.

**Tactile–visual minus visual–visual**

According to our hypothesis, this subtraction should reveal the areas engaged in the tactile exploration of the stimuli and in cross-modal matching. Clusters of activation were found in the parietal lobes, with a big cluster of activation centered on the left postcentral gyrus, extending anteriorly to the cortex lining the central sulcus and posteriorly to the postcentral sulcus and the anterior part of the intraparietal sulcus as well as going more
medially on the superior parietal lobule; a second cluster was situated on the right postcentral gyrus. Another cluster was found at the bottom of the left central sulcus. The left middle cingulate and the left thalamus were activated. The cerebellum showed bilateral foci of activation. Finally, a cluster of activation was situated in the right insula–claustrum region.

**Tactile–visual minus tactile–tactile**

This subtraction image was hypothesized to reveal the areas engaged in the visual perception of form as well as in cross-modal matching. Fields of activation were found in the occipital and parietal cortex as well as in the thalamus and in the right insula–claustrum region (Table 3).

In the occipital cortex, the middle occipital gyri were activated bilaterally; other foci of activation were found in the left lingual gyrus extending on to the collateral sulcus and in the left posterior and anterior fusiform gyri.

The parietal cortex showed areas of activation in the left precuneus, in the right superior parietal gyrus, and in the cortex lining the intraparietal sulcus as well as in the right supramarginal and angular gyri.

The pulvinar nucleus of the thalamus was activated bilaterally, results of the activation in the different paradigms that involved the formation of the visual representation of the stimuli, as explained in Table 1. GOM, gyrus occipitalis medialis; GOS, gyrus occipitalis superior.

**Table 3. Fields of activation associated with visual representation of the stimuli**

<table>
<thead>
<tr>
<th>VV–control</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>Size (mm³)</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>Size (mm³)</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>Size (mm³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L lingual gyrus and collateral sulcus</td>
<td>12</td>
<td>−83</td>
<td>3</td>
<td>9266</td>
<td>−24</td>
<td>−77</td>
<td>−12</td>
<td>846</td>
<td>11</td>
<td>−75</td>
<td>−17</td>
<td>1312</td>
</tr>
<tr>
<td>L GOM</td>
<td>36</td>
<td>−83</td>
<td>6</td>
<td>1991</td>
<td>34</td>
<td>−79</td>
<td>2</td>
<td>881</td>
<td>21</td>
<td>−89</td>
<td>7</td>
<td>8973</td>
</tr>
<tr>
<td>R GOM</td>
<td>−29</td>
<td>−84</td>
<td>−4</td>
<td>1069</td>
<td>−40</td>
<td>−74</td>
<td>−11</td>
<td>937</td>
<td>−42</td>
<td>−72</td>
<td>−4</td>
<td>1078</td>
</tr>
<tr>
<td>R posterior fusiform</td>
<td>−13</td>
<td>−94</td>
<td>−2</td>
<td>978</td>
<td>L cuneus</td>
<td>3</td>
<td>−81</td>
<td>5</td>
<td>5932</td>
<td>L precuneus</td>
<td>18</td>
<td>−62</td>
</tr>
<tr>
<td>R LPS and intraparietal sulcus</td>
<td>−25</td>
<td>−73</td>
<td>41</td>
<td>976</td>
<td>−26</td>
<td>−69</td>
<td>42</td>
<td>963</td>
<td>−13</td>
<td>−68</td>
<td>5</td>
<td>663</td>
</tr>
</tbody>
</table>
| R angular gyrus | −42 | −65 | 40  | 532        | R LPS | 3  | −81 | 5   | 5932       | L angu
dar gyrus | −37 | −47 | 53  | 878        |
| L + R anterior cingulate | 0   | 30  | 30  | 1494       | L + R posterior cingulate | 0  | −40 | 32  | 614        | R supramarginal gyrus | −34 | −45 | 13  | 773        |
| L anterior insula | 32  | 19  | 6   | 751        | L + R posterior cingulate | 0  | −40 | 32  | 614        | R pulvinar | −12 | −30 | 13  | 1133       |
| R anterior prefrontal | −31 | 33  | 25  | 831        | L anterior insula | 32  | 19  | 6   | 751        | L anterior prefrontal | −31 | −10 | 16  | 1465       |
| L anterior prefrontal | 30  | 50  | 5   | 573        | L + R anterior cingulate | 0  | 30  | 30  | 1494       | L cuneus | 3  | −81 | 5   | 5932       |
| L cerebellum | 42  | −66 | −35 | 723        | L precuneus | 18 | −62 | −3  | 1073       | L cerebellum | 42  | −66 | −35 | 723        |

**Table 4. Overlap of the fields of activation associated with somatosensory representation of the stimuli**

<table>
<thead>
<tr>
<th>TT–control ∩ TV–VV</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>Size (mm³)</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>Size (mm³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L postcentral gyrus</td>
<td>30</td>
<td>−27</td>
<td>58</td>
<td>506</td>
<td>L lingual gyrus and collateral sulcus</td>
<td>10</td>
<td>−78</td>
<td>−22</td>
</tr>
<tr>
<td>L LPS</td>
<td>44</td>
<td>−31</td>
<td>47</td>
<td>942</td>
<td>R lingual gyrus and collateral sulcus</td>
<td>−12</td>
<td>−89</td>
<td>−9</td>
</tr>
<tr>
<td>R cerebellum</td>
<td>−17</td>
<td>−47</td>
<td>−19</td>
<td>665</td>
<td>−5</td>
<td>−95</td>
<td>−8</td>
<td>109</td>
</tr>
<tr>
<td></td>
<td>−18</td>
<td>−52</td>
<td>−48</td>
<td>301</td>
<td>L GOM</td>
<td>20</td>
<td>−94</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>32</td>
<td>−83</td>
<td>5</td>
<td>609</td>
</tr>
</tbody>
</table>

**Table 5. Overlap of the fields of activation associated with visual representation of the stimuli**

<table>
<thead>
<tr>
<th>VV–control ∩ TV–TT</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>Size (mm³)</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>Size (mm³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L lingual gyrus and collateral sulcus</td>
<td>10</td>
<td>−78</td>
<td>−22</td>
<td>344</td>
<td>R lingual gyrus and collateral sulcus</td>
<td>−12</td>
<td>−89</td>
<td>−9</td>
</tr>
<tr>
<td>R lingual gyrus and collateral sulcus</td>
<td>−5</td>
<td>−95</td>
<td>−8</td>
<td>109</td>
<td>L GOM</td>
<td>20</td>
<td>−94</td>
<td>6</td>
</tr>
<tr>
<td>R GOS</td>
<td>48</td>
<td>−71</td>
<td>−3</td>
<td>193</td>
<td>16</td>
<td>−92</td>
<td>21</td>
<td>238</td>
</tr>
<tr>
<td>R fusiformis</td>
<td>−42</td>
<td>−73</td>
<td>−15</td>
<td>95</td>
<td>L fusiformis</td>
<td>49</td>
<td>−68</td>
<td>−15</td>
</tr>
<tr>
<td>L cuneus</td>
<td>11</td>
<td>−70</td>
<td>11</td>
<td>198</td>
<td>L cuneus</td>
<td>11</td>
<td>−70</td>
<td>11</td>
</tr>
</tbody>
</table>
and a cluster of activation was found in the left hippocampus. At a more liberal level of significance \((p < 0.6)\) a small cluster \((r = 6.3)\) was found in the left superior colliculus (Talairach and Tournoux, 1988; coordinate 6, \(-30, 2\)). Fields of activation were found in the left cerebellum.

A cluster of significant activation was found in the right insula–claustrum region, with a center of gravity situated in the claustrum.

**TT–control \(\cap\) TV–VV**

This intersection of two independent cluster images was supposed to reveal the areas specifically activated in the formation of the tactile representation of the stimuli (Fig. 5). A cluster of activation was found in the left parietal lobe, on the left postcentral gyrus, which was extending into the cortex lining the postcentral sulcus and the anterior intraparietal sulcus, and in the posterior parietal cortex. A second one was situated more anteriorly on the postcentral gyrus, extending to the cortex lining the central sulcus (Table 4).

Two clusters of activation were found in the right cerebellum. That two activated fields, \(X_{TV},A\) and \(X_{TV},B\), originating from the respective cluster images \(A\) and \(B\), reflect activity in approximately the same synaptic cortical field can be addressed in a forward way. Let the estimated volumes and centers of gravities of a cluster in TT–control and a cluster in TV–VV be \(V_{i,TT}\) and \(V_{i,TV–VV}\), and \(TT\) and \(TV–VV\), respectively. In this case, \(V_{TT} < V_{TV–VV}\) (Table 2). A reasonable criterion of judging whether two clusters reflect activation in the same location is that the overlap, \(Q\), produced by the two clusters is equal to or greater than half of the volume of the smallest of the two clusters, and the centers of gravity of the two clusters are included in the overlap (Ledberg et al., 1995), i.e.:

\[
TT\cap TV–VV = Q, V_{0.5}TT \subseteq Q \text{ and } V–VV \subseteq Q.
\]

This was fulfilled for all three intersections (overlaps) as seen by comparing Tables 2 and 4.

**VV–control \(\cap\) TV–TT**

This intersection of two independent cluster images was supposed to show the areas specifically engaged in the formation of the visual representation of the stimuli (Fig. 6). We found all the clusters of activation in the occipital lobes. (Table 5). The lingual gyri and collateral sulci were activated bilaterally, as were the fusiform gyri. Other foci of activation were found in the right hemisphere in the superior occipital gyrus and in the left hemisphere in the middle occipital gyrus, the left cuneus, and the cortex lining the parieto-occipital sulcus, and in the cortex lining the calcarine sulcus. Of these overlaps the centers of gravity and the volume of overlap in the middle occipital gyri produced by the two clusters was greater than half of the volume of the smallest of the two clusters, and the centers of gravity of the two clusters were included in the overlap.

**TV–TT \(\cap\) TV–VV**

This intersection image was supposed to show the areas engaged in cross-modal transfer of information (Fig. 7). Only one cluster of activation was found, which was situated in the right insula–claustrum, with a center of gravity situated toward the claustrum (Fig. 8, Table 6). Here also the centers of gravity of the two clusters were included in the overlap.

**DISCUSSION**

This experiment, designed for each subject with three runs of four different tasks, gave the opportunity to study, within the same group of subjects, different aspects of the somatosensory and visual processing of shape and to isolate the structures involved in cross-modal transfer of shape.

The motor measurements showed that the matching conditions were mutually balanced and balanced against the control condition for motor activity of the right hand. Accordingly, no activations appeared in any of the motor cortices in any of the subtractions.

In addition, the signal energies for the somatosensory and visual shape stimulation (Roland and Mortensen, 1987) were also balanced across the three matching conditions. Accordingly, we observed no changes in visual cortices in TV–VV and no changes in somatosensory cortices in TT–TV. One might argue that the attention in TV was divided between the somatosensory and the visual modalities, whereas it was allocated to the visual modality in VV and the somatosensory modality in TT. However, to balance the allocation of attention toward the somatosensory and the visual modality between conditions TV and TT and TV and VV, the number of matchings in TV were twice the number during TT and VV.

The performance of the subjects, reflected by their \(d'\) values, were balanced across matching conditions. This made it unlikely that differences between conditions could be interpreted as differences in attention and/or task difficulty. The attentional effects or rCBF may vary between TT and VV. Cross-modal attentional effects in visual tasks tend to decrease the rCBF in somatosensory areas, and somatosensory tasks tend to decrease rCBF in visual areas (Haxby et al., 1994; Kawashima et al., 1995). It is not unlikely that the greater pace of the TV compared with the TT condition and the VV condition might imply a higher rate of switching attention from the tactile to the visual modality and vice versa. This could be a possible explanation for the pulvinar activation seen in TT–TV and TV–VV (Petersen et al., 1985, 1987).

Tables 2 and 4 show the fields activated every time the subjects perceived the ellipsoids tactually. Fields specifically engaged in the haptic processing of the ellipsoids were found by computing the intersection of the cluster images of TT–control and TV–VV. They were located to the contralateral postcentral gyrus, superior parietal lobule, and the cortex lining the anterior part of the intraparietal sulcus. The cortex lining the postcentral sulcus, anterior part of the intraparietal sulcus, and the cortex of the anterior part of the superior parietal lobule have in other studies been activated specifically during haptic processing of shape and length of objects (Roland and Larsen, 1976; Seitz et al., 1991; O’Sullivan et al., 1994; Roland et al., 1996). Together with the

---

### Table 6. Overlap of the fields associated with cross-modal transfer of information

<table>
<thead>
<tr>
<th>Intersection</th>
<th>(x)</th>
<th>(y)</th>
<th>(z)</th>
<th>Size (mm(^3))</th>
</tr>
</thead>
<tbody>
<tr>
<td>TV–TT (\cap) TV–VV</td>
<td>9</td>
<td>271</td>
<td>13</td>
<td>271</td>
</tr>
</tbody>
</table>

This is the result of the overlap between the two paradigms involving cross-modal transfer of information. Only one area was found to be constantly activated in both tasks, which was situated in the right insula–claustrum region.

\[ TT–control \cap TT–control \cap VV–control \]
present results, this strongly indicates that neurons in these regions are engaged in the formation of the haptic shape. We use the expression haptic processing of shape to note that although the experimental design attempted to balance allocation of attention between the somatosensory and visual modalities, we cannot exclude that in the intersection TT–control \( \cap \) TV–VV, the effect of somatosensory attention to shape might not be separable from the processing of shape information.

Structures involved in visual processing of the shape of the ellipsoids, revealed in VV–control, TV–control, and TV–TT, showed similar patterns of activation of the visual cortex. Fields solely engaged in visual processing of the shapes of the objects, and to some extent attention to visual object shape, were isolated in the intersection of the two independent cluster images of VV–control with TV–TT, which revealed several areas in the occipital cortex. All areas—lips of the calcarine sulcus, cuneus,
lingual gyrus, and the cortex lining the collateral sulcus, fusiform gyrus, and the middle occipital gyrus—have been described with different methods as being visual areas (Clarke and Miklossy, 1990; Zilles and Schleicher, 1993; Clarke, 1994a,b; Hadjikhani et al., 1994; Hadjikhani, 1995; Clarke et al., 1995; Sereno et al., 1995; Van Essen et al., 1995a,b). Of these regions, the lingual, fusiform, and occipital gyri have been activated by perception or discrimination of visual form and geometrical patterns (Gulyas and Roland, 1994; Gulyas et al., 1994; Roland and Gulyas, 1995).

We did not find any polymodal areas, i.e., fields of activation present consistently in tactile–tactile, tactile–visual, and visual–visual matching versus control. But we found an area consistently activated in the two subtractions of task involving cross-modal...
transfer, i.e., TV–TT and TV–VV, that was situated in the right insula–claustrum.

The neural structures participating in cross-modal transfer have been for a long time matters of debate, because lesion studies were never able to point to a particular structure consistently involved when cross-modal deficits were present. With notable exceptions (Ettlinger and Wilson, 1990), cross-modal research has been generally based on the assumption that there must be a special process to deal with the confluence of different sensory input in “polysensory convergence areas” (Pandya and Kuypers, 1969; Jones and Powell, 1970; Petrides and Iversen, 1976). The earliest attempts to study the effects of brain lesions on cross-modal performance used the cross-modal recognition method of Cowey and Weiskrantz (1975). Sahgal et al. (1975) and Petrides and Iversen (1976) reported impairment in cross-modal (tactile–visual) matching abilities after posterior temporal and prestriate removal and after lesions of the arcuate sulcus cortex. In more recent studies, authors have used a different cross-modal recognition paradigm (Jarvis and Ettlinger, 1977) in which monkeys learn cross-modality (vision or touch) discrimination tasks. Cortical lesions involving the superior temporal sulcus and the lateral prefrontal region in monkeys did not produce deficits (Ettlinger and Garcha, 1980). Streicher and Ettlinger (1987) examined cross-modal performance for entirely new and unfamiliar objects. Lesions in the frontal, temporal, and parietal cortex gave rise to impairment of cross-modal recognition of unfamiliar objects, despite normal performance on familiar objects.

Thus lesions to the cortex claimed as polymodal, e.g., the cortex lining the superior temporal sulcus, the intraparietal sulcus, the amygdala, and the lateral prefrontal cortex, have failed to abolish cross-modal matching consistently and specifically (Cowey and Weiskrantz, 1975; Sahgal et al., 1975; Petrides and Iversen, 1976; Jarvis and Ettlinger, 1977; Ettlinger and Garcha, 1980; McNally et al., 1982; Murray and Mishkin, 1984; Streicher and Ettlinger, 1987; Nahm et al., 1993). With the exception of the prefrontal cortex, none of these areas was activated by TV, and neither was the superior colliculus, also claimed a polymodal structure (Stein et al., 1976). The amygdala and the cortex lining the superior temporal sulcus were not activated in cross-modal matching (i.e., TV–TT e TV–VV), even if the threshold was set quite liberally. Ettlinger and

Figure 7. Activation in the cross-modal paradigms. Shown are the areas of activation in the two different paradigms involving cross-modal transfer of information. On the left is the cluster image of TV–TT, which shows areas involved in the formation of the visual representation of the stimulus and in the cross-modal transfer of information (coronal slice, y = 8; horizontal slice, z = −9); on the right is the cluster image of TV–VV, which shows areas engaged in the formation of the tactile representation of the stimulus and in the cross-modal transfer of information (coronal slice, y = 8; horizontal slice, z = −13).
Wilson (1990) suggested an alternative model for the mechanism of cross-modal performance, claiming a so-called “leakage” between perceptual and memory systems. On the basis of a 2-deoxyglucose study in monkeys trained to a high level of cross-modal performance, they suggested that one pathway for such leakage may be through the ventral claustrum (Hörster et al., 1989).

Then how does the brain match visual shape with somatosensory shape? Is it possible that the cortical fields representing visual shape communicate with the cortical fields representing somatosensory shape? From studies in monkeys there seems to be no support for such anatomical arrangement. The possible candidates, areas 7a and 7b, are not interconnected to any significant extent (Cavada and Goldman-Rakic, 1989; Andersen et al., 1990; Neal et al., 1990), and we do not know where the homologs of areas 7a and 7b are in humans. Neurophysiological studies in monkeys support the notion of parallel processing of visual and tactile shape, because neurons in a TV task in the somatosensory cortex react only to the tactile components immediately and during the short delay, and neurons in the visual association cortex react only to the visual components (Maunsell et al., 1991; Zhou and Fuster, 1996).

The left lateral prefrontal cortex was activated in TT–control, in VV–control, and in TV–control as the only consistent example of activations in nonsomatosensory nonvisual cortex (Tables 3, 4). Although these functional fields were located near each other, they did not overlap. Thus the lateral prefrontal cortex may be engaged in matching of two stimuli, but stimuli originating from different modalities do not engage the same prefrontal functional field. However, that the active fields did not overlap does not exclude that they might communicate. Indeed, bilateral cooling of a larger lateral prefrontal region in monkeys reversibly hampers not only TV and VT cross-modal matchings but also TT matchings (Shindy et al., 1994).

Direct communication between cortical fields representing visual shape and cortical fields representing somatosensory shape or the communication of each of these sets of representative fields with a third common cortical field might not be necessary for the matching. However, if one wants to advocate a fully parallel processing of somesthesis and vision, it is difficult to envisage how the matching is actually achieved. On the basis of our results, a speculative solution is that the populations of neurons associated with somatosensory processing of shape synchronize their activity with the populations associated with the visual processing (Singer, 1995). If this is so, then communication must exist at least one location to facilitate this synchronization.

In previous experimental studies, short-term memory components have been a confounding factor. In our approach the TV did not contain any memory component, whereas TT and VV had short delay between the first and second object. Because we observed no changes or decreases in the claustrum–insula region in TT–control or in TV–control, the consistent activation of the claustrum–insula cannot be attributed to this. Although unlikely, it cannot be excluded that the claustrum–insula activation may be partly attributable to the fact that the number of matchings were twice those of TT and VV. Against this is the fact that the claustrum–insula was not active in TT–control and VV–control.

Other studies may support the idea of the involvement of the claustrum in cross-modal transfer of information. The claustrum is best developed in primates, cetaceans, and carnivores. Its size is roughly in proportion with cortical volume. The claustrum is connected with virtually all of the cerebral cortex. Its connections have been studied mostly in cats and nonhuman primates (Neal et al., 1986; Hinova-Palova et al., 1988; Hörster et al., 1989; Cortiglighia et al., 1991; Boussaoud et al., 1992; Clasca et al., 1992; Morecraft et al., 1992; Baizer et al., 1993; Steele and Weller, 1993; Tokuno and Tanji, 1993; Updyke, 1993; Webster et al., 1993) (for review, see Sherk, 1986). The different studies show that the claustrum receives and gives rise to direct cortical projections and that it contains maps of different sensory (visual, auditory, and somatosensory) and motor systems.

A recent study of projection by retrograde labeling from claustrum to S1 and V1 done by Minciacci et al. (1995) in the cat shows a clear topographic organization, composed of two parts. In the somatosensory claustrum, there is a progression of cells...
projecting to hindpaw, forepaw, and face representation. The visual claustrum has a retinotopical organization, and claustral neurons project in a retinotopical manner to corresponding parts of V1. A second pattern of claustrum projections is composed of neurons distributed diffusely throughout the nucleus. In both somatosensory and visual claustrum, they intermingle with the topographically projecting cells.

In the monkey, Webster et al. (1993) demonstrated that portions of the claustrum connected with TEO and TE appear to overlap portions connected with other cortical areas, including V1, V2, V4, MT, MST, inferior prefrontal cortex, frontal eye fields, and posterior parietal cortex. Tokuno et al. (1993) showed reciprocal connections between the primary motor area in the monkey and the claustrum, and Baizer et al. (1993) demonstrated in the monkey that cells in the claustrum project both to temporal and parietal cortex, and that there are two representations of face and hand.

**Conclusion**

We found the insula–claustrum consistently active only when somatosensory shape representations were compared with visual shape representations, whereas we did not find any polymodal areas active during the processing of somatosensory as well as visual shape information. This and other studies support the involvement of the insula–claustrum in cross-modal transfer of information. The claustrum may be a site of organized and direct interaction between modality-specific areas. Because only the somatosensory areas were specifically active in the formation of the visual representation of shape, we propose here that, instead of being based on modality-nonspecific representations in polysensory areas, cross-modal transfer takes place between modality-specific areas, and that those modality-specific areas can communicate via the claustrum. This does not, however, exclude that communications for the purpose of matching may exist at other locations, for example, in the prefrontal cortex. Because the claustrum is a small nucleus difficult to distinguish from the insula with PET, more studies are needed with more sensitive techniques to confirm our hypothesis that the claustrum plays a crucial role in cross-modal transfer of information.

**REFERENCES**


Hadjikhani and Roland • Role of the Claustrum in Cross-Modal Transfer in Human


