Neural substrates for dividing and focusing attention between simultaneous auditory and visual events

Jennifer Adrienne Johnson* and Robert J. Zatorre

Montreal Neurological Institute, McGill University, 3801 University St., Montreal, QC H3A 2B4, Canada

Received 10 November 2005; revised 24 January 2006; accepted 16 February 2006

With information constantly bombarding the human sensory systems, how is it that we attend to behaviorally relevant information? The present study examined the behavioral and neural bases of attending to one sense while ignoring another sense (bimodal selective attention) contrasted with attending simultaneously to two senses (bimodal divided attention). In a functional magnetic resonance imaging (fMRI) study, subjects simultaneously heard novel melodies and viewed geometric shapes. They were instructed to actively attend to only one or to both senses (selective or divided attention) or as a baseline condition, passively observe both (bimodal passive). Memory tests for both attended and unattended stimuli validated that subjects were following the attention instructions. Selective attention led to increased activity in relevant sensory cortices while simultaneously leading to decreased activity in irrelevant sensory cortices. The divided attention instruction did not lead to a global increase in sensory cortex activity compared to the bimodal passive baseline condition. However, divided attention did recruit heteromodal areas in the left dorsolateral prefrontal cortex, while selective attention did not recruit any frontal areas. We propose that sustained selective and divided bimodal attention were achieved via distinct neural processes. Selective attention was achieved primarily via modulation of the sensory cortices. Divided attention was achieved for most individuals via recruitment of the middle-dorsolateral prefrontal cortex. However, there was also a trade-off between activity in posterior-dorsolateral prefrontal cortex and sensory regions, such that individuals who demonstrated the best performance during divided attention also showed the greatest recruitment of sensory cortices.

© 2006 Elsevier Inc. All rights reserved.

Introduction

Attention, a cognitive mechanism for dealing with the limited processing capacity of the brain (Pashler, 1998), is used, in part, to sort out behaviorally relevant information from the flood of events reaching all senses simultaneously. Researchers have investigated the neural mechanisms of within-modality attention in the visual (Desimone and Duncan, 1995; Corbetta and Shulman, 2002; Gazzaley et al., 2005), auditory (Alho et al., 1999; Jancke and Shah, 2002), and tactile (Craig and Rollman, 1999) modalities separately. However, we know far less about the neural processes underlying attention to multimodal information. In this fMRI study, we explored the influence of selective and divided attention on the perception and encoding of bimodal information.

In a previous study, we demonstrated that sustained bimodal selective attention leads to crossmodal inhibition between the sensory cortices (Johnson and Zatorre, 2005), in accord with other findings (Haxby et al., 1994; Laurienti et al., 2002; Shomstein and Yantis, 2004). Specifically, we found an enhancement of activity in attended sensory cortices coupled with a decrease in activity in unattended sensory cortices. But how do the sensory cortices react during bimodal divided attention? Whether or not attentional capacity is limited between modalities is widely contested (Pashler, 1990; Duncan et al., 1997; Jolicoeur, 1999); some find dividing attention between modalities leads to a limitation in the recruitment of sensory cortices (Klingberg, 1998; Loose et al., 2003); others do not (Herath et al., 2001; Szameitat et al., 2002). Our previous finding of crossmodal inhibition during selective attention suggests that neural resources are stringently allocated across sensory processing areas. Therefore, we predict that activity in sensory cortices during divided attention will be similarly limited, and this limitation will lead to a behavioral deficit. Moreover, we predict that distinct heteromodal cortices will subserve the two cognitive processes, as has been demonstrated during within-modality selective and divided attention (Corbetta et al., 1991).

Although most bimodal divided attention studies utilize dual-task paradigms requiring subjects to make two motor responses simultaneously, an output response is not required of everyday divided attention. Moreover, motor responses often lead to response interference. Therefore, we adapted the previously used paradigm (Johnson and Zatorre, 2005) to be devoid of this feature. Subjects in an fMRI scanner simultaneously heard novel melodies and viewed abstract shapes with instructions to attend to either the melodies or shapes (selective attention), attend to both modalities (divided attention), or as a baseline, passively observe both (bimodal passive). In contrast to our previous protocol, subjects were not required to perform a continuous response task to...
features of attended stimuli. We measured subjects’ ability to follow our attention instructions by implementing a forced-choice recognition memory task for both attended and ignored information with the prediction that attended information would be remembered better than ignored. We know of no other paradigm that includes dependent measures for both attended and ignored information. Furthermore, while most attention studies have used simple (Loose et al., 2003), or verbal (Shomstein and Yantis, 2004) stimuli, we are interested in the processing of complex, novel, and not easily verbalizable stimuli.

Materials and methods

Subjects

Twelve healthy, right-handed English-speaking volunteers (7 male; 20 to 33 years of age, mean 26 years) with normal hearing and corrected-to-normal vision participated in the experiment after giving written informed consent for a protocol approved by the Montreal Neurological Institute Ethics Review Board. Because nearly all university students have some degree of musical training, we attempted to model the normal distribution of musical experience in our subject pool by including subjects with a variety of musical backgrounds (1 to 16 years of musical training, mean 7.4 years).

Stimuli

Auditory stimuli were 7-s novel tonal melodies presented binaurally at a level of about 80-dB sound pressure level via Siemens MR-compatible pneumatic sound transmission headphones. Visual stimuli were novel, abstract, closed, two-dimensional shapes comprising 14 black line segments, each of which appeared sequentially to form the shape, “drawn” on a white background. The shapes were abstract in that they were created to not resemble any easily verbalizable known objects. They were displayed on a screen at the foot of the scanner bed using a projector and were viewed through an angled mirror placed on the head coil. When presented simultaneously, the auditory and visual stimuli started and stopped at exactly the same time but the individual elements of the two stimuli never synchronized. Because the melodies and shapes were not meant to be integrated into a unitary percept, we created stimuli devoid of overlapping properties and concurrent temporal changes. For more information about the stimuli used in this experiment, see Johnson and Zatorre (2005).

Procedure

During a 10-s trial, instructions were displayed on the screen corresponding to four conditions: attend to melodies and ignore shapes (auditory selective), attend to shapes and ignore melodies (visual selective), attend to both melodies and shapes (bimodal divided), or as a baseline, passively listen to melodies and view shapes (bimodal passive). The instruction for each condition was applied to the next five 10-s trials, termed the encoding phase. During the encoding phase, each 10-s trial comprised 500 ms of silence, then 7 s of a simultaneously presented melody and shape, followed by 630 ms of silence and finally 1870 ms of scan acquisition (Fig. 1). Following all encoding phases except the bimodal passive condition, there was an immediate 20-trial memory phase. Subjects were made aware that a memory test would follow all but the bimodal passive condition. The memory test was a forced-choice recognition test for the 5 previously encoded trials. In the memory test, five pairs (one target and one foil) of melodies and shapes were presented individually and subjects had to choose, via a mouse button press, which stimulus they had previously encountered in the encoding phase. For example, subjects viewed two shapes over 20 s and decided which one they remembered, then they heard two melodies, made a decision, viewed two shapes, made a decision, and so on. Each of the 20 trials of the memory phase was 10 s in duration, including 500 ms silence, 7-s stimulus presentation, 630 ms silence and 1870 ms scan acquisition. Each memory test was always followed by 9130 ms of silence with a blank, white screen and 1870 ms of scan acquisition (baseline acquisitions). The paired encoding-memory phases were repeated until all conditions were completed (Table 1). Then it was repeated twice more for a total of three runs for each subject. Order of stimuli and conditions were counterbalanced across subjects and runs. All conditions and runs included unique blocks of stimuli.

fMRI parameters and analysis

Scanning was carried out in a 1.5-T Siemens Sonata imager. First, a high-resolution T1-weighted anatomical scan was obtained (voxel size: \(1 \times 1 \times 1 \text{ mm}^3\), matrix size: \(256 \times 256\)), and then three series of 142 T2*-weighted gradient-echo planar images of blood–oxygen-level-dependent (BOLD) signal were acquired (head coil, TE = 50 ms, TR = 10 s, matrix size: \(64 \times 64 \times 22\), voxel size: \(5 \times 5 \times 5 \text{ mm}^3\)). The long inter-acquisition time minimized the effects of scanner noise on subjects’ ability to hear the auditory stimuli and avoided contaminating the BOLD signal response to the stimuli in the auditory cortices (Belin et al., 1999).

BOLD signal images were smoothed using a 12-mm FWHM Gaussian kernel and motion corrected to the third frame of each run using AFNI software (Cox, 1996). Then the data were statistically analyzed using a suite of matlab programs (fmristat available at http://www.math.mcgill.ca/keith/fmristat/). Each subjects’ run was fit to a linear model that accounted for stimulus conditions set up in a design matrix corresponding to each acquisition (only encoding phase trials were included), temporal drift, the hemodynamic response function, and temporally correlated errors (Worsley et al., 2001). This yielded the effects, standard deviations, and \(t\) statistics for each run and for each contrast. The three runs for each subject were then combined using the effects and standard deviations from the previous analysis. This step involved using a mixed effects model (smoothing, with a Gaussian filter, the ratio of the random effects variance divided by the fixed effects variance) in order to stabilize the variance estimate and increase the degrees of freedom. The combined functional runs of each subject were then transformed to MNI305 space (Collins et al., 1994) using a function derived from transforming each individual subjects’ anatomical scan to MNI305 space. In a final step, the standardized, combined runs of all 12 subjects were pooled, yielding group statistical \(t\) maps for each contrast. The
threshold for significance was set at $t(90) = 4.46$ (corrected for multiple comparisons $P < 0.05$) for a whole-brain search (250,000 8 mm$^3$ voxels).

Results from the aforementioned analysis raised further research questions that we sought to answer using a functional connectivity analysis. The purpose of the functional connectivity analysis was to find activity in one region of the brain that was correlated with activity in another region of the brain over the time course of the fMRI experiment. One voxel is chosen (referred to as a reference voxel), and the time course of activity in that voxel is correlated with the time course of activity in all other voxels in the brain. However, rather than searching all voxels in the brain that correlated with the time course of activity in the reference voxel, we prechose specific regions of interest that we expected to be functionally connected to the reference voxel. The reference voxels and regions-of-interest were chosen based on the results of the aforementioned subtraction contrasts. Functional connectivity between areas outside of the prechosen regions of interest was not assessed. Therefore, correcting for multiple comparisons was not necessary.

The functional connectivity analysis was performed using a general linear model:

$$Y_{ij} = R_i\beta_{ij} + C_i\beta_j + e$$

where $R$ refers to BOLD signal in the reference voxel, $C$ refers to confounds such as spatial and temporal drift, $i$ refers to the frame, and $j$ refers to the voxel. This was implemented with a suite of matlab programs called fmristat (see http://www.math.mcgill.ca/keith/fmristat/ for more details). First, the reference voxel (in MNI305 space) was converted to each subject’s native space. Then, after correcting for the time at which slices were acquired, the linear regression was solved for each run of each subject in native space. This step yielded a $t$ statistic map, with each voxel’s $t$ value corresponding to the strength of the linear relationship (or correlation) between BOLD in that voxel and BOLD in the reference voxel over the time course of each run. Next the three runs were combined and converted to back MNI305 space, and these files were then pooled across subjects (in the same manner as aforementioned).

Finally, voxel-of-interest analyses were performed by extracting BOLD signal values at specific tags of interest, which were to the frame, and $j$ refers to the voxel. This was implemented with a suite of matlab programs called fmristat (see http://www.math.mcgill.ca/keith/fmristat/ for more details). First, the reference voxel (in MNI305 space) was converted to each subject’s native space. Then, after correcting for the time at which slices were acquired, the linear regression was solved for each run of each subject in native space. This step yielded a $t$ statistic map, with each voxel’s $t$ value corresponding to the strength of the linear relationship (or correlation) between BOLD in that voxel and BOLD in the reference voxel over the time course of each run. Next the three runs were combined and converted to back MNI305 space, and these files were then pooled across subjects (in the same manner as aforementioned).

Finally, voxel-of-interest analyses were performed by extracting BOLD signal values at specific tags of interest, which were

| Stimuli Instruction Memory test |
|-------------------------------|-----------------|------------------|
| Bimodal auditory selective |
| Melodies and shapes | Attend to melodies and ignore shapes | Yes |
| Bimodal visual selective |
| Melodies and shapes | Attend to shapes and ignore melodies | Yes |
| Bimodal divided attention |
| Melodies and shapes | Attend to melodies and shapes | Yes |
| Bimodal passive baseline |
| Melodies and shapes | Passively observe melodies and shapes | No |
| Baseline condition |
| None | No instruction | No |

Table 1: Description of bimodal and baseline conditions

Fig. 1. Representation of fMRI protocol. Each 10s trial comprised 500 ms of silence, followed by 7 s of stimulus presentation, 630 ms of silence and 1870 ms of scan acquisition. Although the design is event-related, conditions are blocked into 5 trials of an encoding phase and 20 trials of a memory test, with baseline acquisitions interspersed.

Fig. 2. Recognition memory test performance for melodies and shapes in the three attention conditions. When attending to melodies, fewer shapes are remembered; when attending to shapes, fewer melodies are remembered. There is no significant decrement in performance from selective to divided attention conditions. Data reported as means and standard errors.
voxel coordinates in Talaraich/MNI305 space. A voxel, rather than region, of interest analysis was performed because with 12-mm spatial smoothing, activity in a single voxel provides a reasonable index of activity in a region. BOLD was extracted for each condition from the effect files generated using fmristat. Combining runs and subjects yielded the average percent change in BOLD signal associated with each condition compared to the silence/blank screen baseline acquisitions.

**Results**

**Behavioral results**

Percent of correctly remembered stimuli was compared across the attention conditions (Fig. 2). To compare the two selective attention and divided attention conditions, we performed a 2 by 3 repeated measures analysis of variance. The first factor was the

---

Fig. 3. Top portion (A) shows sensory cortices active during bimodal passive condition compared to silence/blank screen baseline. Color bar indicates level of significant activity (t values) at each voxel. Middle portion (B) shows the results of the functional connectivity analysis. When the right superior temporal sulcus (R STS) was used as a reference voxel, BOLD in the right superior parietal lobe (R SPL) was found to be negatively correlated with its time course. When the left SPL was used as a reference voxel, BOLD in posterior and anterior portions of the left superior temporal sulcus (pSTS and aSTS) were found to be negatively correlated with its time course. Bottom portion (C) shows the interaction of auditory (R STS, reference voxel from functional connectivity analysis) and visual (right SPL, found to be functionally connected to R STS) cortices during selective attention conditions. Also note the similarities in BOLD response in the bimodal passive and bimodal divided attention conditions. BOLD response reported as mean percent change from baseline condition with standard error. *P < 0.05.
sensory modality (auditory melodies or visual shapes), and the second factor was the attention condition (auditory selective, visual selective, divided). Overall, we found a main effect of attention condition \( (F(2, 22) = 7.1, P < 0.01) \) and an interaction effect \( (F(2, 22) = 21.7, P < 0.01) \). Planned comparisons showed a strong interaction between melody and shape memory in the selective attention conditions \( (F(1, 22) = 42.9, P < 0.01) \), such that attended information, regardless of modality, was remembered better than unattended information. Furthermore, there was no difference between melody memory in the auditory selective and bimodal divided attention conditions \( (F(1, 22) = 0.03, P > 0.05) \) nor was there a difference between shape memory in the visual selective and bimodal divided attention conditions \( (F(1, 22) = 3.1, P > 0.05) \).

To test for a ceiling effect, melody and shape memory in the divided attention conditions were separately compared to 100% using single-sample \( t \) tests. Both \( t \) tests showed memory performance to be significantly lower than 100% \( (P < 0.01) \).

fMRI results

Table 2
Selective attention peaks of activation

<table>
<thead>
<tr>
<th>Region of peak voxel</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>L BA 9 posterior-DLPFC</td>
<td>−46</td>
<td>36</td>
<td>38</td>
<td>4.5</td>
</tr>
<tr>
<td>L BA 8/9 posterior-DLPFC</td>
<td>−46</td>
<td>32</td>
<td>40</td>
<td>4.3</td>
</tr>
<tr>
<td>L BA 9/46 middle-DLPFC</td>
<td>−46</td>
<td>46</td>
<td>16</td>
<td>4.2</td>
</tr>
<tr>
<td>L BA 8 posterior-DLPFC</td>
<td>−46</td>
<td>24</td>
<td>42</td>
<td>4.2</td>
</tr>
</tbody>
</table>

Coordinates are reported in Talairach/MNI space (Talaraich and Tournoux, 1988). L, left; BA, Brodmann’s area; DLPFC, dorsolateral prefrontal cortex.

The contrast between auditory and visual selective attention conditions revealed that, on average, activity increased in relevant sensory cortices while simultaneously decreasing in irrelevant sensory cortices during selective attention. To elucidate the strength of this negative relationship between auditory and visual sensory cortices, we performed a functional connectivity analysis. We sought to determine if the time course of BOLD response in auditory cortices demonstrated a negative correlation with the time course of BOLD response in visual cortices within selective attention conditions. We combined the auditory and visual selective attention conditions into one functional connectivity analysis because (1) we expected the same negative correlation between sensory cortices over the time course of both conditions and (2) it increased the power of the analysis. Two reference voxels were chosen (one auditory and one visual) to assess functional connectivity between the modalities. The right upper bank of the STS \( (x = 48, y = −28, z = 2) \) and the left superior parietal lobe \( (x = −22, y = −66, z = 52) \) were chosen because they were the most significant peaks of activation in each of the selective attention contrasts (see Table 2). Furthermore, because one reference voxel was in the right hemisphere and the other was in the left, we were able to assess functional connectivity in each hemisphere. When testing the auditory reference voxel, functional connectivity was only assessed within visual regions; when testing the visual reference voxel, functional connectivity was only assessed within auditory regions.

The visual region showing the greatest negative correlation with the time course of BOLD response in the auditory reference voxel was the right SPL \( (x = 22, y = −58, z = 68; t = −4.7; P < 0.0001 \text{uncorrected for multiple comparisons}) \), while the auditory regions showing the greatest negative correlation with the time course of BOLD in the visual reference voxel were the left anterior superior temporal sulcus \( (x = −48, y = −2, z = −16; t = −5.3; P < 0.0001 \text{uncorrected for multiple comparisons}) \) and the left posterior superior temporal sulcus \( (x = −50, y = −32, z = −8; t = −4.7; P < 0.0001 \text{uncorrected for multiple comparisons}) \) (Fig. 3B). While we only examined functional connectivity between predicted regions, the predicted regions were actually the areas with the highest levels of significance. Additionally, although the results we report are from a functional connectivity analysis with auditory and visual selective attention conditions combined, we performed the analysis for both the auditory and visual conditions separately and found nearly identical results.

To further demonstrate the interaction of auditory and visual sensory cortices, a voxel-of-interest (VOI) analysis was performed. The voxels that we chose were the right superior temporal sulcus (auditory reference voxel from functional connectivity analysis) and the right SPL (outcome of the functional connectivity analysis). BOLD was extracted at these two voxels for each bimodal condition (Fig. 3C). A 2 (VOI) by 4 (condition) repeated measures analysis of variance shows the interaction effect is significant \( (F(2, 22) = 10.7, P < 0.01) \). Tukey’s honestly significant difference post hoc tests demonstrate that percent BOLD change from silence/blank screen baseline in right STS and right SPL is equivalent during bimodal passive and bimodal divided attention conditions \( (P > 0.05) \), but percent BOLD change is significantly different in right STS and right SPL during bimodal
auditory selective attention ($P < 0.05$) and bimodal visual selective attention conditions ($P < 0.05$). Fig. 3C demonstrates the similar responses of the sensory cortices in the bimodal passive and bimodal divided attention conditions. In fact, the contrast, bimodal divided minus bimodal passive, shows no significant changes in BOLD response in sensory cortices suggesting that the sensory cortices reacted similarly in the two conditions despite the attention instruction.

Although divided attention did not recruit sensory cortices to a greater extent than the bimodal passive condition, divided attention did additionally recruit areas within the dorsolateral prefrontal cortex (DLPFC) (Table 3). A statistically significant peak falls within a cluster of three left-lateralized activations that start in Brodmann’s area 8 and continue into Brodmann’s area 9 (Petrides and Pandya, 1999; Petrides, 2005). Activity in these three posterior peaks is positively correlated with each other ($r = 0.85$), suggesting that they are reacting similarly. Because of their close proximity, given 12-mm spatial smoothing, we will interpret these peaks as one region in the discussion (called the posterior-DLPFC). Another frontal-lobe peak of activity was noted 24 mm anterior and ventral to this cluster in Brodmann’s area 9/46 (called the middle-DLPFC). A voxel-of-interest analysis was performed on the middle-DLPFC peak (Brodmann’s area 9/46) and a peak in the posterior-DLPFC (Brodmann’s area 9). Fig. 4 clearly demonstrates these areas are active only during divided attention and not during all bimodal conditions of equal stimulus presentation (bimodal selective attention conditions and bimodal passive condition). In fact, in the contrasts examining differences between selective attention conditions and the bimodal passive condition, no statistically significant peaks were found in any heteromodal cortices (this remained true after lowering the threshold to $t = 3.0$ $P > 0.5$ and was true when comparing selective attention conditions to the silence/blank screen baseline). Furthermore, the functional connectivity analyses revealed no heteromodal areas functionally connected to auditory or visual sensory cortices during selective attention conditions. One-way repeated measures ANOVAs show an effect of condition within each prefrontal region (BA 9/46, $F(3,33) = 5.5$, $P < 0.01$; BA 9, $F(3,33) = 9.0$, $P < 0.01$). Post hoc single-sample $t$ tests show BOLD-signal is significantly different from zero in the middle-DLPFC during divided attention ($P = 0.001$, significant with Bonferroni correction), and not significantly different from zero during other bimodal conditions ($P > 0.4$). In the posterior-DLPFC, single-sample $t$ tests show BOLD-signal during divided attention is significantly different from zero ($P = 0.003$, significant with Bonferroni correction), BOLD-signal during bimodal selective attention is not significantly different from zero ($P > 0.8$), and BOLD-signal during the bimodal passive condition is significantly different from zero ($P = 0.009$, significant with Bonferroni correction).
To elucidate the behavioral function of these prefrontal cortices, we correlated each subjects’ memory performance (average of melody and shape percent correct in the divided attention condition) with BOLD at the middle-DLPFC and posterior-DLPFC. The results showed no correlation between BOLD and behavior in the middle-DLPFC (Brodmann’s area 9/46; \( r = 0.12, P > 0.05 \)), but a significant negative correlation in the posterior-DLPFC (Brodmann’s area 9; \( r = -0.55, P < 0.05 \)). Fig. 5 shows the negative correlation between BOLD and memory performance in the posterior-DLPFC. Fig. 5 also shows that there is a positive correlation between BOLD in the sensory cortices (average of percent BOLD change in R STS and R SPL) and averaged shape and melody memory performance (\( r = 0.8, P < 0.05 \)).

Discussion

Selective attention

When subjects focused attention on one modality while ignoring another, they remembered more of the attended modality than the unattended. This was accompanied by increased activity in sensory cortices subserving the attended modality and decreased activity in ignored sensory cortices. This is not only a replication of previous work (Johnson and Zatorre, 2005) but an extension of it. In our previous study, participants performed a continuous response motor task to features of the attended stimuli while encoding them. It is possible that some aspect of performing that task resulted in the behavior and activation patterns previously observed. Our current study did not include a continuous motor task, and we observed the same behavior and activation patterns. This demonstrates that directing attention to a modality requires neither an explicit judgment nor a motor response. Rather, an attention instruction alone can alter memory for bimodal stimuli and can modulate sensory cortex activity.

Compared to bimodal passive and silence/blank screen baselines, bimodal selective attention did not recruit additional heteromodal cortices. Other bimodal selective attention studies have demonstrated the involvement of heteromodal cortices, such as posterior parietal and superior prefrontal cortices, in voluntary shifts between auditory and visual modalities (Shomstein and Yantis, 2004), and a frontal–parietal–cingulate network for involuntary shifts of attention between modalities to salient events (Downar et al., 2000). However, our paradigm does not examine quick shifts of attention between modalities; rather, we investigated sustained within-modality attention. Initial shifts of attention between modalities may recruit heteromodal cortices. Here, we found no strong evidence of the involvement of heteromodal cortices during sustained attention to one modality while ignoring another. Rather we hypothesize that it is possible that bimodal selective attention can be achieved via crossmodal inhibition between the sensory cortices, with little involvement of heteromodal cortices at least under some circumstances.

Crossmodal inhibition, or increased activity in one sensory area with simultaneous decreased activity in another, has been demonstrated with presentation of unimodal stimuli (Shulman et al., 1997; Zatorre et al., 1999; Laurienti et al., 2002), as well as bimodal stimuli (Shomstein and Yantis, 2004; Johnson and Zatorre, 2005). The utility of such a mechanism for allocating neural resources between sensory modalities seems obvious: inhibiting activity in nonrelevant sensory cortices frees additional resources for use by relevant cortices. It is still unclear how crossmodal inhibition is achieved. The prefrontal cortex has been implicated to play a top-down role in enhancing the processing of relevant information and inhibiting the processing of irrelevant information in posterior sensory areas (Chao and Knight, 1998; Miller and Cohen, 2001). However, as mentioned, we observed no prefrontal cortex activity during our bimodal selective attention task. Another possible mechanism for achieving crossmodal inhibition is through direct reciprocal connections between the sensory cortices. This hypothesis is supported by recent findings of direct projections between auditory and visual sensory cortices in nonhuman primates (Falchier et al., 2002; Rockland and Ojima, 2003). Furthermore, our finding of functional connectivity between visual and auditory regions, while not direct evidence in support of crosstalk between the sensory cortices, also supports the possibility.

Divided attention

Contrary to our hypothesis, when dividing attention between auditory and visual stimuli, subjects remembered as many of the attended stimuli as during bimodal selective attention. This cannot be attributed to a “ceiling effect”, as memory for both melodies and shapes was significantly lower than 100% during divided attention, and there was a wide range of performance from 73% to 100% as demonstrated in Fig. 5. This finding suggests that splitting attention between modalities does not necessarily lead to a memory decrement compared to selectively attending to one modality and ignoring another. We were surprised by this result, considering others have found behavioral deficits comparing dual-task to single-task performance (Bonnel and Hafer, 1998; Jolicoeur, 1999; Spence and Read, 2003), including divided attention during encoding (Craik et al., 1995; Iidaka et al., 2000). However, studies that demonstrate dual-task decrements tend to require subjects to perform manual responses during divided attention. As decrements are often attributed to two or more tasks competing for similar neural substrates (e.g., dual-response tasks overlap in motor demands and/or recruitment of motor cortices) (Klingberg, 1998), it may not be surprising that in our study, which does not require motor responses, the capacity to perceive and remember two simultaneous stimuli presented in separate modalities is possible without a behavioral decrement.

With no behavioral deficit during divided attention, one could hypothesize that sensory cortex activity during divided attention would be equal to the summed activity in attended sensory cortices during the two selective attention conditions. However, sensory cortices were significantly less active during divided than selective attention, suggesting a neural resource limitation between modalities, as others have found (Klingberg, 1998; Loose et al., 2003). This interpretation is further supported by our finding of equivalent sensory cortex activity during divided attention and the bimodal passive baseline. In other words, the divided attention instruction did not lead to increased sensory cortex activity above that of just passively observing the bimodal stimuli, suggesting both conditions saturated between-modality resources. We did not probe memory following the bimodal passive condition out of concern that subjects would then attempt to remember stimuli during subsequent passive conditions, making the condition no longer “passive.” Nonetheless, we deemed the bimodal passive baseline an improvement over a pure, no-stimulus baseline which lacks any control over subjects’ cognitive activity. While one could conjecture that equivalent
sensory cortex activity in divided and passive conditions was due to subjects treating the two conditions identically, our findings of frontal contributions to divided attention refute this interpretation.

**Frontal contributions to divided attention**

Because the extent to which sensory cortices could be recruited during divided attention was limited, one would expect heteromodal cortical regions to be recruited in order to maintain task performance equivalent to that during selective attention. We found that only bimodal divided attention, and not selective or passive conditions, recruited additional heteromodal prefrontal cortices. So, although sensory cortices reacted similarly in bimodal passive and divided conditions, the recruitment of prefrontal cortices during divided attention further supports the conclusion that subjects were not treating the conditions equally.

The prefrontal cortices recruited during divided attention were primarily left-lateralized and in the DLPFC. This result is consistent with findings from patients with left-lateralized lesions of the superior PFC who demonstrated grave behavioral deficits during divided and focused attention (Godefroy and Rousseau, 1996). Additionally, the finding of left-lateralized PFC activity during encoding corroborates previous findings showing recruitment of this region when encoding recollective memories (Kensinger et al., 2003), countering the interpretation that subjects implicitly encoded during divided attention, compared to explicitly encoding during selective attention. Also, left-lateralized PFC activity, when comparing divided to selective attention, corroborates findings of Loose et al. (2003), although they found activity more ventrally, in the inferior frontal gyrus (IFG). However, in that study IFG was simply less active during selective than divided attention, whereas we find no left DLPFC activity during all other bimodal conditions. Loose and colleagues findings possibly resulted from the compounded effect of dual-task performance, rather than recruitment of an additional area for dividing attention, as we found.

We consider recruitment of the left DLPFC as compensating for the lack of global increased activity in sensory cortices during divided attention in order to maintain performance equivalent to that during selective attention. Most agree the DLPFC subserves executive functions, particularly working memory (Curtis and D’Esposito, 2004). However, some assert the DLPFC is specifically involved in maintenance of spatial information in working memory (Levy and Goldman-Rakic, 2000), while others emphasize its role in manipulating information in working memory (Petrides, 2000). Our findings fit the latter model, as spatial information remained consistent between all conditions, whereas divided attention required handling multiple independent inputs in working memory compared to the other conditions.

To further elucidate the functional role of sub-regions of the DLPFC active during divided attention, we correlated middle-DLPFC and posterior-DLPFC activity with divided attention performance. While middle-DLPFC (Brodmann’s area 9/46) activity was not correlated with performance, posterior-DLPFC (particularly Brodmann’s area 9) activity was negatively correlated with performance. As all subjects recruited middle-DLPFC equally, this area likely subserved the handling of multiple components in working memory. The finding of a negative correlation between activity in posterior-DLPFC and performance contrasts with the notion that increased PFC activity leads to increased performance; yet others have demonstrated similar negative relationships between performance and PFC activity (Rypma and D’Esposito, 1999; Herath et al., 2001; Szameit et al., 2002). Compared to the middle-DLPFC’s role in subserving working memory, the posterior-DLPFC has been implicated in modulating the processing of perceptual information in posterior sensory cortices (Milham et al., 2003), which fits well with the present findings. We found that during divided attention, posterior-DLPFC activity demonstrated an interaction with sensory cortices as a function of performance (Fig. 5). Subjects with poor divided attention performance (particularly the bottom two performers) recruited posterior-DLPFC more and sensory cortices less; good performers (particularly the top three performers) demonstrated the opposite trend. One interpretation of these findings is that good performers were able to recruit more of the sensory cortices to perform the divided attention task and thereby did not need to recruit frontal cortices. Poor performers, on the other hand, did not seem able to recruit additional sensory cortices for the task and therefore may have recruited the PFC in an attempt to improve performance. These results demonstrate that individual differences in divided attention performance can at least in part be explained by functional interactions between prefrontal and sensory regions of the brain.

**Acknowledgments**

We thank Marc Bouffard for his technical expertise, Keith Worseley for sharing his knowledge of statistics, and the personnel of the McConnell Brain Imaging Center for their assistance. Supported by a grant from the Canadian Institutes for Health Research.

**References**


