

Cross-Modal Processing in Early Visual and Auditory Cortices depends on Expected Statistical Relationship of Multisensory Information

Bernhard Baier,^{1,2} Andreas Kleinschmidt,^{1,3} and Notger G. Müller¹

¹Cognitive Neurology Unit and Brain Imaging Center, Clinic for Neurology, Johann Wolfgang Goethe-University, 60590 Frankfurt, Germany, ²Department of Neurology, Johannes-Gutenberg-University, Mainz, 55101 Mainz, Germany, and ³Institut National de la Santé et de la Recherche Médicale, Unité 562, Service Hospitalier Frédéric Joliot, Commissariat à l'Énergie Atomique, 91401 Orsay, France

Previous studies have shown that processing information in one sensory modality can either be enhanced or attenuated by concurrent stimulation of another modality. Here, we reconcile these apparently contradictory results by showing that the sign of cross-modal interactions depends on whether the content of two modalities is associated or not. When concurrently presented auditory and visual stimuli are paired by chance, cue-induced preparatory neural activity is strongly enhanced in the task-relevant sensory system and suppressed in the irrelevant system. Conversely, when information in the two modalities is reliably associated, activity is enhanced in both systems regardless of which modality is task relevant. Our findings illustrate an ecologically optimal flexibility of the neural mechanisms that govern multisensory processing: facilitation occurs when integration is expected, and suppression occurs when distraction is expected. Because thalamic structures were more active when the senses needed to operate separately, we propose them to serve gatekeeper functions in early cross-modal interactions.

Key words: cross-modal perception; visual cue; auditory cue; fMRI; early visual cortex; early auditory cortex; human

Introduction

It has long been assumed that the early sensory processing chains in the brain are unimodal and operate primarily independently from each other with activity in one system exerting little if any influence on processing in another. This view has been challenged by the demonstration of cross-modal effects on perception: a sound is misallocated toward the semantically matching visual stimulus, auditory speech perception can be modulated by lip reading, and effects of intermodal attention in visual and auditory event-related potentials showed that intermodal attention operates by a selective modulation of modality-specific stimulus-driven responses (Eimer and Schröger, 1998). Furthermore, it was shown that spatially nonpredictive peripheral cues (e.g., a sound) could attract covert visual attention to specific locations and that sounds were misallocated at their apparent visual source (Driver and Spence, 1998).

These behavioral observations have been complemented recently by neuroimaging studies showing that activity in one sensory system can be altered by input to another (Macaluso et al., 2000; McDonald et al., 2000; Laurienti et al., 2002; Weissman et al., 2004; Johnson and Zatorre, 2005). The level at which these cross-modal interactions occur seems to depend on the categor-

ical quality of the stimuli used, but it is far less clear which parameters control the sign of modulation, i.e., enhancement versus suppression. For example, Macaluso et al. (2000) demonstrated that tactile stimulation enhanced activity in the visual cortex, but only when it was administered to the same side as the visual target. Conversely, Laurienti et al. (2002) found that activity in visual cortex was reduced when subjects listened to sounds, whereas activity in auditory cortex was reduced during visual processing. Hence, cross-modal interactions can be both mutually suppressive as well as facilitating. Here, we provide evidence that nature and sign of cross-modal neural interaction depend not only on congruence but also on whether sensory inputs are expected to convey associated or unrelated and thus potentially conflicting information, i.e., whether stimulus processing in one modality will be helped or disturbed by taking information in another modality into account.

We designed an experiment in which subjects had to differentiate either between two tones or between two visual objects. Visual and auditory stimuli were presented simultaneously in each trial, but a preceding cue indicated which modality subjects should base their response on. Although the overall sensory input across the two conditions was identical, the crucial manipulation concerned the overall statistical relationship between auditory and visual targets. In the “associated” condition, tones with a certain pitch were almost always paired with pictures of a bar with a certain tilt, whereas in the “non-associated” condition, tones and bars were combined at random. Thus, in the associated condition, discriminating stimuli in the cued modality could be fa-

Received Nov. 27, 2005; revised Oct. 16, 2006; accepted Oct. 17, 2006.

Correspondence should be addressed to Notger Müller, Cognitive Neurology Unit and Brain Imaging Center, Clinic for Neurology, Johann Wolfgang Goethe-University, Theodor-Stern-Kai 7, 60590 Frankfurt, Germany. E-mail: n.mueller@em.uni-frankfurt.de.

DOI:10.1523/JNEUROSCI.1457-06.2006

Copyright © 2006 Society for Neuroscience 0270-6474/06/2612260-06\$15.00/0

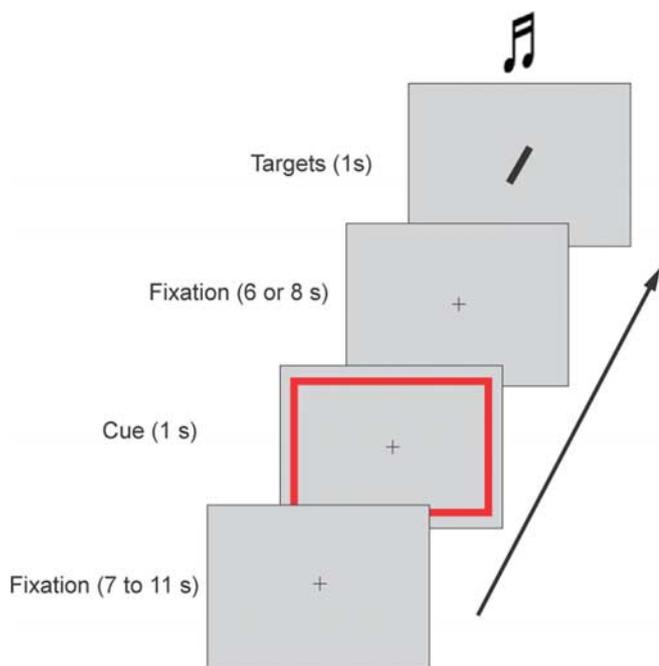


Figure 1. Experimental paradigm. A blue or red peripheral frame served as cue and instructed subjects whether the auditory (red frame) or the visual (blue frame) target stimulus would be relevant.

cilitated by processing the item in the other modality, whereas in the non-associated condition, subjects could not reliably benefit from but rather be distracted by the other modality. To focus on the strategically mediated modulation of cross-modal interactions, we assessed neural activity in visual and auditory cortex after the cue but preceding the subsequent period of target processing during which other aspects as mismatch detection and response conflict may come into play. To enhance effects in early sensory areas, we used primarily meaningless but challenging low-level stimuli. In particular, for such low-level multisensory interactions, it is conceivable that cross-modal interactions might be governed by subcortical gating. It has been shown that the thalamus houses a local circuitry that would be suitable to efficiently mediate cross-modal interactions (Crabtree and Isaac, 2002). We therefore speculated that we might observe an effect of condition in the thalamus as well.

Materials and Methods

Subjects. The eight subjects participating in the functional magnetic resonance imaging (fMRI) study were 23–33 years old (six females) with normal hearing and normal or corrected-to-normal vision. Participants were paid for participation in the study conducted in conformity with the Declaration of Helsinki.

Stimuli. Each trial started with a fixation period of variable duration (7–11 s). Then a cue indicated to the subjects whether their subsequent response should be based on auditory or visual information. Cues were presented for 1 s and consisted of framing in blue or red the visual display ($24 \times 16^\circ$) (Fig. 1). A red frame signaled that the auditory target stimulus would be relevant, and a blue frame signaled that the visual target stimulus would be relevant. The cue was followed by another fixation period of 6 or 8 s duration, and then visual and auditory target stimuli were presented simultaneously for 1 s. Auditory target stimuli were tones at 700 or 600 Hz that were bilaterally presented through custom-made MR-compatible headphones. Visual target stimuli were bars tilted with a steep (50°) or flat (30°) angle. The bars were presented at the center of the screen and comprised $7 \times 2^\circ$ (Fig. 1). When cued to the visual modality, subjects were instructed to press the left button for the bar with the steep

angulation and the right button for the one with the shallow angulation. In auditory task trials, the left button was associated with the higher tone, and the right button was associated with the lower tone. Maximum response time was set to 2 s.

There were two conditions for this paradigm: associated and non-associated. In the associated condition, 88.9% of trials contained fixed cross-modal stimulus pairs: in 24 trials, the high tone was paired with a steep bar, and, in another 24 trials, the low tone was paired with a flat bar. Only six trials contained the opposite pairings, i.e., high tone and flat bar and low tone and steep bar. The latter trials were catch trials aimed at avoiding that subjects would focus on the same modality regardless of cue type, for example by basing their responses on the auditory target regardless of whether the visual or the auditory modality had been cued. In the non-associated condition, tones and bars were combined at random, i.e., there were 13 trials each with the combination high–steep, high–flat, low–steep, and low–flat, respectively. Associated and non-associated conditions were tested in different sessions at least 2 d apart. Although this across-session comparison may have introduced additional variance, we chose this approach over a within-session comparison of blocked trials for the following reason: subjects had to learn whether the two modalities conveyed consistent information or not. To do so, extensive training was necessary before scanning to learn the respective rule and in the second session to unlearn the former rule. Training sessions were repeated until subjects achieved a hit rate of at least 80%. On average, this involved five runs of the experiment in which, however, the fixation periods were reduced to 3 s for time saving. If we instead had presented the different conditions in short consecutive blocks, subjects would either have had to change their strategy over and over again or they might have adopted a common single strategy for both tasks (i.e., always ignore the uncued modality resulting in a carryover from the non-associated into the associated condition) so that our task modulations would have become useless. Across subjects, the order of associated and non-associated sessions was randomized.

Although most subjects felt that the combination in the associated condition was intuitive, we took special caution to also ensure that subjects had explicitly understood this association. First, subjects were informed about the respective rule as well as the occurrence of catch trials. Second, training sessions were administered before scanning of each condition for the subjects to familiarize themselves with the new rule and “unlearn” the old rule.

fMRI procedure. fMRI data were acquired with a 3T-MRI system (Siemens Trio; Siemens, Erlangen, Germany). Stimuli were back-projected onto a screen that subjects looked at via a mirror mounted onto a standard head coil. During each session, 2×518 (associated condition) or 2×504 (non-associated condition) volumes of 32 axial slices (3 mm thickness, 0.3 mm gap) were collected using a gradient echo–echo planar imaging sequence [repetition time (TR), 2000 ms; echo time (TE), 30 ms; flip angle, 60° ; voxel size, $3.3 \times 3.3 \times 3$ mm]. Structural three-dimensional datasets were acquired using a T1-weighted sagittal magnetization-prepared rapid-acquisition gradient echo sequence (TR, 2250 ms; TE, 2.6 ms; flip angle, 9° ; inversion time, 900 ms; voxel size, $1.1 \times 1.0 \times 1.1$ mm).

Data analysis. fMRI data were analyzed with the Brainvoyager software (Brain Innovation, Maastricht, The Netherlands). After correction for slice scan time differences within a volume, functional volumes were coregistered with the three-dimensional normalized structural datasets to generate volume–time courses that then were motion corrected and temporally high-pass filtered at 336 s.

Region of interest analysis. Regions of interest (ROIs) were determined by first selecting those areas responding to the simultaneously presented auditory and visual target stimuli and then by labeling the respective activation clusters as early visual cortex (VC), early auditory cortex (AC), and thalamus based on anatomical knowledge.

Our analysis avoided confounding the top-down driven cue-induced expectancy of interest with bottom-up sensory processing of the cue: the cue was a frame presented at the outer edge of the screen, whereas the visual target was a small bar at the center of the screen. In other words, we chose a cue that, as a sensory stimulus, activated different subareas in early visual cortex than those responding to the target. Accordingly, any

cue-related effects in those latter areas could be assigned to a nonsensory effect induced by the cue signal (Müller and Kleinschmidt, 2003; Müller et al., 2003). Because the targets and not the cues served to determine the ROIs, activity in response to the cue in these ROIs should reflect expectancy-driven modulation rather than sensory processing of the visual cue.

The auditory and visual target stimuli served as a regressor to functionally determine the ROIs by calculating for each subject a fixed-effects general linear model. This contrast served to identify on a subject-by-subject basis those regions responding during the sensory target input, in other words, the candidate regions in which we expected cue-induced activity modulations that would be relevant for subsequent sensory target processing. This analysis revealed activation clusters in early visual and auditory cortices as well as in the thalamus. The resulting ROIs were hence labeled as VC and AC based on anatomical criteria: clusters along the calcarine sulcus were defined as VC, and clusters along Heschl's gyrus were defined as AC. Clusters in the diencephalon on the central base of the brain directly on top of the mesencephalon were regarded as belonging to the thalamus. Although we could not assign the latter to specific thalamic nuclei, they were clearly remote from the specific visual and auditory thalamic relay nuclei. In cases in which the selection of an ROI remained ambiguous in a subject, e.g., when encountering closely neighboring clusters in the same overall region, the cluster closest to ROI of the group average was selected (for the Talairach coordinates, see Table 1). Given our use of foveal stimuli, we could not differentiate effects in primary cortex from those in secondary visual cortex. Furthermore, because only early visual cortex has receptive fields small enough to ensure that subareas representing the cue could be differentiated from those representing the target, we refrained from analyzing areas such as visual area 4.

Analysis of cue-induced activity. Within each ROI, the voxel with the peak activation was selected, and the signal was then averaged across the surrounding $3 \times 3 \times 3$ voxels. The 2 s preceding the cue served as baseline. For the subsequent group analysis, the event-related data from each ROI were averaged within a time interval from 3 to 9 s after cue onset and thus did not include contributions from target processing, e.g., mismatch detection. The resulting mean blood oxygen level-dependent (BOLD) level of each ROI (AC, VC, and thalamus) was entered into a repeated-measure ANOVA with the factors cued modality (visual vs auditory), condition (associated vs non-associated), and brain region (visual, auditory, and thalamic) by collapsing the two hemispheres.

Results

Behavioral data

As predicted, subjects made more errors (20.4 vs 10.4%; $t = 2.5$; $p < 0.05$) and were slower (671 vs 630 ms; $t = 4.1$; $p < 0.01$) in the non-associated than in the associated conditions, indicating that the non-associated condition was more demanding and more prone to distraction. Regarding the catch trials in the associated condition, subjects responded correctly in the majority of trials (76%). There was no difference between visual and auditory trials ($t = 2.6$, $p = 0.8$ for accuracy; $t = 1.2$, $p = 0.3$ for reaction time). Analyzing separately visual and auditory conditions in the associated and non-associated conditions, we found no difference between auditory and visual modality in terms of accuracy ($F_{(1,7)} = 1.05$; $p = 0.4$) or reaction time ($F_{(1,7)} = 3.9$; $p = 0.09$). There was also no sign of an interaction between condition and modality regarding the number of errors ($F_{(1,7)} = 1.23$; $p = 0.3$) or reaction time ($F_{(1,7)} = 0.01$; $p = 0.1$).

fMRI data

The primary goal of our study was to compare activity in early sensory structures depending on cross-modal relationships.

Table 1. Mean and range (in parentheses) of Talairach stereotaxic coordinates of brain areas assessed in the study

Region	Side	x	y	z
Auditory cortex	Right	51 (37 to 64)	-15 (-10 to -20)	5 (2 to 10)
	Left	-52 (-42 to -63)	-21 (-12 to -33)	4 (3 to 6)
Visual cortex, cue	Right	9 (6 to 13)	-75 (-68 to -84)	-4 (-2 to -5)
	Left	-8 (-6 to -10)	-76 (-70 to -85)	-5 (-3 to -7)
Visual cortex, target	Right	6 (1 to 10)	-82 (-58 to -92)	-3 (9 to -14)
	Left	-7 (-1 to -17)	-79 (-69 to -86)	-6 (9 to -17)
Thalamus	Right	8 (5 to 12)	-19 (-14 to -24)	7 (6 to 8)
	Left	-10 (-8 to -14)	-20 (-16 to -26)	7 (5 to 9)

We found that cueing compared with baseline enhanced activity in all sensory cortices assessed (Fig. 2). Although cueing was implemented as a visual signal for both modalities, we found that cueing the visual modality enhanced activity in target-related VC more than cueing the auditory modality; conversely, cueing the auditory modality resulted in greater activity in AC than cueing the visual modality ($F_{(1,7)} = 11.90$; $p < 0.01$). The sensory responses to cues and targets could be spatially separated thanks to their retinotopic distance (Fig. 3, Table 1).

Our crucial finding was obtained by testing the triple interaction that corresponds to the hypothesis that had motivated our experiment. The interaction between condition (non-associated vs associated), sensory area (visual vs auditory cortex), and cued modality (visual vs auditory task) was highly significant ($F_{(1,7)} = 12.75$; $p < 0.01$). In VC, activity was most pronounced in the non-associated condition if the visual modality had been cued as task relevant and lowest when the auditory modality had been cued as task relevant in the non-associated condition ($t = 2.9$; $p < 0.03$). In the associated condition, cue-induced activation of VC remained at intermediate levels and was nearly the same regardless of which modality had been cued as task relevant ($t = 0.38$; $p = 0.7$). In AC, the reverse pattern was observed: activity was highest when, in the non-associated condition, the auditory modality was cued and lowest when, in the non-associated condition, the visual modality was cued ($t = 4.09$; $p < 0.01$). Again, the associated condition yielded intermediate levels of activation that did not differ with respect to the cued modality ($t = 0.41$; $p = 0.7$).

Figure 4 summarizes the findings in early sensory cortices by collapsing data from different sensory areas. Mean activity across sensory areas is plotted as a function of whether the sensory area matched the cued modality (i.e., VC after cueing visual and AC after cueing auditory) or not (i.e., VC after cueing auditory and AC after cueing visual). In the associated condition, activity in early sensory areas was approximately the same regardless of whether the related modality had been cued as task relevant or the other modality instead. In the non-associated condition, activity was relatively enhanced in the matching area and relatively suppressed in the nonmatching area. In other words, if subjects expected that performing the task in one modality would be supported by the input of the other modality, preparatory activity was enhanced in both sensory systems. If, however, the other modality was expected to convey distracting information, then activity was enhanced only in the task-relevant system, whereas it was reduced in the nonrelevant system.

Our analysis of target-driven responses across the early sensory pathways had also robustly revealed thalamic activation foci. Analyzing cue-induced preceding activity changes in these ROIs as we did for early sensory cortices, we found that regional

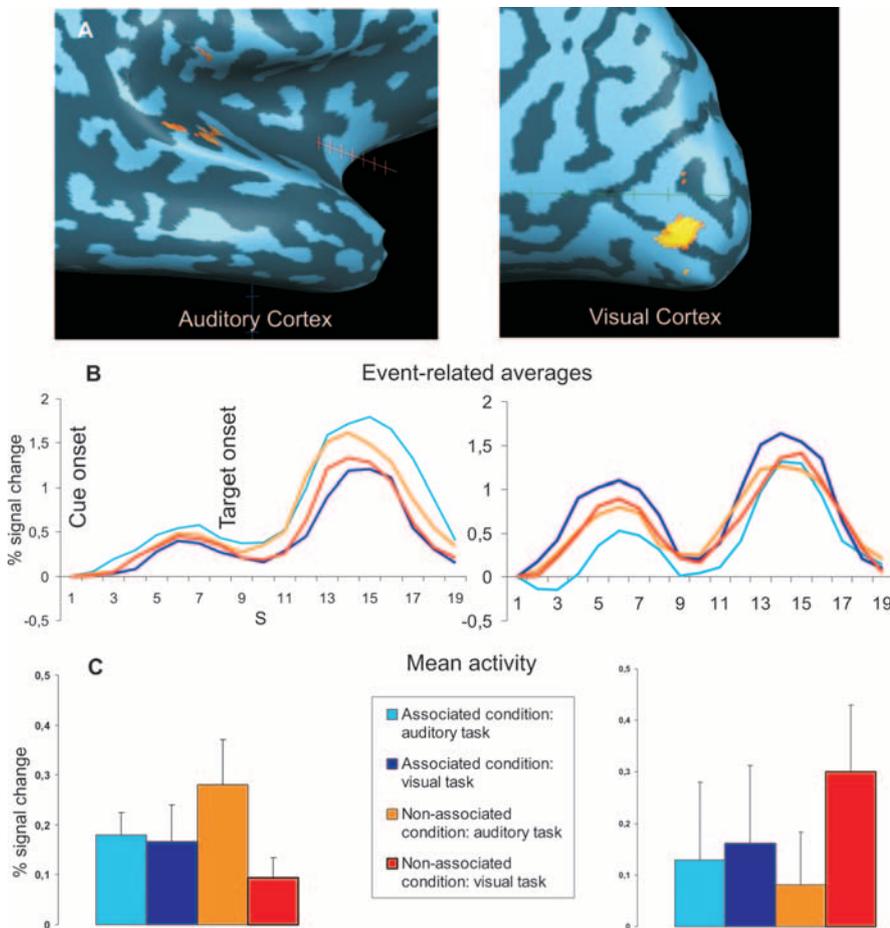


Figure 2. *A*, Regions identified as early auditory and visual cortex in a single inflated hemisphere of one representative subject. *B*, Event-related BOLD response in response to cue and target in the same subject (8 s trials only). *C*, Bar plots of group-averaged results.

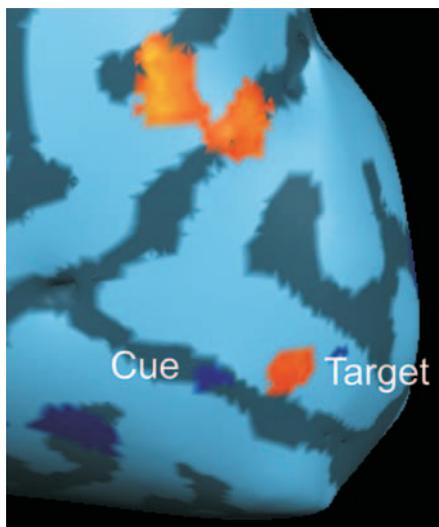


Figure 3. Activation clusters induced by the target and cue (group analysis). Note that the two stimuli activated different subareas in primary visual cortex.

thalamic activity displayed a main effect of condition ($t = 11.12$; $p < 0.001$) because of higher activity in the non-associated vs associated condition (Fig. 5). No additional main effects or interactions were observed in thalamic ROIs.

Discussion

To the best of our knowledge, this is the first study to isolate neural correlates of expectancy in cross-modal processing. This was achieved while maintaining equivalent sensory input across conditions and by measuring cue-induced instead of target-related activations. We manipulated expectancy via the statistical relationship between meaningless arbitrary items presented to two different senses and made subjects aware of this relationship. We could demonstrate that the expected association between visual and auditory input modulates preparatory activity in early sensory cortices as well as in the thalamus. When subjects expected a random association of auditory and visual items, activity in the cued sensory area became strongly enhanced, whereas there was little activity in the task-irrelevant sensory system. When auditory and visual stimuli were paired in a reliable systematic manner, such that they conveyed concordant information, activity was enhanced in both systems regardless of which modality had been cued.

We also found that, in thalamic regions that responded during target processing but did not seem to cover the modality-specific relay nuclei, activity was enhanced during the non-associated over the associated version of the task. The enhanced activity in the thalamus that we observed during the non-associated condition may reflect a selection or “gate-keeping” process, but we cannot assign it to a specific thalamic substructure or even ascertain that it is confined to a single such structure. It has been suggested that suppression of sensory activity in primary visual cortex (V1) when multiple stimuli are present is not mediated through intracortical connections but originates from feed-forward thalamic signals (Freeman et al., 2002). Other single-neuron studies have found evidence for mutual suppression between modality-related dorsal thalamic nuclei (Crabtree et al., 1998, 2002), and the thalamic reticular nucleus codes retinotopic information required for spatial orienting but may act on geniculocortical transmission instead of projecting to V1 (McAlonan et al., 2006). Alternatively, the enhanced thalamic activity in our non-associated task may be related to task difficulty and increased alertness.

As the items used in our study were meaningless, so was their association: there is no real-life congruence or incongruence for our target stimuli. Modulation of sensory activity occurred while subjects were waiting for the targets to appear, i.e., before the actual targets had to be processed. Although we exclusively applied a visual cue and scanner noise certainly drives baseline auditory cortical activity, our results were quite symmetrical for the two modalities. This suggests that neither the visual cue induced preference for the visual modality nor the scanner noise drowned effects in auditory cortex. Moreover, we could exclude the possibility that activity in visual cortex was simply bottom-up driven by the cue: the assessed ROI retinotopically represented the target

and did not overlap with visual cortex corresponding to the cue (Fig. 3).

Condition-dependent effects were assessed throughout the preparatory period, but the initial transient cue-induced response may indicate a contribution from an additional endogenous signal that very recently has been described in the literature (Jack et al., 2006). Our observation of a similar transient cue-induced component in auditory cortex furthermore suggests that the observation made by Jack et al. also applies to auditory and not only visual cortex. A thalamic source of those cue-induced transient modulations has been suggested, and it is conceivable that the thalamic effect we observed here reflects a modulation of, for instance, alerting levels by strong versus random association of multisensory input.

At first sight, our results in cortical areas resemble those from classical studies of selective versus divided attention, in which subjects have to discriminate different attributes of the same set of visual stimuli (Corbetta et al., 1990). These studies show higher sensitivity to detect subtle stimulus differences when subjects can focus on one attribute instead of dividing attention among several attributes. At the same time, attention enhances the activity in different regions of extrastriate visual cortices that are specialized for processing information related to the selected attribute. With respect to our study, one could assume that subjects divided attention between the auditory and visual modality in the associated condition and focused attention on one modality in the non-associated condition. This interpretation is unlikely for the following reasons. First, our subjects never actually had to divide attention; instead, in all tasks, they could perform the task by focusing on a single modality. Second, the behavioral results suggest that, even in the associated condition, subjects did not evenly divide attention across both visual and auditory input but remained biased toward the cued modality.

Our experimental approach also differs profoundly from those previous studies that reported effects of congruence of spatial or item-based properties in cross-modal processing. Not only did we analyze neural effects occurring before targets and thus potential congruence, it is impossible to define congruence or incongruence for our stimuli because they were meaningless. This difference is important because the effects of congruence can only be demonstrated by testing for a modulation of target processing. Accordingly, effects of congruence reported previously (Macaluso et al., 2000; Weissman et al., 2004) may reflect the detection of a “semantic” or spatial mismatch between two stimuli requiring additional processing or a response conflict that might amplify target-relevant processing.

Finally, our experimental design also differs from those previous studies that compared unimodal and bimodal processing. In the study by Laurienti et al. (2002), unimodal selective attention was associated with reductions of ongoing activity in cortex belonging to the nonstimulated modality, whereas divided attention to bimodal input without a behavioral task yielded activation in sensory cortices of both modalities. Johnson and Zatorre (2005) used melodies as auditory stimuli and abstract shapes as visual stimuli, which were either presented alone or together. They could show that both conditions, the unimodal as well as the bimodal one, led to an increase of the BOLD responses in the relevant sensory areas and a decrease in the irrelevant sensory areas. Behaviorally, attended stimuli were remembered better than nonattended stimuli.

Hence, although our experimental approach differs from previous work in several crucial features, our findings are compatible with these previous results and may in fact point at the mecha-

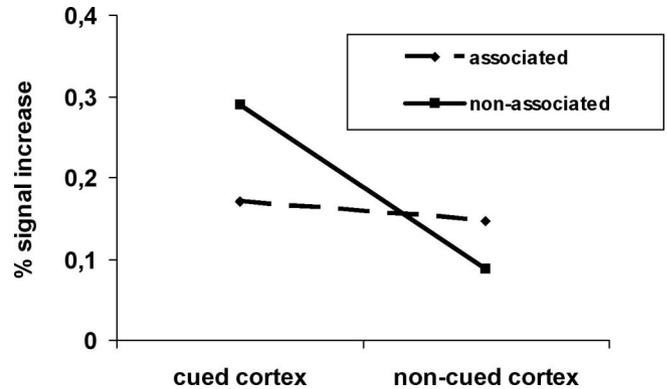


Figure 4. Mean activity averaged across early sensory areas (visual and auditory) as a function of whether the cued modality matched the sensory cortex or not.

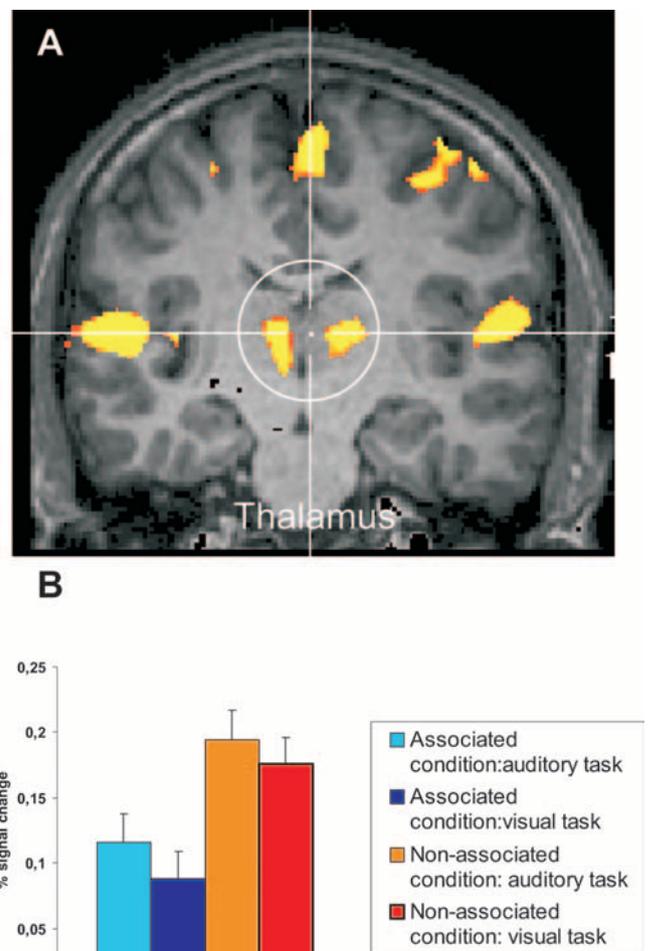


Figure 5. *A*, Regions identified as thalamus in a single subject. *B*, Bar plots of group averaged results.

nisms that account for these effects. Congruence of sensory input, for instance, may activate a feedback loop that results in a cross-modal facilitation of processing in the unattended modality. The cue-induced effects we observed would then come into play during the period of target processing and remain undistinguishable from other response components. Conversely, first-pass detection of incongruence could result in a suppression of ongoing input processing in the unattended modality. Similarly, the tran-

sition from unimodal to bimodal stimulation would be expected to result in a suppression of the modality added on unless attention was instructed to be divided across the two (Laurienti et al., 2002; Johnson and Zatorre, 2005).

Our experiment shows some resemblance to a recently published study by Weissman et al. (2004). In that study, subjects were presented with written and spoken letters that either matched or not. In case of incongruent letter pairings and when subjects had been cued to the visual modality, activity in visual cortex but not in auditory cortex was increased compared with congruent pairings. Conversely, a conflict between auditory and visual letter increased activity in auditory but not in visual cortex when the auditory letter served as target. Again, however, this study differs from ours in several crucial points. First, our visual and auditory stimuli, unlike visual and auditory letters, were unrelated before training, and an association was only established by the current task. Second, we used simple stimuli that do not necessarily rely on higher-order areas (serving language-oriented processes) but may be classified within primary sensory areas. Third, and most crucial, Weissman et al. assessed neural activity when the actual targets were presented. Hence, their effects may reflect the detection of a mismatch between the two stimuli, a response conflict, or may simply be based on differences in sensory input between congruent and incongruent stimulus pairs. In our study, conversely, the same stimulus pairs were used in the associated and non-associated condition, and neural activity was measured before target onset so that the observed effects in sensory areas were most likely top-down mediated, in line with the conclusion proposed by Johnson and Zatorre (2005). The latter assumption is based on fMRI studies demonstrating that the flexible adjustment of sensory activity relies on recently acquired information maintained in working memory whereby frontal and parietal cortices (Macaluso and Driver, 2003) play crucial roles. Indeed, several studies have reported attention-driven preparatory activity in sensory areas specialized for the expected stimulus that was controlled by a frontoparietal network. The hypotheses motivating our study specifically targeted effects in early sensory cortical and subcortical areas. Because our design was tailored accordingly, we refrained from exploratory analyses of effects in higher cortical association areas.

Whereas Weissman et al. (2004) only reported enhanced activity in the relevant sensory system with conflicting bimodal inputs, Laurienti et al. (2002) and Johnson and Zatorre (2005) also observed reduced activity in the irrelevant system. We found cross-modal suppression only in conditions in which auditory and visual inputs were expected to convey unrelated information.

If sound and vision were linked with the same response, activity was enhanced in both sensory systems regardless of which modality had been cued, although to a lesser degree than in the cued system during the non-associated condition. In other words, if subjects are likely to benefit from the task-irrelevant modality, activity in the respective system will be enhanced instead of being suppressed. The brain hence appears to dispose of a rather flexible system for modulating cross-modal interactions.

References

- Crabtree JW, Isaac JT (2002) New intrathalamic pathways allowing modality related and cross-modality switching in the dorsal thalamus. *J Neurosci* 22:8754–8761.
- Crabtree JW, Collingridge GL, Isaac JT (1998) A new intrathalamic pathway linking modality-related nuclei in the dorsal thalamus. *Nat Neurosci* 1:389–394.
- Corbetta M, Miezin FM, Dobmeyer S, Shulman GL, Petersen SE (1990) Attentional modulation of neural processing of shape, color and velocity in humans. *Science* 248:1556–1559.
- Driver J, Spence C (1998) Cross-modal links in spatial attention. *Philos Trans R Soc Lond B Biol Sci* 353:1319–1331.
- Eimer M, Schröger E (1998) ERP effects of intermodal attention and cross-modal links in spatial distributions. *Psychophysiology* 35:313–327.
- Freeman TC, Durand S, Kiper DC, Carandini M (2002) Suppression without inhibition in visual cortex. *Neuron* 35:759–771.
- Jack AI, Shulman GL, Snyder AZ, McAvoy M, Corbetta M (2006) Separate modulations of human V1 associated with spatial attention and task structure. *Neuron* 51:135–147.
- Johnson JA, Zatorre RJ (2005) Attention to simultaneous unrelated auditory and visual events: behavioural and neural correlates. *Cereb Cortex* 15:1609–1620.
- Laurienti PJ, Burdette JH, Wallace MT, Yen YF, Field AS, Stein BE (2002) Deactivation of sensory-specific cortex by cross-modal stimuli. *J Cogn Neurosci* 14:420–429.
- Macaluso E, Driver J (2003) Multimodal spatial representations in the human parietal cortex: evidence from functional imaging. *Adv Neurol* 93:219–233.
- Macaluso E, Frith CD, Driver J (2000) Modulation of human visual cortex by crossmodal spatial attention. *Science* 289:1206–1208.
- McAlonan K, Cavanaugh J, Wurtz RH (2006) Attentional modulation of thalamic reticular neurons. *J Neurosci* 26:4444–4450.
- McDonald JJ, Teder-Salejarvi WA, Hillyard SA (2000) Involuntary orienting to sound improves visual perception. *Nature* 407:906–908.
- Müller NG, Kleinschmidt A (2003) Dynamic interaction of object- and space-based attention in retinotopic visual areas. *J Neurosci* 23:9812–9816.
- Müller NG, Bartelt OA, Donner TH, Villringer A, Brandt SA (2003) A physiological correlate of the “zoom lens” of visual attention. *J Neurosci* 23:3561–3565.
- Weissman DH, Warner LM, Woldorff MG (2004) The neural mechanisms for minimizing cross-modal distraction. *J Neurosci* 24:10941–10949.