Behavioral/Systems/Cognitive

Sound-Induced Enhancement of Low-Intensity Vision: Multisensory Influences on Human Sensory-Specific Cortices and Thalamic Bodies Relate to Perceptual Enhancement of Visual Detection Sensitivity

Toemme Noesselt,[1,2]* **Sascha Tyll,**[2]* **Carsten Nicolas Boehler,**[2] **Eike Budinger,**[2,3] **Hans-Jochen Heinze,**[2,3] **and Jon Driver**[4]

[1]Department of Psychology, Otto-von-Guericke-Universität, 39016 Magdeburg, Germany, [2]Department of Neurology II, Otto-von-Guericke-Universität, 39120 Magdeburg, Germany, [3]Leibniz-Institut für Neurobiologie, 39118 Magdeburg, Germany, and [4]Wellcome Centre for Neuroimaging and University College London Institute of Cognitive Neuroscience, University College London, London WC1N 3BG, United Kingdom

Combining information across modalities can affect sensory performance. We studied how co-occurring sounds modulate behavioral visual detection sensitivity (d'), and neural responses, for visual stimuli of higher or lower intensity. Co-occurrence of a sound enhanced human detection sensitivity for lower- but not higher-intensity visual targets. Functional magnetic resonance imaging (fMRI) linked this to boosts in activity-levels for sensory-specific visual and auditory cortex, plus multisensory superior temporal sulcus (STS), specifically for a lower-intensity visual event when paired with a sound. Thalamic structures in visual and auditory pathways, the lateral and medial geniculate bodies, respectively (LGB, MGB), showed a similar pattern. Subject-by-subject psychophysical benefits correlated with corresponding fMRI signals in visual, auditory, and multisensory regions. We also analyzed differential "coupling" patterns of LGB and MGB with other regions in the different experimental conditions. Effective-connectivity analyses showed enhanced coupling of sensory-specific thalamic bodies with the affected cortical sites during enhanced detection of lower-intensity visual events paired with sounds. Coupling strength between visual and auditory thalamus with cortical regions, including STS, covaried parametrically with the psychophysical benefit for this specific multisensory context. Our results indicate that multisensory enhancement of detection sensitivity for low-contrast visual stimuli by co-occurring sounds reflects a brain network involving not only established multisensory STS and sensory-specific cortex but also visual and auditory thalamus.

Introduction

There is a growing literature on how combining information from different senses may enhance perceptual performance. The principle of "inverse effectiveness" (PoIE) originally introduced by Stein and colleagues for cell recordings (for an overview, see Stein and Meredith, 1993) suggests that co-occurrence of stimulation in two modalities may lead to enhanced neural responses, particularly for stimuli that produce a weak response in isolation [but see Holmes (2009) for critique and Angelaki et al. (2009) for reconsideration within a Bayesian framework]. It has been suggested that one behavioral consequence of a putative PoIE might be enhanced detection sensitivity for near-threshold stimuli in one sense when co-occurring with an event in another sense (Stein and Meredith, 1993; Frassinetti et al., 2002).

Many (but not all) behavioral studies of multisensory integration have used relatively intense suprathreshold stimuli, hence could not test for near-threshold detection sensitivity (but see McDonald et al., 2000; Frassinetti et al., 2002). Some audiovisual studies did assess cross-modal effects in relation to stimulus intensity but studied intensity matching (Marks et al., 1986), audiovisual changes (Andersen and Mamassian, 2008), or reaction time (Doyle and Snowden, 2001), rather than unimodal detection sensitivity (d'). Given PoIE proposals that multisensory benefits should arise particularly in near-threshold detection for low-intensity stimuli, we focused on d' for lower-intensity (vs higher-intensity) visual stimuli when co-occurring with a sound or alone.

Recent results on the neural basis of audiovisual interactions indicate that these may affect not only brain areas traditionally considered as multisensory convergence zones, such as cortically the superior temporal sulcus (STS) (Bruce et al., 1981; Beauchamp et al., 2004; Barraclough et al., 2005), but also areas traditionally considered as modality specific [e.g., visual striate/extrastriate cortex (Calvert, 2001; Noesselt et al., 2007; Watkins et al., 2007), plus core/belt auditory cortex (Brosch et al., 2005;

Kayser et al., 2007; Noesselt et al., 2007)] [see Ghazanfar and Schroeder (2006) and Driver and Noesselt (2008) for reviews on multisensory integration; and Falchier et al. (2002), Rockland and Ojima (2003), Cappe and Barone (2005), and Budinger and Scheich (2009) for possible anatomical pathways]. Subcortical structures have also long been implicated in multisensory integration (Stein and Arigbede, 1972). Recent work indicates possible thalamic involvement in some multisensory effects (Baier et al., 2006; Cappe et al., 2009), including audiovisual speech processing (Musacchia et al., 2006) or training-induced plastic changes in speech processing (Musacchia et al., 2007). But here we focus on nonsemantic stimuli (cf. Sadaghiani et al., 2009), with particular interest in whether sensory-specific thalamus can be implicated.

We studied an audiovisual situation in which co-occurring sound bursts might enhance detection sensitivity (d') for visual Gabor patches. Behaviorally, given PoIE proposals, we predicted d' benefits attributable to a co-occurring sound for lower- but not higher-intensity visual targets. We sought to identify the neural basis of any such behavioral pattern in the human brain using event-related functional magnetic resonance imaging (fMRI). We tested within localized sensory-specific or heteromodal regions for patterns of blood oxygen level-dependent (BOLD) activation (or interregional coupling) during our audiovisual paradigm that corresponded to the behavioral pattern for visual detection sensitivity.

Materials and Methods

Behavioral study

Fourteen subjects (eight females; age range, 19–33 years) participated in the initial psychophysical experiment. An additional 12 participated in the fMRI study (see below) and also yielded behavioral data. Subjects provided informed consent in accordance with local ethics.

Outside the scanner, visual stimuli were presented on a monitor using Presentation 9.13 (Neurobehavioral Systems). Visual targets comprised rectified Gabor patches that differed in luminance. These stimuli subtended 1.5° visual angle, had a duration of 16.6 ms, a spatial frequency of 3 cycles per degree, and were presented at 5° horizontal and 1° vertical eccentricity in the upper right quadrant. Auditory stimuli were presented from a loudspeaker located just above the visual stimulus position and comprised a 3 kHz sound burst (70 dB; duration, 16 ms); see below for sounds during scanning.

On each trial, subjects performed a signal detection task for visual targets, indicating the presence or absence of a visual target stimulus by pressing one of two buttons, regardless of whether a co-occurring sound was presented or not. They had to maintain central fixation and respond as accurately and quickly as possible [we collected reaction times (RTs) for completeness, but they showed a different outcome compared with the critical signal-detection measure of visual sensitivity, namely d' (see below)]. A visible outline square on the monitor (1.7 × 1.7°, 13.05 cd/ m², 16.6 ms duration), surrounding the possible target position, always appeared to signal when a response was required (see Fig. 1a). This visual square was present in all stimulus conditions, so was not predictive of target presence, and will be subtracted out by our contrasts of the different conditions in the later fMRI experiment (see below), since it appeared in all conditions. We introduced it to signal when a response was required, and thereby to serve also as an onset marker for the no-sound no-target condition, which otherwise could not have been estimated straightforwardly for the fMRI response. Likewise, no RTs for the no-sound no-target condition could have been collected without the square frame to indicate a response was required.

The very brief target duration (16.6 ms) was chosen to obtain hit rates in the range of 60 –80% without a mask and also because similarly short stimulus durations have recently been used in several other studies of multisensory integration (Bonath et al., 2007; Lakatos et al., 2007). We acknowledge that longer presentation durations might in principle yield

different results (for some neural and behavioral effects of manipulating audiovisual stimulus durations, see Meredith et al., 1987; Boenke et al., 2009). The possible role of stimulus duration might be studied in future variants of the paradigm introduced here.

Visual targets were presented on a random one-half of trials, and on a random one-half of these target-present trials the auditory sound burst also occurred concurrently. But the sound was equally likely to appear on target-absent trials also (see Fig. 1c), thus conveying no information about the visual nature of the trial. Initially, there were three visual-threshold-determination runs, in which targets at eight intensity levels (12.30, 8.50, 8.12, 7.93, 7.71, 7.31, 7.09, or 6.91 cd/m², against a 6.01 cd/m² gray background) could occur. In each run, there were 30 trials per intensity condition, plus 240 nontarget trials. These runs were used to determine the two intensity levels that yielded 55– 65% correct detection and 85–95% correct detection for each subject (note that even the latter value was consistently below 100% ceiling). The two selected intensity values were then used in the main experiment, with one-half of the target presentations at the higher luminance and one-half at the lower, in random order. Across subjects, the two mean luminances used were 7.09 and 7.71 cd/m² (see Fig. 1b).

The main experiment had a 3 × 2 factorial event-related design, with factors of visual condition (no target, lower-intensity target, or higher-intensity target) and sound condition (present/absent). The same number of nontarget trials and target trials were presented in three runs (90 higher intensity target trials, 90 lower intensity target trials, and 180 nontarget trials per subject) with a mean intertrial interval of 1500 ms; range, 1200 –1800 ms randomized. We used standard signal-detection analyses (Stanislaw and Todorov, 1999) to separate perceptual sensitivity (d') from criterion (c) for detection of the visual target. This separation is important; for example, some recent studies attributed behavioral effects of an uninformative sound on visual judgments to a response bias rather than increased sensitivity (Marks et al., 2003; Odgaard et al., 2003; Lippert et al., 2007), whereas other groups have reported enhanced perceptual sensitivity instead (McDonald et al., 2000; Vroomen and de Gelder, 2000; Frassinetti et al., 2002; Noesselt et al., 2008), but sometimes only with informative sounds (Lippert et al., 2007), unlike here. We treated correct detections of targets as hits, reports of target absence when present as misses, reports of target absence when absent as correct rejections, and reports of target presence when absent as false alarms. Any responses later than 1500 ms after stimulus onset were discarded. We then used a z-transformed ratio to compute sensitivity (d') separately from criterion (c) (Stanislaw and Todorov, 1999). Note that all of the visual signal detection scores were calculated separately for conditions with the (task-irrelevant) sound present, versus absent.

Behavioral results for all measures (d'; response criterion c; accuracy; and for completeness, RT) were analyzed with mixed ANOVAs, having the within-subject orthogonal factors of visual condition and of sound, plus the between-subject factor of inside/outside scanner (for unscanned vs scanned groups of subjects, respectively). The latter factor had no significant impact, so behavioral data from inside and outside scanner could be pooled. Nevertheless, for completeness, we also plot the behavioral results separately for inside and outside the scanner, to illustrate the replicability of the pattern (see supplemental Fig. S1a, available at www. jneurosci.org as supplemental material).

fMRI experiment

fMRI data acquisition. fMRI data were acquired for the MRI group ($n =$ 12; six females; aged 21–29 years) using a circular-polarized whole-head coil (Bruker BioSpin) on a Siemens whole body 3T MRI Trio scanner (Siemens). The procedure was as for the behavioral measures in the unscanned group, except as follows.

First, data from an fMRI localizer run were collected before the behavioral visual-threshold determination process. These runs used blocked high-intensity unisensory visual or auditory stimuli to identify auditorily responsive brain areas; visually responsive areas plus candidate heteromodal areas (i.e., those responding to both visual and auditory stimuli) (255 volumes; field of view, 200 × 200 mm; repetition time/echo time/ flip angle, 2000 ms/30 ms/80°; 30 slices; spatial resolution, 3.5 × 3.5 × 3.5 mm). Six auditory blocks (tones at 70 dB, as used in the main experi-

ment) and six visual blocks with high-intensity visual stimuli (18.56 cd/m^2, at 4 Hz stimulation rate to maximize the visual responses) were presented, each in alternating ON–OFF sequences of 20 s each. Candidate heteromodal brain areas were defined by the overlap between auditory and visual responsive areas (i.e., with a conjunction analysis) (Friston et al., 1999).

Second, eye position was recorded throughout scanning using a custom-made infrared eye-tracking device (Kanowski et al., 2007). Importantly, eye position did not differ between our auditory conditions (<1° from the central point in 95% of trials), so the auditory enhancement of lower-intensity visual detection d' that we report cannot be attributable to changed eye position.

Third, auditory stimuli were recorded for each subject before scanning using in-ear microphones while presenting sounds as for the nonscanned participants. These recordings were then played back via MR-compatible headphones during scanning, to create the perception of a sound coming from a location close to the visual stimulus. Behavioral results did not differ inside versus outside the scanner (see supplemental Fig. S1, available at www.jneurosci.org as supplemental material).

Fourth, behavior together with concurrent fMRI during the task (see information on preceding localizers runs above for technical details of the MR scanning runs, which were the same during the main fMRI experiment) were then acquired for each scanned subject in six experimental runs (each with mean intertrial interval, 2500 ms; range, 2000–6000 ms; Poisson distributed; 180 trials per target condition and 360 trials per nontarget condition).

fMRI analyses of localizer runs and identification of regions of interest. After discarding the first five volumes of each imaging run, data were slice acquisition time corrected, realigned, normalized, and smoothed (6 mm full-width at half-maximum) using SPM2 (www.fil.ion.ucl.ac.uk/spm). After preprocessing, the data from the localizer runs were analyzed with a model comprising the two blocked stimulus conditions (high-intensity visual or auditory) plus the interleaved no-stimulus baseline blocks, for each subject. Subject-specific local maxima within lateral and medial geniculate bodies (LGB, MGB), plus primary visual or auditory cortex, were identified by a combination of functional and anatomical criteria. Functional criteria comprised a preference (at $p < 0.01$ or better) for visual stimulation in the case of LGB and V1; or for auditory stimulation in the case of MGB and A1. Structural criteria comprised posterior-ventral thalamic site for the relevant maxima in the case of LGB or MGB, or the calcarine fissure for V1, or the anterior medial part of Heschl's gyrus for A1/core region in auditory cortex. Anatomical structures were identified in subject-specific inversion-recovery echo planar images that have the same distortions as functional data. Averaged subject-specific local maxima within A1 and V1 were also compared with probability maps of primary visual and auditory cortex (www.fz-juelich.de/ime/spm_anatomy_toolbox) to further corroborate these localizations (for identified local maxima in LGB, MGB, A1, and V1 for all subjects, see supplemental Figs. S2, S3, available at www.jneurosci.org as supplemental material). Finally, unisensory and candidate multisensory regions at the group voxelwise stereotactic level (as a supplement to individual anatomical criteria) were further identified using a random-effects group ANOVA, with comparisons of auditory and visual stimulation plus their respective baselines (thresholded at $p < 0.01$; corrected for cluster level) (see supplemental Table S1, available at www.jneurosci.org as supplemental material), to define "inclusive masks" for use in the statistical parametric mapping (SPM) analysis of the main experiment, as described below.

fMRI analysis of activations in main experiment. For the main multisensory experiment (i.e., the visual detection task, with or without co-occurring auditory events on each trial), all trials with correct responses for the six experimental conditions (and separately all those trials with incorrect responses, regardless of condition) were modeled for each participant using the canonical hemodynamic response function in SPM2. The incorrect responses were not further split by condition because of their infrequent occurrence (see below). Since only correct trials were considered for each condition in the fMRI results we present, the observed modulations of BOLD response by condition cannot be readily explained by different error rates in specific conditions.

Voxel-based group results were assessed for the 3×2 factorial design of the main experiment with SPM, using a within-subject ANOVA with all six experimental conditions as provided by SPM2, at the random-effects group level. To focus initially on sensory-specific (visual or auditory) or candidate heteromodal brain areas (responding to both vision and audition), results from the main event-related fMRI experiment were initially masked with the visual, auditory, or heteromodal SPM masks as identified by the separate (and thus independent) blocked localizer runs (masked at $p < 0.01$, cluster level corrected for cortical structures). But whole-brain results are also reported where appropriate (see supplemental material and tables, available at www.jneurosci.org). Significance levels in the main experiment were set to $p < 0.01$ [small volume corrected [false discovery rate (FDR)] using the localizer masks] unless otherwise mentioned. Cluster levels for cortical regions were set to $k > 20$. The cluster level criterion for MGB/LGB was lowered to $k > 5$ because of the smaller size of these subcortical structures (D'Ardenne et al., 2008).

Finally, in addition to the group voxelwise SPM analysis in normalized stereotactic space, we also ran corroborating additional analyses on subject-specific regions of interest (ROIs) (for LGB, MGB, V1, and A1), as defined by the localizer runs but now separately for each individual, via the combination of functional and structural criteria described earlier, independent of the main experiment. As will be seen, these provide an important form of corroborating analysis that is not subject to some of the selection issues that can inevitably arise in normalized SPM group contrasts (for which interaction contrasts inevitably select the most significant peak voxel for an interaction term, which might therefore show an exaggerated pattern compared with an independently defined ROI). ROIs were centered on subject-specific local maxima, and each had a radius of 2 mm. We then tested the extracted BOLD signals per condition across the group, but now for individually defined ROIs, which could thus correspond to somewhat different voxels in different subjects, albeit for the same defined region, unlike the voxelwise normalized group analysis with the normalized inclusive group-masks.

Interregional coupling analysis and brain–behavior relationships. For effective-connectivity analyses, time courses from the LGB and MGB sites identified in each individual subject's thalamic ROIs (via their individual combined structural and blocked-functional criteria, see above) were extracted and analyzed in extended models with single-subject seeds. The regressors for analyzing each subject comprised six experimental conditions, plus (separately) all incorrect trials as above, now plus LGB or MGB time courses also (see below) and the derived "psychophysiological interaction" (PPI), within the standard PPI approach (Friston et al., 1997) for assessing condition-dependent interregional coupling. The PPI approach is a well established and relatively assumption-free approach that measures covariation of residual variance across regions as the index of effective connectivity, in the context of an experimental factor or factors. More specifically, PPIs test for changes in interregional effective connectivity (i.e., for higher or lower covariation in residuals) between a given "seed" region and other brain regions, for a particular context relative to others. Here, we used this standard approach to test specifically for higher interregional coupling when a sound (minus no sound) co-occurred with a lower-intensity (vs higher-intensity) visual target, using seeds as described below. In particular, our single-subject SPM model now included all seven experimental conditions plus five additional regressors for the physiological response from the individually defined seed within the thalamus (LGB or MGB, in separately seeded analyses) and the products of this physiological response with the four conditions yielded by crossing high/low visual intensity with the presence or absence of sounds. We then took the individually computed PPI results to the second random-effects, group level (Noesselt et al., 2007) to assess common PPI effects emerging from individual seeds within MGB or LGB, again using within-subject ANOVAs with four conditions (connection strength for higher- or lower-intensity stimuli with vs without sounds), at the random-effects level.

Finally, multiple regressions were used to assess any relationships between the subject-by-subject size of the critical behavioral interaction effects, in relation to BOLD activations; or separately in relation to the strength of interregional effective-connectivity (individually seeded PPI)

effects in the fMRI data. Here, one regressor was defined for the main effect of fMRI activation (or coupling strength) for higher- or lower-intensity targets with versus without sounds (interaction term), whereas another regressor was defined for the subject-specific behavioral interaction effect. Any brain–behavior relationships were computed for the whole brain with SPM, but we focused on significant regression coefficients within regions that had been independently defined by the sensory-specific blocked localizers and their conjunction, to avoid the potential problem of "double dipping" (Kriegeskorte et al., 2009). We also tested whether outliers potentially influenced the regression analysis (Nichols and Poline, 2009; Vul et al., 2009).

Results

Behavioral data

Behaviorally, we tested whether a co-occurring sound could enhance visual target detection sensitivity (d'), even though the presence of a sound carried no information about whether a visual target was present or absent (since sounds were equally likely on target and nontarget trials). Given some previous multisensory research (McDonald et al., 2000; Vroomen and de Gelder, 2000; Frassinetti et al., 2002; Noesselt et al., 2008), and past proposals associated with the possible function of the putative principle of inverse effectiveness (Stein et al., 1988; Stein and Meredith, 1993; Kayser et al., 2008), we predicted that co-occurrence of a sound should benefit detection sensitivity (d') for the lower-intensity but not higher-intensity visual targets, compared with their respective no-sound conditions.

Figure 1e plots the critical visual sensitivity scores, in formal signal detection terms (i.e., d' scores). As predicted, co-occurrence of a sound enhanced visual detection sensitivity, but only for lower-intensity not higher-intensity visual targets. This led to a significant interaction between visual intensity level and sound presence ($F_{(1,25)} = 8.97$; $p = 0.006$) in a repeated-measures ANOVA. Visual detection sensitivity (d') was affected by sound presence only for the lower-intensity visual targets ($t_{(25)} = 5.49$; $p < 0.001$).

Figure 1d shows a comparable outcome for raw accuracy data, rather than signal detection d'. Supplemental Figure S1a (available at www.jneurosci.org as supplemental material) also shows the accuracy data when separated for subjects tested inside or outside the scanner, who did not differ (as further confirmed by a mixed-effects ANOVA that found no impact of the between-subject inside/outside scanner factor, $p > 0.1$; and likewise for all our other behavioral measures). As with visual d', accuracy in the visual detection task revealed that the co-occurrence of a sound enhanced detection for the lower-intensity but not for the higher-intensity visual targets (even though the latter were not completely at ceiling). This led once again to a significant interaction between visual intensity level and sound presence ($F_{(1,25)} = 20.79$; $p < 0.001$) in a repeated-measures ANOVA. Accuracy increased only when a sound was paired with a lower-intensity visual target ($t_{(25)} = 7.16$; $p < 0.001$).

For completeness, we also analyzed criterion and reaction time measures (see supplemental Fig. S1b,c, available at www.jneurosci.org as supplemental material). Interestingly, these showed a very different pattern to our more critical measures of sensitivity (d') and hit rate. Co-occurrence of a sound merely speeded responses overall (see supplemental Fig. S1b, available at www.jneurosci.org as supplemental material) regardless of visual intensity and even of the presence/absence of a visual target (see supplemental Fig. S1b legend, available at www.jneurosci.org as supplemental material). Thus, RTs were simply faster in the presence of a potentially alerting sound, regardless of visual condition. Hence this particular RT result need not, strictly speaking,
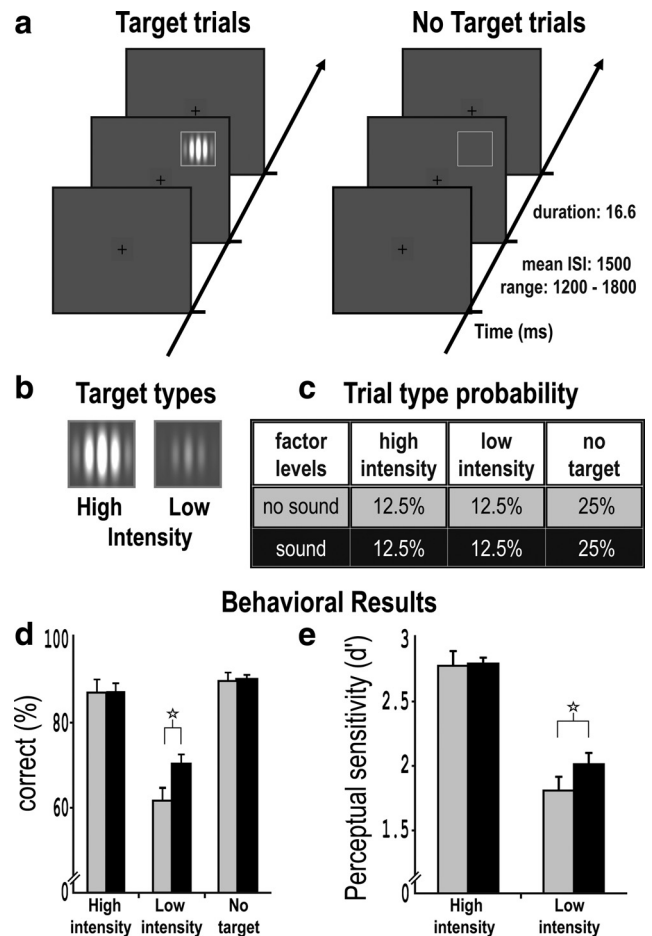

Figure 1. Experimental design and behavioral results. *a*, Schematic illustrative display sequences for single trials, shown with higher-intensity target grating either present (sequence illustrated on left) or absent (sequence illustrated on right). *b*, Schematic examples of higher- or lower-intensity targets, with absolute luminances individually defined for each subject (see main text for details). *c*, Percentage of different trial types. Sound or no-sound was equiprobable and unrelated to the visual conditions. *d*, Behavioral mean accuracy, with 95% confidence intervals for data pooled across outside-scanner and inside-scanner groups (for separate plots of subgroups inside and outside scanner, see supplemental Fig. S1a, available at www.jneurosci.org as supplemental material). The gray bars depict no-sound conditions. The black bars depict sound conditions. Performance improved significantly only for lower-intensity targets when paired with a sound (see main text), indicated by the star symbol above the pairs of values. *e*, Perceptual sensitivity (d'), as derived from signal-detection theory, for higher- or lower-intensity visual targets presented with (black bars) or without (gray bars) a co-occurring sound. Perceptual sensitivity in detecting the visual event was enhanced by a co-occurring sound only for the lower-intensity visual target (even though performance for the higher-intensity target was not at ceiling; see also accuracy data in *d*).

be considered multisensory in nature, as only the auditory factor influenced the RT pattern. In terms of possible fMRI analogs, the RT pattern would therefore correspond simply to the main effect of sound presence (which we found, as reported below, to activate auditory cortex, STS, and MGB, as would be expected). Hence we did not explore the RT effect any further. Nevertheless, we note that the overall speeding because of a sound is broadly consistent with a wide literature showing that a range of visual tasks, both manual and saccadic, can be speeded by sound occurrence (Hughes et al., 1994; Doyle and Snowden, 2001). In the past, such overall speeding by a sound has often been discussed in the context of possible nonspecific alerting effects (Posner, 1978). Some behavioral studies have found more complex RT patterns when varying both visual and auditory stimulus intensities (Marks et al., 1986; Marks, 1987), but here with a constant (rela-

tively high) auditory intensity, we found only overall speeding of manual RTs in the visual task, regardless of visual condition (see supplemental Fig. S1*b*, available at www.jneurosci.org as supplemental material).

Turning to the final behavioral measure of criterion, we found that participants adopted a higher criterion for reporting low-intensity target presence (see supplemental Fig. S1*c*, available at www.jneurosci.org as supplemental material). But since within the signal detection framework criterion is strictly independent of sensitivity, d', this criterion effect cannot contaminate our critical d' results. Moreover, the criterion effect as a function of visual intensity applied here regardless of auditory condition (see supplemental Fig. S1*c*, available at www.jneurosci.org as supplemental material), so unlike the d' and accuracy results need not be interpreted as reflecting any multisensory phenomenon.

Thus, only the critical behavioral measures of d' and accuracy showed differential multisensory effects (i.e., that depended on both auditory and visual conditions), with co-occurrence of a sound genuinely enhancing perceptual sensitivity (d') and accuracy for lower-intensity but not higher-intensity visual targets. This pattern of multisensory outcome for detection sensitivity and accuracy appears compatible with the idea, long associated with the putative PoIE for multisensory integration, that co-occurrence of events in multiple modalities might particularly benefit near-threshold detection (as for the lower-intensity, but not higher-intensity, visual targets here). Our analyses of fMRI data below test for neural consequences of the co-occurring sounds, for visual targets of lower- versus higher-intensity.

fMRI results

We used separate passive blocked fMRI localizers to predetermine potential candidate "sensory-specific" brain regions (responding to our high-intensity visual stimuli more than our auditory, or vice versa); and for determining potential candidate "heteromodal" regions (those areas responding significantly to both our auditory and our high-intensity visual stimuli, on a conjunction test). As expected, passive viewing of our high-intensity visual gratings activated left occipital cortex contralateral to the (right) visually stimulated hemifield, plus the contralateral LGB (see supplemental Table S1*a*, available at www.jneurosci.org as supplemental material); with bilateral parietal, frontal, and temporal regions also activated. Activation caused by passive listening to our auditory tones arose in bilateral MGB, plus middle temporal cortical areas including the planum temporale, Heschl's gyrus, planum polare, and extending ventrally into medial STS (see supplemental Table S1*b*, available at www.jneurosci.org as supplemental material). Finally, candidate heteromodal regions that responded significantly both to visual stimuli and also to sounds included posterior STS, plus parietal and dorsolateral prefrontal regions, all in accord with previous studies (Beauchamp et al., 2004; Noesselt et al., 2007) (see supplemental Table S1*c*, available at www.jneurosci.org as supplemental material).

Below, we present results from our main event-related fMRI experiment divided into three sections. First, we present results from a conventional group voxelwise SPM analysis of BOLD activations, supplemented by results from individually defined ROIs. Second, we present results from interregional effective-connectivity analyses of functional coupling between brain areas as a function of experimental condition. Third, we present brain–behavior regression analyses testing whether local BOLD signals in the implicated areas, or interregional coupling strength, covary with subject-by-subject psychophysical

benefits specific to combining a sound with a lower-intensity visual target.

Modulation of local BOLD response to lower-intensity (vs higher-intensity) visual events by concurrent sound

For the event-related results from the main fMRI experiment, the most important contrast concerns a greater enhancing impact of the sound on lower-intensity than higher-intensity visual targets, analogous to the behavioral effect on visual detection d' and accuracy. The critical interaction contrast is as follows: (lower-intensity light with sound) minus (lower-intensity light alone) > (higher-intensity light with sound) minus (higher-intensity light alone). This two-way interaction-contrast subtracts out any trivial effects attributable to visual intensity per se, attributable to visual-frame presentation that signaled when a response was required on every trial, or attributable to sound presence per se. See supplemental Table S2 (available at www.jneurosci.org as supplemental material) for details of the outcome for the visual-intensity or sound presence contrasts, which all turned out as expected [i.e., higher activation of visual cortex attributable to increased visual intensity (main effect of high intensity > low intensity) (supplemental Table S2*a*, available at www.jneurosci.org as supplemental material), and of auditory cortex attributable to sound presence (main effect of sound > no sound conditions) (see supplemental Table S2*b*, available at www.jneurosci.org as supplemental material)].

The critical two-way interaction contrast (as per the formula above) is analogous to the critical behavioral interaction that affected accuracy and d' (Fig. 1*d,e*), so may reveal the neural analog of the sound-induced boosting of visual processing. We interrogated the visual, auditory, and candidate heteromodal audiovisual regions (via SPM inclusive masking) that had been defined independently of the main experiment by the separate localizers. We found the critical interaction to be significant not only in STS (Fig. 2, top right; Table 1, In multisensory areas), a known multisensory brain region, but also in extrastriate visual regions contralateral to the visual target (Fig. 2, top left; Table 1, In visual areas), plus in posterior insula/Heschl's gyrus (i.e., likely to correspond with low-level auditory cortex) (Fig. 2, top middle; Table 1, In auditory areas). Note that, while showing the critical interaction effect, the response patterns within visual and auditory cortex also shows the overall modality preferences one would expect for high-intensity visual or auditory stimuli, as confirmed also by the independent localizers. Table 1 (Outside visual/auditory and multisensory areas) lists additional areas showing an interaction outside the visual, auditory, and heteromodal regions of main interest, for completeness.

The plot for the group interaction contrast within voxelwise normalized space in Figure 2, top middle plot, shows for insula/Heschl's gyrus not only the anticipated (PoIE-like) increase in response when the sound co-occurs with a low-intensity visual target but also an apparent lack of auditory response when the same sound is paired with a high-intensity visual target (although please note that zero on the *y*-axis in the plots of Fig. 2 represents the session mean, rather than absolute zero). Although in principle the latter unexpected outcome might potentially reflect subadditive responses for high-intensity pairings (cf. Angelaki et al., 2009; Sadaghiani et al., 2009; Stevenson and James, 2009; Kayser et al., 2010), alternatively it might reflect the inevitable tendency for SPM interaction contrasts to highly the most significant voxels showing the strongest interaction pattern (so in the present context, not only an enhanced response to the sound when paired with a lower-intensity visual target, but also some

reduction in this response when paired with a higher-intensity visual target, at the peak interaction voxel in SPM).

Accordingly, we next extracted the β weights from the main experiment for subject-specific, individually defined A1 ROIs, as derived from the independent localizer runs (Fig. 3, bottom), thereby circumventing any selection bias. This individual ROI analysis confirmed the interaction pattern for A1, showing significantly increased BOLD signal for low-intensity visual targets when paired with sounds versus without sounds there ($F = 8.33$, $p < 0.05$ for the interaction; *post hoc* $t = 2.3$, $p < 0.05$, for the pairwise contrast) (Fig. 3, bottom). In this more sensitive individual ROI analysis of A1, free from any voxelwise selection biases, the ROI results showed robust auditory responses from primary auditory cortex, even for the significantly reduced response found when paired with a higher-intensity visual stimulus (Fig. 3, bottom).

Importantly, the critical sound-induced enhancement of visual responses was also found subcortically in the LGB (Fig. 2, bottom left; Table 1, In visual areas) and in the MGB (Fig. 2, bottom right; Table 1, In auditory areas) in the group voxelwise SPM analysis. Thus, in addition to multi-sensory STS, not only did visual fusiform and auditory A1/Heschl's gyrus show the critical interaction pattern cortically, but so did subcortical thalamic stages of the visual and auditory pathways.

As shown in the plots of Figure 2 from the group voxelwise SPM analysis, for all of the affected areas (i.e., STS, visual cortex, auditory cortex, LGB, and MGB) the co-occurrence of a sound enhanced the BOLD response for the lower-intensity visual target condition more than for the higher-intensity visual target condition. In principle, one must consider whether the latter outcome could reflect some "ceiling" effect for the BOLD signal in the high-intensity condition. However, the bar graphs in Figure 2 show that the BOLD responses in the affected regions (with the exception of the fusiform gyrus) were typically higher for low-intensity stimuli paired with sounds, than for any of the high-intensity conditions, which thus argues against any ceiling concerns in terms of BOLD level per se. Moreover, even our higher-intensity visual stimuli had modest absolute intensities. Other work (Buracas et al., 2005) suggests that visual BOLD signals typically saturate only for much higher luminance levels than those used here. Nonetheless, we found the expected pattern of enhanced BOLD responses for higher- versus lower-intensity visual stimuli in fusiform, LGB, plus additional visual regions (see supplemental Ta-
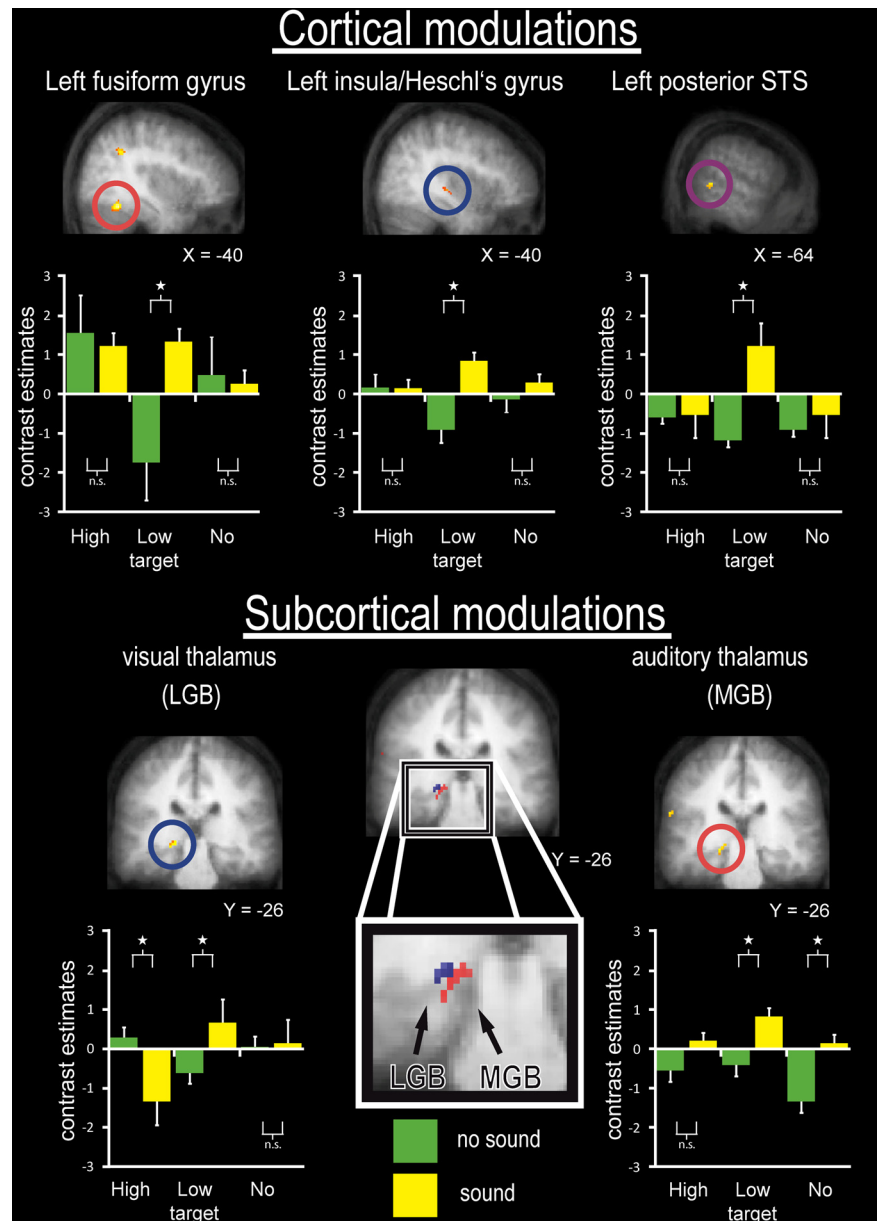

Figure 2. Group results: fMRI results specific to co-occurring sounds boosting the response for lower-intensity visual targets more than higher-intensity targets. This pattern was formally tested for with an interaction contrast: (lower-intensity visual targets with sound minus without sounds > higher-intensity visual target with sound minus without) (i.e., same interaction as for behavioral accuracy and d' effects in Figure 1, d and e). Note that this fMRI interaction term is not contaminated by any differences within visual or auditory stimulation. We further confirmed whether this interaction involved a significant increase of the low-intensity plus sound condition, relative to the low-intensity without sounds [see bar graphs, with yellow bars for sound conditions and green for no-sound; pairwise significance (or nonsignificance, n.s.) is indicated by brackets above/below pairs of values in bar graphs]. The top half of the figure shows cortical effects, within visual cortex (here fusiform gyrus) as independently identified in the localizer run (left column); and likewise for auditory cortex/insula (middle column), plus multisensory STS (right column). The bar graphs below brain sections depict the corresponding BOLD estimates (plus 95% confidence intervals), relative to the global mean as zero, for no-sound conditions in green and sound conditions in yellow, for all three visual conditions (high intensity, low intensity, no intensity) (for details, see Table 1). The bottom half of the figure shows subcortical effects in visual thalamus (shown in blue) and auditory thalamus (shown in red). Their localization as visually (LGB) or auditorily (MGB) responsive thalamus was confirmed independently by our unisensory localizer runs, and further corroborated by individual region-of-interest analyses and interregional coupling analyses (see main text and supplemental Fig. S2, available at www.jneurosci.org as supplemental material). The bar graphs again depict BOLD estimates for all six conditions, in LGB (left plot) and MGB (right plot).

ble S2, available at www.jneurosci.org as supplemental material) when presented without sounds, as expected.

One aspect of the specific pattern of BOLD responses in group-normalized LGB (Fig. 2, bottom left) may appear

Table 1. fMRI interaction effect (mirroring the behavioral benefit)

Brain region	Peak *T*	Value of *p* <	Cluster size	MNI coordinates		
				X	*Y*	*Z*
In visual areas[a]						
Occipital lobe						
Left fusiform gyrus	3.56	0.01	99	−40	−54	−20
Left superior occipital gyrus	2.86	0.01	64	−20	−68	28
Right superior occipital gyrus	2.80	0.01	42	12	−68	24
Frontal lobe						
Left inferior frontal gyrus	4.00	0.01	81	−46	24	18
Parietal lobe						
Left precuneus	3.91	0.01	339	−2	−58	60
Left inferior parietal lobule	3.25	0.01	317	−58	−58	34
Left intraparietal sulcus	3.25	0.01		−36	−50	38
Temporal lobe						
Left superior temporal sulcus	2.79	0.01	119	−52	−42	2
Subcortical						
Left lateral geniculate body	2.83	0.01	12	−20	−26	−12
In auditory areas[b]						
Temporal lobe						
Left posterior superior temporal sulcus	2.86	0.01	112	−52	−42	2
Right posterior superior temporal sulcus	2.82	0.01	134	70	−42	0
Left transverse temporal gyrus (Heschl's gyrus)	2.79	0.01	23	−40	−26	−4
Frontal lobe						
Left inferior frontal gyrus	4.00	0.01	81	−46	24	18
Right middle cingulate cortex	3.96	0.01	40	2	−16	32
Parietal lobe						
Left intraparietal sulcus	3.61	0.01	221	−36	−50	38
Subcortical						
Left medial geniculate body	2.51	0.05	19	−16	−26	−8
In multisensory areas[c]						
Temporal lobe						
Left posterior superior temporal sulcus	2.78	0.01	34	−64	−40	2
Right posterior superior temporal sulcus	2.75	0.01	51	70	−42	0
Frontal lobe						
Left inferior frontal gyrus	4.00	0.01	81	−46	24	18
Right middle cingulate cortex	3.96	0.01	40	2	−16	32
Parietal lobe						
Left angular gyrus	3.25	0.01	35	−58	−58	34
Left precuneus	3.92	0.01	22	−6	−50	46
Outside visual/auditory and multisensory areas[d]						
Parietal lobe						
Right precuneus	3.38	0.01	45	6	−52	64
Frontal lobe						
Left anterior cingulate cortex	3.05	0.01	65	−2	10	28
Subcortical						
Right parahippocampal gyrus	3.85	0.01	56	24	−24	−18

The table provides local maxima for the critical interaction involving sound presence and visual intensity, namely (low intensity with sound minus low intensity without sound) > (high intensity with sound minus high intensity without sound).

[a]Interaction effect in visual areas (cf. supplemental Table S1*a*, available at www.jneurosci.org as supplemental material) thresholded at *p* < 0.01 (FDR-corrected for multiple comparisons within inclusive masks) (see supplemental Table S1, available at www.jneurosci.org as supplemental material); *k* > 20 and (*) *k* > 5 in case of LGB.

[b]Interaction effect in auditory areas (cf. supplemental Table S1*b*, available at www.jneurosci.org as supplemental material) thresholded at *p* < 0.01 (FDR-corrected for multiple comparisons within inclusive masks) (see supplemental Table S1, available at www.jneurosci.org as supplemental material); *k* > 20 and (*) *k* > 5 in case of MGB.

[c]Interaction effect in candidate heteromodal areas (cf. supplemental Table S1*c*, available at www.jneurosci.org as supplemental material) thresholded at *p* < 0.01 (FDR-corrected for multiple comparisons within inclusive masks) (see supplemental Table S1, available at www.jneurosci.org as supplemental material); *k* > 20.

[d]Interaction effect outside any of the masks (i.e. beyond those regions predefined as visual, auditory, or candidate-heteromodal by our localizers), thresholded at *p* < 0.001, *k* > 40 (since no a priori hypotheses applied for those).

somewhat counterintuitive, with an apparent decrease for high-intensity visual stimuli when paired with sounds (as had been found in group results for the interaction pattern in A1 above).

To address this point and also provide additional validation of the novel results at the thalamic level, we again supplemented the group voxelwise analyses by identifying visual and auditory thalamic body ROIs in each individual subject (supplemental Fig. S2, available at www.jneurosci.org as supplemental material) (see above for the rationale of using complementary ROI analyses). We then assessed the experimental effects in these individual thalamic ROIs (i.e., which could now correspond to somewhat dif-ferent voxels in different subjects, but for the same defined area). Since these ROIs were defined independently of the main experiment, via the separate localizers, this again circumvents any selection bias for the interaction contrast. This individual approach corroborated our group voxelwise SPM results, while also removing the one unexpected outcome (for the LGB) with this more unbiased ROI approach. Again, we found enhanced BOLD signals when a sound is added to a lower-intensity visual target, but no significant change in response when the same sound is added to a higher-intensity visual target (for LGB ROI results from all subjects, see Fig. 3, top left; for a confirmatory analysis on the subset of subjects who showed the most unequivocal LGB and

MGB localization, see supplemental Fig. S5, top plot, available at www.jneurosci. org as supplemental material).

The pattern of activity for the fusiform interaction peak in the voxelwise SPM analysis (Fig. 2, top left) showed one unexpected trend, namely a tendency for lower activation in the absence of a sound for a low-intensity visual target versus none. But this trend was nonsignificant ($p > 0.1$) so need not be considered further. In any case, it may again simply reflect a selection bias for peaks of SPM interaction contrasts to highlight voxels that show apparent "crossover" patterns, as explained above for other regions.

To summarize the BOLD activation results so far (Figs. 2, 3), predefined (inclusively masked) heteromodal cortex in STS showed a pattern of enhanced BOLD signal by co-occurring sounds only for lower-intensity but not higher-intensity visual targets (thus analogous to the impact on visual detection d' found behaviorally) (compare Fig. 1e). Similar patterns were found in sensory-specific visual cortex, in sensory-specific insula/Heschl's gyrus, and even in sensory-specific thalamus (LGB and MGB). Additional analysis of individually defined ROIs confirmed the interaction pattern in primary auditory cortex, MGB, and LGB (Fig. 3; supplemental Fig. S5, available at www. jneurosci.org as supplemental material). This ROI analysis also confirmed that the few unexpected aspects of the interaction results from group-normalized space (i.e., apparent crossover interaction pattern for LGB; apparent loss of auditory response in presence of higher-intensity visual targets for insula/Heschl's gyrus) were no longer evident for the more sensitive individual analyses of independently defined ROIs. Those unexpected aspects of the group-normalized results should thus be treated with caution. By contrast, all of our critical activations were found robustly in the individual ROIs, as well as for the voxelwise group-normalized analysis.

Although the observed pattern for MGB and A1 ROIs (Fig. 3, top right and bottom) was not identical, both showed the critical interaction, with strongest responses when the sound was paired with a low-intensity visual target. It appears that the MGB ROI tended to be somewhat more responsive to high-intensity visual targets in the absence of sound (albeit only as a nonsignificant trend) than for A1. This may reflect the fact that some subnuclei within the MGB receive visual inputs (Linke et al., 2000), given that the BOLD signal will aggregate across different subnuclei; and/or it could reflect possible feedback signals from heteromodal STS.

Mechanistically, on the level of neuronal firing rates audiovisual integration may be rather complex, as different frequency bands of neural response can be differentially modulated by audiovisual stimuli (e.g., in the STS) (Chandrasekaran and
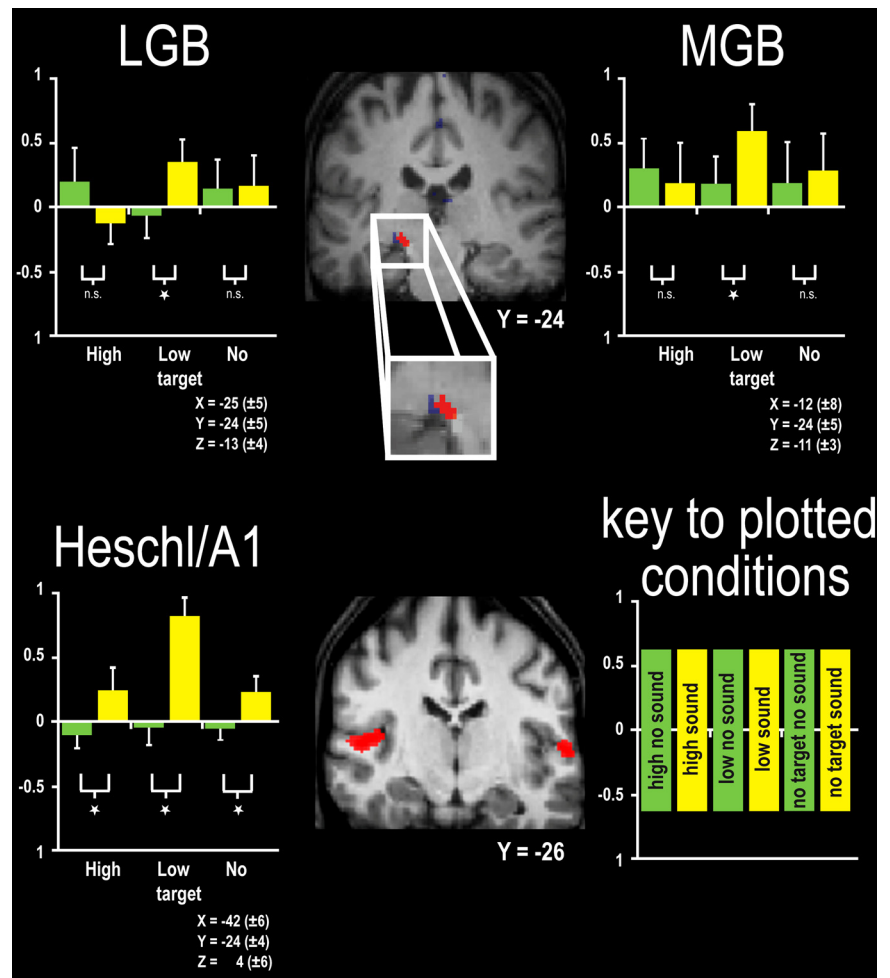

Figure 3. fMRI responses in subject-specific, individually defined visual and auditory thalamus plus A1/Heschl's gyrus. Top, Brain section depicts BOLD effects in visual (blue) and auditory (red) thalamus for one illustrative individual subject. The bar graphs on either side of brain sections depict the BOLD estimates for all six experimental conditions as derived from subject-specific ROIs for visual (left bar graphs) and auditory thalamus (right bar graph), as defined by separate localizers combining functional and anatomical criteria (see Materials and Methods), with the extracted signals from subject-specific regions then averaged across all subjects (shown with 95% confidence intervals), unlike the voxelwise averaging of Figure 2. Note once again the increased signal when a low-intensity visual target was paired with a sound. See also supplemental Figure S2 (available at www.jneurosci.org as supplemental material) for subject-specific LGB and MGB localization in each of the individual 12 subjects. Bottom, Brain section depicting BOLD effects in auditory cortex for one illustrative individual subject. The bar graph on left of brain section depicts the BOLD estimates for all six experimental conditions as derived from subject-specific ROIs for Heschl's gyrus, as defined by separate localizers combining functional and anatomical criteria (see Materials and Methods), with the extracted signals from subject-specific regions then averaged across all subjects (shown with 95% confidence intervals), unlike the voxelwise averaging of Figure 2. See also supplemental Figure S3 (available at www.jneurosci.org as supplemental material) for subject-specific A1 localization in each of the individual 12 subjects.

Ghazanfar, 2009). More generally, one mechanism potentially underlying multisensory integration in the time-frequency domain was proposed by Schroeder/Lakatos and colleagues in recent influential work (Lakatos et al., 2007, 2008; Schroeder et al., 2008; Schroeder and Lakatos, 2009) that primarily concerned tactile-auditory situations, rather than audiovisual as here. Tactile stimulation can phase reset neural signaling in auditory cortex, thereby enhancing response to synchronous auditory inputs. Moreover, some overlapping audiotactile representations in the thalamus (Cappe et al., 2009) have now been reported, as have learning-induced plastic changes of auditory and tactile processing attributable to (musical) training (Schulz et al., 2003; Musacchia et al., 2007). Related effects might conceivably impact on an audiovisual situation like our own, but potential
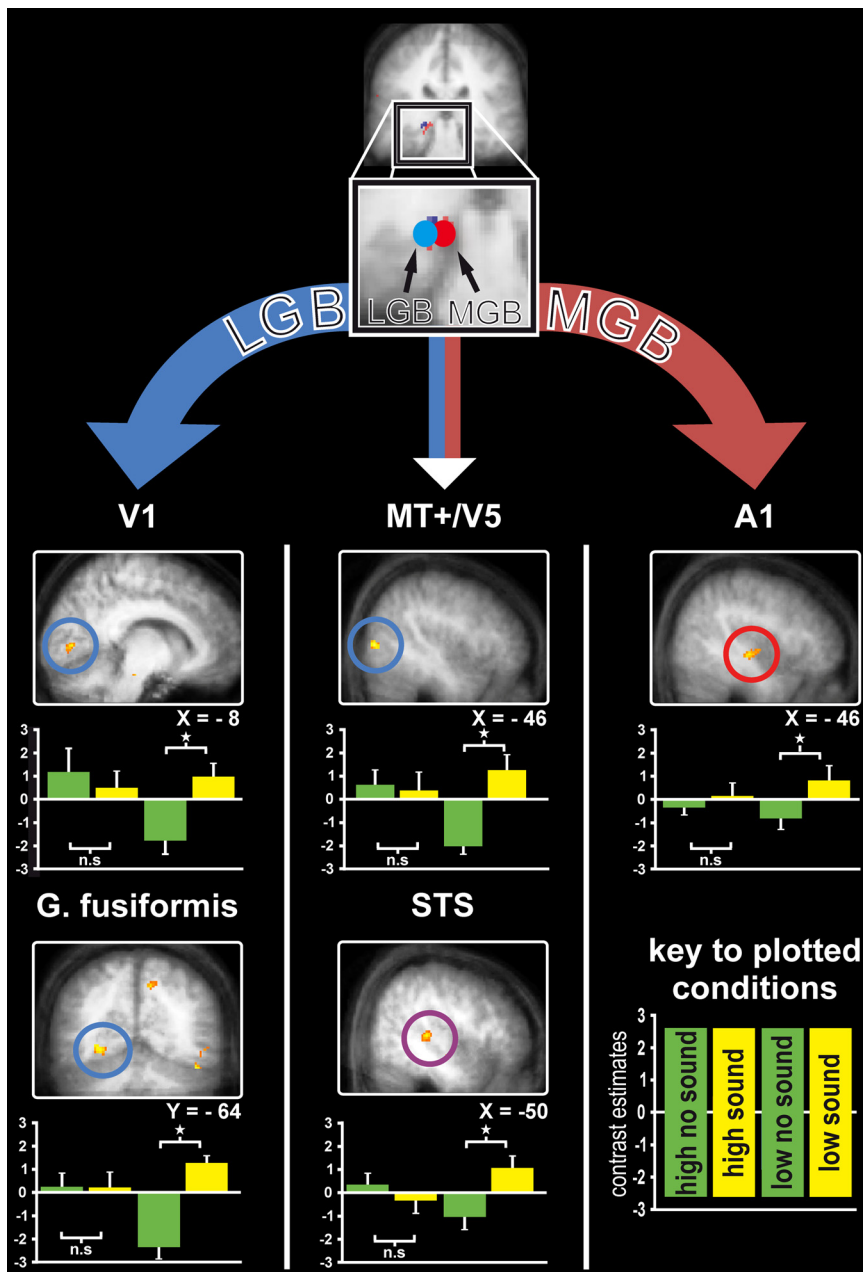
Figure 4. Enhanced interregional functional coupling for lower-intensity visual targets when paired with a concurrent sound (vs without a sound), relative to higher intensity targets with sound minus without sounds (i.e., analogous interaction pattern as for d' in Fig. 1e). Top, Separate seeds for our PPI analyses of interregional coupling are schematically illustrated by red and blue circles overlaid on activation in sensory-specific left visual thalamus (blue spot) and left auditory thalamus (red spot). Enhanced functional coupling (specifically for the interaction test as stated above) with LGB seed is revealed for left calcarine fissure and left fusiform gyrus (shown in left column), and with MGB seed for left low-level auditory cortex (shown in right column). The LGB and MGB seeds for PPI analyses thus differentially highlighted visual or auditory cortex, respectively, for condition-dependent coupling, thus providing additional confirmation of their visual or auditory status, respectively. Middle column, The different seed sites (LGB or MGB) nevertheless also produced some overlap in functional coupling with common remote regions for the two different seeds, as found when now testing for the conjunction of seeding the PPI analysis in either. The common remote coupling effects were observed in medial STS (bottom middle panel) and also in occipital gyrus in possible vicinity of human MT + (top middle panel). See Table 2 for details.

phase resetting would seem to require visual signals to precede auditory signals sufficiently to overcome the different transduction times (Musacchia et al., 2006; Schroeder et al., 2008; Schroeder and Lakatos, 2009). This seems somewhat unlikely for the present concurrent audiovisual pairings. To our knowledge, the earliest impact of concurrent visual stimuli on auditory

event-related potential components has been found to emerge at ~50 ms after stimulus, well beyond the initial phase of auditory processing (Giard and Peronnet, 1999; Molholm et al., 2004). We note that Lakatos et al. (2009) report phase resets in macaque auditory cortex attributable to visual stimulation only after the initial activation.

The modulations we observe in visual cortex (and LGB) might in principle reflect phase resetting there [cf. Lakatos et al. (2008) for attention-related phase resetting of visual cortex], and/or involve projections from auditory or multisensory cortex, which serve to increase the signal-to-noise ratio for the trials pairing a concurrent sound with low-intensity visual targets. In accord, Romei et al. (2007) recently reported an enhancement of transcranial magnetic stimulation (TMS)-induced "phosphene perception" when sounds were combined with near-threshold TMS over visual cortex. This phase resetting could potentially be the underlying mechanism of our regional fMRI effects and may reflect the functional coupling of distant brain regions.

Effects specific to pairing lower-intensity (vs higher-intensity) visual stimuli with sound on interregional effective connectivity

To assess functional coupling between brain regions, we next tested for potential condition-dependent changes in "effective connectivity" between areas (i.e., interregional coupling) for the affected thalamic bodies with cortical sensory-specific and heteromodal structures. Note that possible changes in interregional coupling are logically distinct from effects on local BOLD activations as described above, so can produce a different outcome. We tested for interregional coupling using the relatively assumption-free PPI approach (Friston et al., 1997). We seeded the PPI analyses in (individually defined) left LGB or left MGB, and tested for enhanced "coupling" with other regions, which arose specifically in the context of a lower- rather than higher-intensity visual target being paired with a sound (i.e., analogous interaction pattern to that found for behavioral sensitivity, d'; and for local BOLD activations above; but now testing for analogously condition-dependent changes in the strength of functional interregional coupling, rather than for local activations as in the preceding fMRI results section).

We found such enhanced coupling for the critical interaction effect (Fig. 4; Table 2, LGN specific coupling, LGB specific coupling) between left LGB with ipsilateral occipital areas including

Table 2. Effects for psychophysiological interaction (PPI analysis) [(low intensity visual with sound minus low intensity without sound) > (high intensity visual with sound minus high intensity without sound)]

Brain region	Peak T	Value of $p <$	Cluster size	MNI coordinates		
				X	Y	Z
LGN-specific coupling in visual areas[a]						
Occipital lobe						
Left middle occipital gyrus (MT+/hV5)	4.48	0.001	334	−50	−80	4
Left fusiform gyrus	3.53	0.01	190	−26	−64	−12
Left calcarine gyrus (V1)	3.53	0.01	56	−8	−84	6
Parietal lobe						
Left anterior cingulate cortex	3.05	0.01	65	−2	10	28
Frontal lobe						
Right inferior frontal gyrus	4.37	0.001	394	38	24	8
Left middle frontal gyrus	3.36	0.01	73	−24	0	56
Left inferior frontal gyrus	2.85	0.01	179	−34	18	6
LGB-specific coupling in auditory areas[b]						
Temporal lobe						
Right temporal lobe	3.40	0.01	21	38	6	−28
Left temporal lobe	3.20	0.01	25	−36	4	−26
LGB-specific coupling in multisensory areas[c]						
Frontal lobe						
Right inferior frontal gyrus	3.24	0.01	23	54	22	14
MGN-specific coupling in auditory areas[d]						
Temporal lobe						
Left anterior transverse gyrus (Heschl)	4.83	0.01	250	−46	−24	−6
Right planum polare	4.12	0.01	213	46	−16	−16
Left superior temporal gyrus	2.87	0.01	72	−44	−38	22
MGN-specific coupling in visual areas[e]						
Occipital lobe						
Left middle occipital gyrus (MT+/hV5)	3.74	0.01	63	−48	−78	2
MGN-specific coupling in multisensory areas[f]						
No significant effects						
Overlap LGB- and MGN-specific coupling in visual areas[g]						
Occipital lobe						
Left middle occipital gyrus (MT+/hV5)	3.89	0.01	63	−48	−80	0
Temporal lobe						
Left superior temporal gyrus	3.72	0.01	29	−46	−38	20
Frontal lobe						
Left anterior cingulate gyrus	2.81	0.01	51	−8	24	30
Overlap LGB- and MGN-specific coupling in auditory areas[h]						
Temporal lobe						
Right superior temporal sulcus	4.12	0.01	28	46	−16	−16
Left superior temporal sulcus	2.99	0.01	23	−46	−22	−6
Overlap LGB- and MGN-specific coupling in multisensory areas[i]						
No significant effects						

The table provides local maxima for the functional coupling outcome, with either an LGB or MGB seed, for the PPI specifically testing for higher condition-dependent coupling as follows: (low intensity visual with sound minus low intensity without sound) > (high intensity visual with sound minus high intensity without sound).

[a,b,c]Local maxima for LGB-seed-specific coupling in [a]predefined visual areas, [b]auditory areas, and [c]candidate heteromodal areas, all with PPI thresholded at $p < 0.01$ (FDR-corrected for multiple comparisons within inclusive masks) (see supplemental Table S1, available at www.jneurosci.org as supplemental material); $k > 20$.

[d,e,f]Local maxima for MGB-seed-specific coupling in [d]predefined visual areas, [e]auditory areas, and [f]candidate heteromodal areas all thresholded at $p < 0.01$ (FDR-corrected for multiple comparisons within inclusive masks) (see supplemental Table S1, available at www.jneurosci.org as supplemental material); $k > 20$.

[g,h,i]Local maxima for the overlap (conjunction) of LGB- and MGB-seeded coupling in [g]predefined visual areas, [h]auditory areas, and [i]candidate multisensory areas, all thresholded at $p < 0.01$ (FDR-corrected for multiple comparisons within inclusive masks) (see supplemental Table S1, available at www.jneurosci.org as supplemental material); $k > 20$.

primary visual cortex (consistent with the visual nature of the LGB, and thus providing additional confirmation of that functional localization). Analogously, we also found such condition-dependent enhanced coupling of left MGB with ipsilateral Heschl's gyrus (consistent with the auditory nature of MGB) (Fig. 4; Table 2, MGN specific coupling). Beyond these sensory-specific coupling results for LGB or MGB seeds, we also found enhanced coupling of both MGB and LGB with STS and putative MT+ (Campana et al., 2006; Eckert et al., 2008) (Fig. 4; Table 2, Overlap LGB and MGN specific coupling). This enhanced coupling (i.e., higher covariation in residuals) with STS and putative MT+ was again specific to the context of a sound being paired with a lower-intensity (rather than higher-intensity) visual target (i.e., to the very condition that had led to the critical behavioral enhancements of d' and accuracy).

These new effects for MGB are in line with anatomical and electrophysiological studies reporting that subnuclei within the MGB receive some visual inputs (Linke et al., 2000) and can respond to visual stimulation (Wepsic, 1966; Benedek et al., 1997; Komura et al., 2005) and demonstrations that the MGB is connected with STS (Burton and Jones, 1976; Yeterian and Pandya, 1989). To our knowledge, no direct connections of auditory regions nor of STS with LGB have been reported to date, although there is some evidence for direct connections of LGB with extrastriate regions (Yukie and Iwai, 1981). Alternatively, the observed modulations in LGB and its condition-dependent coupling with other areas might in principle potentially involve early visual cortex, which is anatomically linked with posterior STS (Falchier et al., 2002; Ghazanfar et al., 2005; Kayser and Logothetis, 2009) and reciprocally connected with LGB.

Thus far, we have shown that (1) co-occurrence of a sound significantly enhances perceptual sensitivity (d') and detection accuracy for a lower-intensity but not a higher-intensity visual target, in apparent accord with the principle of inverse effectiveness; (2) that a related interaction pattern is observed for BOLD activations in STS, visual cortex (plus LGB), and auditory cortex (plus MGB); (3) we also find a logically analogous interaction pattern for interregional coupling. Specifically, on this latter point, we found enhanced coupling of the two thalamic sites (LGB or MGB) with their respective sensory-specific cortices, and also between both of these thalamic sites and STS (plus lateral occipital cortex possibly corresponding to MT+), for the particular context that led to enhanced behavioral sensitivity (i.e., with this effective connectivity being most pronounced when a lower-intensity visual target is paired with a sound).

Relationship of subject-by-subject behavioral benefits to increased local brain activations and (separately) to increased strength of interregional coupling specifically when a lower- rather than higher-intensity visual target is paired with a sound

To test for an even closer link between brain activity and behavior, we next assessed whether our independently localized brain regions (i.e., the visually responsive, auditorily responsive, and candidate heteromodal areas identified by the separate blocked localizers) showed BOLD signals for which the critical interaction pattern correlated with subject-by-subject behavioral benefits for the impact of adding sound to a lower- rather than higher-intensity visual target. We first tested for subject-by-subject brain-behavior relations for the regional BOLD activations (i.e., for the basic contrasts of conditions). We regressed the subject-specific BOLD interaction differences against the analogous behavioral difference. This revealed significant subject-by-subject brain–behavior regression coefficients in left visual cortex, contralateral to the visual target (Fig. 5a; Table 3, In visual areas), plus left auditory cortex (Fig. 5b; Table 3, In auditory areas) and heteromodal STS (Fig. 5c; Table 3, In multisensory areas). See also supplemental Figure S4 (available at www.jneurosci.org as supplemental material) for a confirmatory brain–behavior regression with behavioral outliers removed from the analysis.

Next, we tested whether changes in the interregional coupling of these thalamic structures with cortical areas (analogous to Fig. 4) might relate to the subject-by-subject behavioral interaction outcome. When weighting the PPI analyses (seeded either in left LGB or left MGB, as individually defined) by the parametric, subject-by-subject size of the critical behavioral interaction, we found that both the LGB and the MGB independently showed stronger enhancement of coupling with bilateral STS (Fig. 5d, Table 4) in the specific context of the sound-plus-lower-intensity-light condition, in relation to the impact on performance. The outcome of this weighted PPI analysis reveals that these thalamic-cortical neural coupling effects (for LGB–STS, and also separately replicated for MGB–STS) have some parametric relationship to the corresponding behavioral effect in psychophysics.

When taken together, the different aspects of our fMRI results clearly identify a functional corticothalamic network of visual, auditory, and multisensory regions. These regions are activated more strongly, and become more functionally integrated as shown by the interregional coupling data, when a task-irrelevant sound co-occurs with a lower-intensity visual target, the very same condition that led behaviorally to enhanced d' and hit rate in the visual detection task. This link is further strengthened by the brain–behavior relationships we observed.
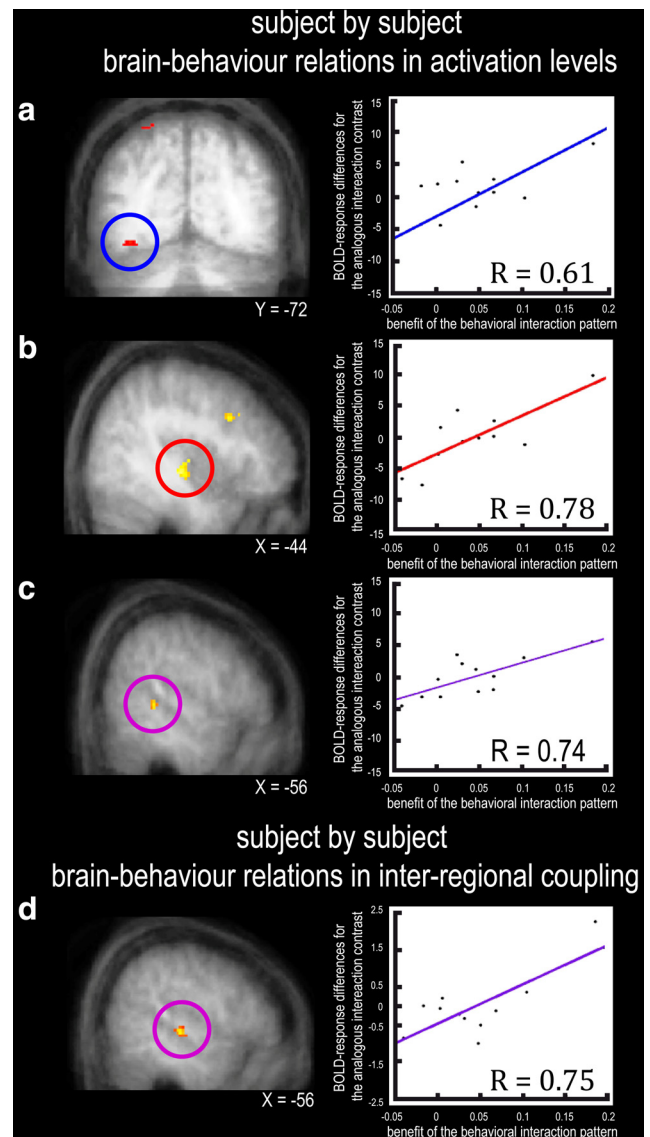

Figure 5. Subject-by-subject brain–behavior relationships. Within the independently localized regions of interest, significant regression coefficients were found between the size of the behavioral interaction pattern (i.e., difference in subjects' visual detection hit rate for sound minus no-sound conditions being more pronounced for lower- than higher-intensity visual targets, as shown along y-axis in scatterplots), with differences in BOLD response [for the analogous interaction contrast: (lower intensity light with sound) minus (lower intensity light alone) > (higher intensity light with sound) minus (higher intensity light alone)]. This arose in visual (**a**), auditory (**b**), and candidate heteromodal (**c**) cortex. The scatterplots on right depict the subject-specific means of corresponding fMRI differences for each participant ($n = 12$) plotted against the subject-specific behavioral interaction effect on hit rate. **d** depicts significant changes in LGB-and-MGB-seeded (conjunction) interregional coupling-strength (PPI) with remote regions, against size of behavioral interaction. This analysis highlights stronger coupling of both LGB and MGB with multisensory STS plus auditory cortex, in parametric relation to the subject-by-subject size of the critical behavioral interaction.

Discussion

Behaviorally, we found that co-occurrence of a sound increased accuracy and enhanced sensitivity (d') for detection of lower-intensity but not higher-intensity visual targets. This psychophysical outcome provides some new evidence apparently consistent with inverse effectiveness proposals for multisensory integration (Meredith and Stein, 1983; Stein et al., 1988; Stein and Meredith, 1993; Calvert and Thesen, 2004). Although other recent studies have shown some auditory influences on visual performance or d' (McDonald et al.,

Table 3. Correlation of subject-specific behavioral effects with BOLD response

Brain region	Peak T	Value of p <	Cluster size	MNI coordinates		
				X	Y	Z
In visual areas[a]						
Occipital lobe						
Left fusiform gyrus	4.25	0.01	50	−36	−72	−20
Parietal lobe						
Left superior parietal lobule	6.03	0.01	71	−36	−64	60
Temporal lobe						
Left superior temporal sulcus	6.37	0.01	46	−66	−40	0
Right superior temporal sulcus	4.47	0.01	84	66	−44	−6
Frontal lobe						
Right middle frontal gyrus	5.64	0.01	29	52	26	40
Left middle frontal gyrus	4.82	0.01	116	−54	20	26
In auditory areas[b]						
Temporal lobe						
Right superior temporal gyrus	8.32	0.01	21	60	−2	−10
Left anterior transverse gyrus (Heschl)	5.32	0.01	61	−44	−18	−6
Left superior temporal gyrus	4.82	0.01	43	−58	−48	2
Right medial temporal pole	4.71	0.01	37	40	6	−30
Frontal lobe						
Right middle frontal gyrus	5.64	0.01	29	52	26	40
In multisensory areas[c]						
Temporal lobe						
Left superior temporal sulcus	6.37	0.01	27	−56	−42	2
Right superior temporal sulcus	4.56	0.01	54	70	−42	−4
Parietal lobe						
Left superior parietal lobule	6.03	0.001	21	−36	−64	60
Frontal lobe						
Right middle frontal gyrus	4.98	0.001	25	44	6	52
Outside visual/auditory and multisensory areas[d]						
Frontal lobe						
Middle cingulate cortex	11.78	0.001	240	0	−26	50

The table provides local maxima for the regression of subjects' differential behavioral performance with differential BOLD responses (as defined by the interaction contrast) (see main text and Table 2).

[a,b,c]Local maxima for BOLD/behavior regression in predefined [a]visual, [b]auditory, [c]candidate-heteromodal areas, all thresholded at $p < 0.01$ (FDR-corrected for multiple comparisons within inclusive masks) (see supplemental Table S1, available at www.jneurosci.org as supplemental material); $k > 20$.

[d]Local maxima for BOLD/behavior regression outside predefined auditory, visual, and candidate heteromodal areas, thresholded at $p < 0.001$; $k > 40$ (since no a priori hypotheses applied for those).

Table 4. Correlation of PPIs of LGB and MGN with subject-specific behavioral effects

Brain region	Peak T	Value of p <	Cluster size	MNI coordinates		
				X	Y	Z
In visual areas[a]						
Parietal lobe						
Left superior parietal lobule	4.39	0.01	39	−24	−66	48
Frontal lobe						
Right superior medial gyrus	4.02	0.01	39	4	36	44
In auditory areas[b]						
Temporal lobe						
Left posterior insula	4.10	0.01	32	−46	−36	10
Left superior temporal gyrus	3.82	0.01	21	−56	−26	10
Frontal lobe						
Right inferior frontal gyrus	4.68	0.01	33	52	22	22
In multisensory areas[c]						
Temporal lobe						
Left superior temporal sulcus	4.10	0.01	31	−56	−26	−2
Frontal lobe						
Right precentral gyrus	3.69	0.01	24	50	4	50

The table provides local maxima for the regression of subjects' differential behavioral performance with the overlap of LGB- and MGB-specific coupling (PPI) (see Table 2 for details of that).

[a]Local maxima for overlap-PPI/behavior regression in predefined visual areas thresholded at $p < 0.01$ (FDR-corrected for multiple comparisons within inclusive masks) (see supplemental Table S1, available at www.jneurosci.org as supplemental material); $k > 20$.

[b]Local maxima for overlap-PPI/behavior regression in auditory areas thresholded at $p < 0.01$ (FDR-corrected for multiple comparisons within inclusive masks) (see supplemental Table S1, available at www.jneurosci.org as supplemental material); $k > 20$.

[c]Local maxima for overlap-PPI/behavior regression in predefined candidate heteromodal areas thresholded at $p < 0.01$ (FDR-corrected for multiple comparisons within inclusive masks) (see supplemental Table S1, available at www.jneurosci.org as supplemental material); $k > 20$.

2000; Frassinetti et al., 2002; Noesselt et al., 2008), following on from pioneering audiovisual studies that did not study sensitivity per se (Marks et al., 1986, 2003), here we specifically showed that only lower-intensity targets benefited in visual-detection d' from co-occurring sounds, whereas higher-intensity did not, thereby setting the stage for our fMRI study.

Our fMRI data provide evidence that both sensory-specific and heteromodal brain regions (as defined by independent fMRI "localizers") showed the same critical interaction in the main event-related fMRI experiment. Thus, STS, visual cortex (and LGB), auditory cortex (and MGB), all showed enhanced BOLD responses when a sound was added to a lower-intensity visual target, but not when the same sound was added to a higher-intensity visual target instead (for which any trends were if anything suppressive instead). The present findings of audiovisual multisensory effects that influence not only heteromodal STS but also auditory and visual cortex accord well with several other recent studies (Ghazanfar and Schroeder, 2006; Kayser and Logothetis, 2007; Kayser et al., 2007; Noesselt et al., 2007; Watkins et al., 2007). Moreover, recent human fMRI studies and recordings in macaque STS and low-level auditory cortex also suggest that the multisensory principle of inverse effectiveness may apply there for some audiovisual situations (Ghazanfar et al., 2008; Stevenson and James, 2009). Here, we extend the possible remit of this principle to implicate sensory-specific visual and auditory thalamus (LGB and MGB), in multisensory effects that can evidently relate to the conditions determining subject-specific psychophysical detection sensitivity.

This expands the recently uncovered principle that cross-modal interplay can affect sensory-specific cortices (Calvert et al., 2000; Macaluso et al., 2000; McDonald et al., 2003; Macaluso and Driver, 2005; Driver and Noesselt, 2008; Fuhrmann Alpert et al., 2008; Wang et al., 2008), to encompass thalamic levels of sensory-specific pathways also [see also Musacchia et al. (2006), albeit while noting that the V-brainstem potentials they measured in a speech task cannot be unequivocally attributed to specific thalamic structures; and Baier et al. (2006) for some nonspecific thalamic modulations]. Here, we find that audiovisual interplay can be observed even at the level of sensory-specific thalamic LGB and MGB. This might potentially arise because of feedback influences (see below) and/or through feedforward thalamic interactions that may guide some of the earlier audiovisual interaction effects (cf. Giard and Peronnet, 1999; Molholm et al., 2002; Brosch et al., 2005). Although subcortically we did not find any significant impact on the human superior colliculus, this might reflect some of the known fMRI limitations for that particular structure (Sylvester et al., 2007).

We also performed analyses of effective connectivity [i.e., interregional functional coupling (as distinct from local regional activation)] for the fMRI data. We found that LGB and MGB showed stronger interregional coupling (residual covariation) with their associated sensory-specific cortices (visual or auditory, respectively), for the particular context of a sound paired with a low-intensity visual target. This enhanced coupling between visual or auditory thalamus and their respective sensory cortices further confirms our ability to separate and dissociate those visual and auditory thalamic structures (as also indicated by the blocked localizers, and by our individual analyses). This aspect of the effective-connectivity pattern goes beyond other recent results showing cross-modal influences that involve sensory-specific auditory cortex (Brosch et al., 2005; Ghazanfar et al., 2005, 2008; Noesselt et al., 2007; Kayser et al., 2008) or visual cortex (Noesselt et al., 2007). An additional notable finding from our coupling results was that the separate analyses seeded in either individually defined LGB or MGB both independently revealed enhanced coupling with the heteromodal STS for the same particular context (i.e., stronger coupling when a sound was paired with a lower-intensity visual target, and thus when sensory detection was enhanced). The observed effective-connectivity patterns might potentially serve to enhance the unisensory features of a bound multisensory object, enhancing an otherwise weak representation in one modality by means of the co-occurrence with a bound strong event in another modality, consistent with the impact on unimodal visual detection sensitivity here from the co-occurring sound.

To demonstrate an even closer link between the audiovisual effect found psychophysically (i.e., enhanced detection for lower-intensity visual targets when paired with a sound) and the fMRI data, we tested for regions showing subject-by-subject brain–behavior relationships. Within our independently defined visually selective, auditorily selective, or heteromodal areas, brain–behavior relationships for BOLD activation were found for STS, auditory cortex, and for visual cortex contralateral to the targets. Moreover, the functional coupling of LGB and MGB with heteromodal STS varied parametrically with the subject-by-subject size of the critical cross-modal behavioral pattern (Fig. 5d).

The visual cortical effects here were consistently contralateral to the low-intensity visual target and consistently highlighted the fusiform gyrus. Additionally, our connectivity analyses revealed enhanced connectivity of LGN with V1 and a lateral occipital region (Fig. 4, bottom left) whose MNI coordinates correspond

reasonably well with putative V5/MT+, as described in many previous purely visual studies (Campana et al., 2006). Although our main conclusions do not specifically depend on identifying this region as true MT+, that appears consistent with reports that MT+ may be particularly involved in detection of low-contrast visual stimuli (Tootell et al., 1995), as for the lower-intensity visual targets here. It might also relate to reports that MT+ may show some auditory modulations of its visual response (Calvert et al., 1999; Amedi et al., 2005; Beauchamp, 2005; Ben-Shachar et al., 2007; Eckert et al., 2008).

Our cross-modal findings arose even though only the visual modality had to be judged here (cf. McDonald et al., 2000; Busse et al., 2005; Stormer et al., 2009). Future extensions of our paradigm could test whether these cross-modal effects are modulated when attention to modality is varied (cf. Busse et al., 2005). The possible impact of attentional load could also be of interest (cf. Lavie, 2005).

The new effects we found for human sensory-specific visual and auditory thalamus highlight the importance of thalamocortical connectivity. A role for nonspecific thalamocortical loops in multisensory integration had been hypothesized recently, with some evidence from animal studies (Fu et al., 2004; Hackett et al., 2007; Lakatos et al., 2007; Reches and Gutfreund, 2009) or human fMRI (Baier et al., 2006). Some early interactions have recently been uncovered for rat MGB with audiovisual stimuli (Komura et al., 2005), as have effects for the tectothalamic pathway in owls (Reches and Gutfreund, 2009), whereas other physiological evidence in rodents points at the potentially multisensory nature of subnuclei within the MGB (Wepsic, 1966; Bordi and LeDoux, 1994; Benedek et al., 1997) (for review, see Budinger et al., 2006, 2008). At a very different level to human fMRI data, one possible mechanism for audiovisual interplay between thalamic and cortical structures might involve calbindin-positive neurons within the thalamus, which could directly link thalamic with cortical structures (Jones, 2001; Lakatos et al., 2007). Alternatively (or in addition), the present thalamic modulations could potentially reflect feedback mechanisms from cortical on thalamic regions (O'Connor et al., 2002), to subserve the audiovisual interactions found. Feedback signals may selectively enhance visual stimulus representations even within LGB (O'Connor et al., 2002). Future invasive animal studies are needed to shed light on the temporal dynamics of multisensory interplay within this thalamocortical network, whereas electrophysiological measures in humans might also prove useful (Musacchia et al., 2006). But the present data already indicate a link between multisensory behavioral benefits in visual detection, found specifically for lower-intensity visual targets paired with a sound, to modulation of sensory-specific auditory and visual thalamus. Moreover, we found enhanced interregional coupling of LGB and MGB with cortical STS, which scaled with behavioral performance, emphasizing the close interplay of subcortical and cortical processing in multisensory integration (Jiang et al., 2007) and the relevance of this for behavioral impacts on perceptual sensitivity.

In conclusion, we uncovered a network of sensory-specific and multisensory regions in the human brain in relation to auditory-visual interactions. We were able to link this network closely to the psychophysical phenomenon of enhanced detection sensitivity for lower- but not higher-intensity visual targets when co-occurring with a sound. Visual and auditory thalamic nuclei were more activated when a sound co-occurred with a lower-intensity (but not a higher-intensity) visual target and also coupled more strongly with the affected cortical regions in this context. Together, these results implicate thalamocortical interplay in enhanced perceptual sensitivity (d') attributable to mul-

tisensory interactions. They demonstrate in humans that cross-modal influences on sensory-specific brain structures can extend even to sensory-specific visual and auditory thalamus, with these influences relating to psychophysical performance.

References

Amedi A, von Kriegstein K, van Atteveldt NM, Beauchamp MS, Naumer MJ (2005) Functional imaging of human crossmodal identification and object recognition. Exp Brain Res 166:559–571.

Andersen TS, Mamassian P (2008) Audiovisual integration of stimulus transients. Vision Res 48:2537–2544.

Angelaki DE, Gu Y, DeAngelis GC (2009) Multisensory integration: psychophysics, neurophysiology, and computation. Curr Opin Neurobiol 19:452–458.

Baier B, Kleinschmidt A, Müller NG (2006) Cross-modal processing in early visual and auditory cortices depends on expected statistical relationship of multisensory information. J Neurosci 26:12260–12265.

Barraclough NE, Xiao D, Baker CI, Oram MW, Perrett DI (2005) Integration of visual and auditory information by superior temporal sulcus neurons responsive to the sight of actions. J Cogn Neurosci 17:377–391.

Beauchamp MS (2005) Statistical criteria in FMRI studies of multisensory integration. Neuroinformatics 3:93–113.

Beauchamp MS, Argall BD, Bodurka J, Duyn JH, Martin A (2004) Unraveling multisensory integration: patchy organization within human STS multisensory cortex. Nat Neurosci 7:1190–1192.

Benedek G, Perény J, Kovács G, Fischer-Szátmári L, Katoh YY (1997) Visual, somatosensory, auditory and nociceptive modality properties in the feline suprageniculate nucleus. Neuroscience 78:179–189.

Ben-Shachar M, Dougherty RF, Deutsch GK, Wandell BA (2007) Contrast responsivity in MT+ correlates with phonological awareness and reading measures in children. Neuroimage 37:1396–1406.

Boenke LT, Deliano M, Ohl FW (2009) Stimulus duration influences perceived simultaneity in audiovisual temporal-order judgment. Exp Brain Res 198:233–244.

Bonath B, Noesselt T, Martinez A, Mishra J, Schwiecker K, Heinze HJ, Hillyard SA (2007) Neural basis of the ventriloquist illusion. Curr Biol 17:1697–1703.

Bordi F, LeDoux JE (1994) Response properties of single units in areas of rat auditory thalamus that project to the amygdala. II. Cells receiving convergent auditory and somatosensory inputs and cells antidromically activated by amygdala stimulation. Exp Brain Res 98:275–286.

Brosch M, Selezneva E, Scheich H (2005) Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkeys. J Neurosci 25:6797–6806.

Bruce C, Desimone R, Gross CG (1981) Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. J Neurophysiol 46:369–384.

Budinger E, Scheich H (2009) Anatomical connections suitable for the direct processing of neuronal information of different modalities via the rodent primary auditory cortex. Hear Res 258:16–27.

Budinger E, Heil P, Hess A, Scheich H (2006) Multisensory processing via early cortical stages: connections of the primary auditory cortical field with other sensory systems. Neuroscience 143:1065–1083.

Budinger E, Laszcz A, Lison H, Scheich H, Ohl FW (2008) Non-sensory cortical and subcortical connections of the primary auditory cortex in Mongolian gerbils: bottom-up and top-down processing of neuronal information via field AI. Brain Res 1220:2–32.

Buracas GT, Fine I, Boynton GM (2005) The relationship between task performance and functional magnetic resonance imaging response. J Neurosci 25:3023–3031.

Burton H, Jones EG (1976) The posterior thalamic region and its cortical projection in New World and Old World monkeys. J Comp Neurol 168:249–301.

Busse L, Roberts KC, Crist RE, Weissman DH, Woldorff MG (2005) The spread of attention across modalities and space in a multisensory object. Proc Natl Acad Sci U S A 102:18751–18756.

Calvert GA (2001) Crossmodal processing in the human brain: insights from functional neuroimaging studies. Cereb Cortex 11:1110–1123.

Calvert GA, Thesen T (2004) Multisensory integration: methodological approaches and emerging principles in the human brain. J Physiol Paris 98:191–205.

Calvert GA, Brammer MJ, Bullmore ET, Campbell R, Iversen SD, David AS (1999) Response amplification in sensory-specific cortices during crossmodal binding. Neuroreport 10:2619–2623.

Calvert GA, Campbell R, Brammer MJ (2000) Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. Curr Biol 10:649–657.

Campana G, Cowey A, Walsh V (2006) Visual area V5/MT remembers "what" but not "where." Cereb Cortex 16:1766–1770.

Cappe C, Barone P (2005) Heteromodal connections supporting multisensory integration at low levels of cortical processing in the monkey. Eur J Neurosci 22:2886–2902.

Cappe C, Morel A, Barone P, Rouiller EM (2009) The thalamocortical projection systems in primate: an anatomical support for multisensory and sensorimotor interplay. Cereb Cortex 19:2025–2037.

Chandrasekaran C, Ghazanfar AA (2009) Different neural frequency bands integrate faces and voices differently in the superior temporal sulcus. J Neurophysiol 101:773–788.

D'Ardenne K, McClure SM, Nystrom LE, Cohen JD (2008) BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. Science 319:1264–1267.

Doyle MC, Snowden RJ (2001) Identification of visual stimuli is improved by accompanying auditory stimuli: the role of eye movements and sound location. Perception 30:795–810.

Driver J, Noesselt T (2008) Multisensory interplay reveals crossmodal influences on "sensory-specific" brain regions, neural responses, and judgments. Neuron 57:11–23.

Eckert MA, Kamdar NV, Chang CE, Beckmann CF, Greicius MD, Menon V (2008) A cross-modal system linking primary auditory and visual cortices: evidence from intrinsic fMRI connectivity analysis. Hum Brain Mapp 29:848–857.

Falchier A, Clavagnier S, Barone P, Kennedy H (2002) Anatomical evidence of multimodal integration in primate striate cortex. J Neurosci 22:5749–5759.

Frassinetti F, Bolognini N, Làdavas E (2002) Enhancement of visual perception by crossmodal visuo-auditory interaction. Exp Brain Res 147:332–343.

Friston KJ, Buechel C, Fink GR, Morris J, Rolls E, Dolan RJ (1997) Psychophysiological and modulatory interactions in neuroimaging. Neuroimage 6:218–229.

Friston KJ, Holmes AP, Price CJ, Büchel C, Worsley KJ (1999) Multisubject fMRI studies and conjunction analyses. Neuroimage 10:385–396.

Fu KM, Shah AS, O'Connell MN, McGinnis T, Eckholdt H, Lakatos P, Smiley J, Schroeder CE (2004) Timing and laminar profile of eye-position effects on auditory responses in primate auditory cortex. J Neurophysiol 92:3522–3531.

Fuhrmann Alpert G, Hein G, Tsai N, Naumer MJ, Knight RT (2008) Temporal characteristics of audiovisual information processing. J Neurosci 28:5344–5349.

Ghazanfar AA, Schroeder CE (2006) Is neocortex essentially multisensory? Trends Cogn Sci 10:278–285.

Ghazanfar AA, Maier JX, Hoffman KL, Logothetis NK (2005) Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. J Neurosci 25:5004–5012.

Ghazanfar AA, Chandrasekaran C, Logothetis NK (2008) Interactions between the superior temporal sulcus and auditory cortex mediate dynamic face/voice integration in rhesus monkeys. J Neurosci 28:4457–4469.

Giard MH, Peronnet F (1999) Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. J Cogn Neurosci 11:473–490.

Hackett TA, Smiley JF, Ulbert I, Karmos G, Lakatos P, de la Mothe LA, Schroeder CE (2007) Sources of somatosensory input to the caudal belt areas of auditory cortex. Perception 36:1419–1430.

Holmes NP (2009) The principle of inverse effectiveness in multisensory integration: some statistical considerations. Brain Topogr 21:168–176.

Hughes HC, Reuter-Lorenz PA, Nozawa G, Fendrich R (1994) Visual-auditory interactions in sensorimotor processing: saccades versus manual responses. J Exp Psychol Hum Percept Perform 20:131–153.

Jiang W, Jiang H, Rowland BA, Stein BE (2007) Multisensory orientation behavior is disrupted by neonatal cortical ablation. J Neurophysiol 97:557–562.

Jones EG (2001) The thalamic matrix and thalamocortical synchrony. Trends Neurosci 24:595–601.

Kanowski M, Rieger JW, Noesselt T, Tempelmann C, Hinrichs H (2007) Endoscopic eye tracking system for fMRI. J Neurosci Methods 160:10–15.

Kayser C, Logothetis NK (2007) Do early sensory cortices integrate cross-modal information? Brain Struct Funct 212:121–132.

Kayser C, Logothetis NK (2009) Directed interactions between auditory and superior temporal cortices and their role in sensory integration. Front Integr Neurosci 3:7.

Kayser C, Petkov CI, Augath M, Logothetis NK (2007) Functional imaging reveals visual modulation of specific fields in auditory cortex. J Neurosci 27:1824–1835.

Kayser C, Petkov CI, Logothetis NK (2008) Visual modulation of neurons in auditory cortex. Cereb Cortex 18:1560–1574.

Kayser C, Logothetis NK, Panzeri S (2010) Visual enhancement of the information representation in auditory cortex. Curr Biol 20:19–24.

Komura Y, Tamura R, Uwano T, Nishijo H, Ono T (2005) Auditory thalamus integrates visual inputs into behavioral gains. Nat Neurosci 8:1203–1209.

Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI (2009) Circular analysis in systems neuroscience: the dangers of double dipping. Nat Neurosci 12:535–540.

Lakatos P, Chen CM, O'Connell MN, Mills A, Schroeder CE (2007) Neuronal oscillations and multisensory interaction in primary auditory cortex. Neuron 53:279–292.

Lakatos P, Karmos G, Mehta AD, Ulbert I, Schroeder CE (2008) Entrainment of neuronal oscillations as a mechanism of attentional selection. Science 320:110–113.

Lakatos P, O'Connell MN, Barczak A, Mills A, Javitt DC, Schroeder CE (2009) The leading sense: supramodal control of neurophysiological context by attention. Neuron 64:419–430.

Lavie N (2005) Distracted and confused?: selective attention under load. Trends Cogn Sci 9:75–82.

Linke R, Braune G, Schwegler H (2000) Differential projection of the posterior paralaminar thalamic nuclei to the amygdaloid complex in the rat. Exp Brain Res 134:520–532.

Lippert M, Logothetis NK, Kayser C (2007) Improvement of visual contrast detection by a simultaneous sound. Brain Res 1173:102–109.

Macaluso E, Driver J (2005) Multisensory spatial interactions: a window onto functional integration in the human brain. Trends Neurosci 28:264–271.

Macaluso E, Frith CD, Driver J (2000) Modulation of human visual cortex by crossmodal spatial attention. Science 289:1206–1208.

Marks LE (1987) sOn cross-modal similarity: auditory-visual interactions in speeded discrimination. J Exp Psychol Hum Percept Perform 13:384–394.

Marks LE, Szczesiul R, Ohlott P (1986) On the cross-modal perception of intensity. J Exp Psychol Hum Percept Perform 12:517–534.

Marks LE, Ben-Artzi E, Lakatos S (2003) Cross-modal interactions in auditory and visual discrimination. Int J Psychophysiol 50:125–145.

McDonald JJ, Teder-Sälejärvi WA, Hillyard SA (2000) Involuntary orienting to sound improves visual perception. Nature 407:906–908.

McDonald JJ, Teder-Sälejärvi WA, Di Russo F, Hillyard SA (2003) Neural substrates of perceptual enhancement by cross-modal spatial attention. J Cogn Neurosci 15:10–19.

Meredith MA, Stein BE (1983) Interactions among converging sensory inputs in the superior colliculus. Science 221:389–391.

Meredith MA, Nemitz JW, Stein BE (1987) Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. J Neurosci 7:3215–3229.

Molholm S, Ritter W, Murray MM, Javitt DC, Schroeder CE, Foxe JJ (2002) Multisensory auditory-visual interactions during early sensory processing in humans: a high-density electrical mapping study. Brain Res Cogn Brain Res 14:115–128.

Molholm S, Ritter W, Javitt DC, Foxe JJ (2004) Multisensory visual-auditory object recognition in humans: a high-density electrical mapping study. Cereb Cortex 14:452–465.

Musacchia G, Sams M, Nicol T, Kraus N (2006) Seeing speech affects acoustic information processing in the human brainstem. Exp Brain Res 168:1–10.

Musacchia G, Sams M, Skoe E, Kraus N (2007) Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. Proc Natl Acad Sci U S A 104:15894–15898.

Nichols T, Poline JB (2009) Commentary on Vul et al.'s (2009) "Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition." Perspect Psychol Sci 4:291–293.

Noesselt T, Rieger JW, Schoenfeld MA, Kanowski M, Hinrichs H, Heinze HJ,

Driver J (2007) Audiovisual temporal correspondence modulates human multisensory superior temporal sulcus plus primary sensory cortices. J Neurosci 27:11431–11441.

Noesselt T, Bergmann D, Hake M, Heinze HJ, Fendrich R (2008) Sound increases the saliency of visual events. Brain Res 1220:157–163.

O'Connor DH, Fukui MM, Pinsk MA, Kastner S (2002) Attention modulates responses in the human lateral geniculate nucleus. Nat Neurosci 5:1203–1209.

Odgaard EC, Arieh Y, Marks LE (2003) Cross-modal enhancement of perceived brightness: sensory interaction versus response bias. Percept Psychophys 65:123–132.

Posner M (1978) Chronometric explorations of mind. Oxford: Oxford UP.

Reches A, Gutfreund Y (2009) Auditory and multisensory responses in the tectofugal pathway of the barn owl. J Neurosci 29:9602–9613.

Rockland KS, Ojima H (2003) Multisensory convergence in calcarine visual areas in macaque monkey. Int J Psychophysiol 50:19–26.

Romei V, Murray MM, Merabet LB, Thut G (2007) Occipital transcranial magnetic stimulation has opposing effects on visual and auditory stimulus detection: implications for multisensory interactions. J Neurosci 27:11465–11472.

Sadaghiani S, Maier JX, Noppeney U (2009) Natural, metaphoric, and linguistic auditory direction signals have distinct influences on visual motion processing. J Neurosci 29:6490–6499.

Schroeder CE, Lakatos P (2009) Low-frequency neuronal oscillations as instruments of sensory selection. Trends Neurosci 32:9–18.

Schroeder CE, Lakatos P, Kajikawa Y, Partan S, Puce A (2008) Neuronal oscillations and visual amplification of speech. Trends Cogn Sci 12:106–113.

Schulz M, Ross B, Pantev C (2003) Evidence for training-induced cross-modal reorganization of cortical functions in trumpet players. Neuroreport 14:157–161.

Stanislaw H, Todorov N (1999) Calculation of signal detection theory measures. Behav Res Methods Instrum Comput 31:137–149.

Stein B, Meredith M (1993) The merging of the senses. Cambridge, MA: MIT.

Stein BE, Arigbede MO (1972) Unimodal and multimodal response properties of neurons in the cat's superior colliculus. Exp Neurol 36:179–196.

Stein BE, Huneycutt WS, Meredith MA (1988) Neurons and behavior: the same rules of multisensory integration apply. Brain Res 448:355–358.

Stevenson RA, James TW (2009) Audiovisual integration in human superior temporal sulcus: inverse effectiveness and the neural processing of speech and object recognition. Neuroimage 44:1210–1223.

Störmer VS, McDonald JJ, Hillyard SA (2009) Cross-modal cueing of attention alters appearance and early cortical processing of visual stimuli. Proc Natl Acad Sci U S A 106:22456–22461.

Sylvester R, Josephs O, Driver J, Rees G (2007) Visual FMRI responses in human superior colliculus show a temporal-nasal asymmetry that is absent in lateral geniculate and visual cortex. J Neurophysiol 97:1495–1502.

Tootell RB, Reppas JB, Kwong KK, Malach R, Born RT, Brady TJ, Rosen BR, Belliveau JW (1995) Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. J Neurosci 15:3215–3230.

Vroomen J, de Gelder B (2000) Sound enhances visual perception: cross-modal effects of auditory organization on vision. J Exp Psychol Hum Percept Perform 26:1583–1590.

Vul E, Harris C, Winkielman P, Pashler H (2009) Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. Perspect Psychol Sci 4:274–290.

Wang Y, Celebrini S, Trotter Y, Barone P (2008) Visuo-auditory interactions in the primary visual cortex of the behaving monkey: electrophysiological evidence. BMC Neurosci 129:79.

Watkins S, Shams L, Josephs O, Rees G (2007) Activity in human V1 follows multisensory perception. Neuroimage 37:572–578.

Wepsic JG (1966) Multimodal sensory activation of cells in the magnocellular medial geniculate nucleus. Exp Neurol 15:299–318.

Yeterian EH, Pandya DN (1989) Thalamic connections of the cortex of the superior temporal sulcus in the rhesus monkey. J Comp Neurol 282:80–97.

Yukie M, Iwai E (1981) Direct projection from the dorsal lateral geniculate nucleus to the prestriate cortex in macaque monkeys. J Comp Neurol 201:81–97.