ORIGINAL ARTICLE

Look Hear! The Prefrontal Cortex is Stratified by Modality of Sensory Input During Multisensory Cognitive Control

Andrew R. Mayer[1,2,3], Sephira G. Ryman[1,3], Faith M. Hanlon[1], Andrew B. Dodd[1] and Josef M. Ling[1]

[1]The Mind Research Network/Lovelace Biomedical and Environmental Research Institute, Albuquerque, NM 87106, USA, [2]Departments of Neurology and Psychiatry, University of New Mexico School of Medicine, Albuquerque, NM 87131, USA and [3]Department of Psychology, University of New Mexico, Albuquerque, NM 87131, USA

Address correspondence to Andrew Mayer, The Mind Research Network, Pete & Nancy Domenici Hall, 1101 Yale Blvd. NE, Albuquerque, NM 87106, USA. Email: amayer@mrn.org

Abstract

Parsing multisensory information from a complex external environment is a fundamental skill for all organisms. However, different organizational schemes currently exist for how multisensory information is processed in human (supramodal; organized by cognitive demands) versus primate (organized by modality/cognitive demands) lateral prefrontal cortex (LPFC). Functional magnetic resonance imaging results from a large cohort of healthy controls ($N = 64$; Experiment 1) revealed a rostral-caudal stratification of LPFC for auditory versus visual attention during an audio-visual Stroop task. The stratification existed in spite of behavioral and functional evidence of increased interference from visual distractors. Increased functional connectivity was also observed between rostral LPFC and auditory cortex across independent samples (Experiments 2 and 3) and multiple methodologies. In contrast, the caudal LPFC was preferentially activated during visual attention but functioned in a supramodal capacity for resolving multisensory conflict. The caudal LPFC also did not exhibit increased connectivity with visual cortices. Collectively, these findings closely mirror previous nonhuman primate studies suggesting that visual attention relies on flexible use of a supramodal cognitive control network in caudal LPFC whereas rostral LPFC is specialized for directing attention to auditory inputs (i.e., human auditory fields).

Key words: ARM, audio-visual, fMRI, lateral prefrontal cortex, multisensory

Introduction

All organisms are constantly bombarded by multisensory information that must be efficiently processed to gain information about the external environment, minimize uncertainty, and guide goal-directed behavior (Ernst and Banks 2002; Weissman et al. 2004; Fetsch et al. 2013). Both unisensory and multisensory processing occur in several stages and can be automatic or heavily influenced by top-down attentional control (Desimone and Duncan 1995; Driver and Noesselt 2008; Talsma et al. 2010; Fetsch et al. 2013). For example, over a half-century ago it was observed that unisensory cortex is positively modulated (i.e., increased neuronal response) for attended stimuli (Hubel et al. 1959), with more recent evidence of suppression (i.e., decreased neuronal response) for ignored stimuli (Johnson and Zatorre 2005; Mayer et al. 2009). Both phenomena are referred to as attention-related modulations (ARM), and likely result from feedback from higher-order cortical areas (Motter 1993; Buffalo et al. 2010). However, the cortical regions responsible for top-down modulation of unisensory cortex and the resolution of conflicting multisensory information remain actively debated, especially within prefrontal cortex (van Atteveldt et al. 2014; Talsma 2015).

Visual information tends to dominate during most situations involving multisensory stimulation (Visual Dominance Theory or

Colavita effect; Colavita 1974; Posner et al. 1976; Donohue et al. 2013), with neuronal evidence of reduced competition from auditory distractors (Schmid et al. 2011) and increased cortical involvement for top-down attention shifts away from visual streams (Spence et al. 2012). However, the resolution of conflicting multisensory information streams ultimately depends on task context (Modality Appropriateness Hypothesis; Ernst and Bulthoff 2004; Yuval-Greenberg and Deouell 2007; Mayer et al. 2009; Schmid et al. 2011), with auditory dominance occurring during tasks with high temporal demands. In addition, the perceived reliability of information (Ernst and Banks 2002; Baier et al. 2006; Fetsch et al. 2013) and stage of processing (Chen and Zhou 2013; Fetsch et al. 2013) also play a role in determining which sensory modality dominates.

The dorsal medial (dMFC; composed of the anterior cingulate gyrus and presupplementary motor area) and dorsolateral (DLPFC) prefrontal cortex are both critical for resolving conflicting multisensory stimuli and top-down attentional control, albeit with different postulated roles. The dMFC signals the need for top-down control via conflict detection or an increase in error-likelihood (Botvinick et al. 2001; Brown and Braver 2005; Shenhav et al. 2013). In contrast, the DLPFC instantiates top-down control through representation of task context, goal maintenance, and by more directly biasing perceptual processing within sensory and other cortical regions (MacDonald et al. 2000; Badre and Wagner 2004; Egner and Hirsch 2005; Braver 2012).

Other theories suggest that the LPFC works in conjunction with the intraparietal sulcus (IPS) during top-down attentional selection (Corbetta and Shulman 2002). Caudal aspects of the LPFC (frontal eye fields; FEF) and IPS have been implicated in modulating visual cortex during transcranial magnetic stimulation (Ruff et al. 2008, 2006), direct stimulation (Moore and Armstrong 2003), and neuroimaging studies of visual spatial attention (Bressler et al. 2008; Lauritzen et al. 2009). However, the majority of studies implicating the FEF and IPS have focused exclusively on visual tasks/visual cortex, and it is not clear if these results would generalize to other sensory modalities (Braga et al. 2013). For example, nonhuman primate data indicate an exclusive representation of auditory cortex in rostral LPFC (i.e., auditory fields) versus a greater representation of visual cortex in caudal LPFC (Barbas and Mesulam 1985; Medalla and Barbas 2014). In contrast, human LPFC has been stratified into rostral-caudal and ventral-dorsal gradients based on various cognitive (e.g., abstract vs. concrete; spatial vs. nonspatial) schemes (Plakke and Romanski 2014; Bahlmann et al. 2015), with limited evidence for sensory specialization. Instead, neuroimaging studies implicate distinct networks for attention to visual (dorsal fronto-parietal cortex) relative to auditory (fronto-temporal cortex) stimuli across major brain lobules, with the DLPFC serving in a supramodal capacity for multiple sensory modalities (Braga et al. 2013).

A multisensory Stroop task was therefore used to investigate whether the prefrontal cortex functions in a supramodal capacity (i.e., conflict resolution regardless of modality for focused attention) or exhibits modality-specific patterns of activation dependent on sensory input. In Experiment 1, 64 healthy controls identified simultaneously presented congruent (i.e., identical auditory and visual numbers) or incongruent (i.e., conflicting auditory and visual numbers) multisensory targets presented at either low (0.33 Hz) or high (0.66 Hz) frequencies. Two frequencies were used to manipulate both cognitive and perceptual load, which we have previously shown to directly influence the magnitude of ARM in unisensory cortex (Mayer et al. 2009). We hypothesized greater involvement of the dMFC and DLPFC during

incongruent trials regardless of whether attention was directed toward auditory or visual inputs (i.e., supramodal role). We also predicted that these supramodal frontal regions from the cognitive control network (CCN) would show increased activation while attending to visual cues and ignoring auditory distractors, along with the FEF (Ruff et al. 2006; Ruff et al. 2008; Mayer et al. 2009). Similar to previous studies (Egner and Hirsch 2005; Lauritzen et al. 2009), Experiments 2 (N = 64) and 3 (N = 42) utilized functional-connectivity magnetic resonance imaging (fcMRI) analyses to provide confirmatory evidence from task results.

## Materials and Methods

Participants

Sixty-seven healthy right-handed adult volunteers between the ages of 18–45 participated in Experiments 1 and 2. One participant was identified as a motion outlier (more than 3 times the interquartile range on 2 of 6 framewise displacement parameters) and removed from subsequent analyses. One participant was excluded for slow reaction times (more than 3 times the interquartile range) and another participant performed below chance levels on accuracy. The final cohort therefore included 64 participants (35 males; mean age = 29.33 ± 8.23 years; mean education = 13.69 ± 2.05 years). Resting-state data were also collected on an independent cohort of 42 right-handed healthy adults (Experiment 3: 30 males; mean age = 28.76 ± 5.94 years; mean education = 15.56 ± 1.95 years) to verify connectivity results obtained in Experiment 2.

Exclusion criteria for the study included self-reported history of neurological disorders, psychiatric disorders, developmental disorders, traumatic brain injury with >5 min loss of consciousness, recent history of substance abuse or dependence and current illicit substance use. None of the participants were taking psychoactive medications at the time of the study, and all participants were encouraged not to drink alcohol the night before the study. Informed consent was obtained from all subjects according to institutional guidelines at the University of New Mexico School of Medicine.

Experiment Task

The multisensory task (Fig. 1A,B; Experiment 1) began with a multisensory (audio-visual) cue (exemplary visual angle = 7.69°; duration = 175 ms) indicating the sensory modality for focused attention ("HEAR" = attend-auditory; "LOOK" = attend-visual). Cues were followed by congruent or incongruent multisensory numeric stimuli (target words = one, two or three; exemplary visual angle = 9.73°; 200 ms duration) at either low (0.33 Hz; 3 trials per block) or higher (0.66 Hz; 6 trials per block) rates of stimulus frequency over a 10-s block. Visual stimuli were presented as word rather than numerical ("three" versus "3") representations to maximize the semantic relationships to aurally presented stimuli (Fias et al. 2001).

Participants responded to target numbers in the attended modality with a right-handed button press. All multisensory stimuli (cues and targets) were presented foveally (eye-centered) and binaurally via headphones (head-centered). Participants were asked to maintain constant head and eye positioning (visual fixation on a centrally presented cross) to consistently maintain spatial correspondence between auditory and visual stimuli. Simultaneous presentation of audio-visual stimuli was achieved with diode and microphone measurements made prior to study start (Donohue et al. 2013). A total of 432 trials were presented across 6 separate imaging runs during the multisensory Stroop task (total scan time = 32 min and 24 s).
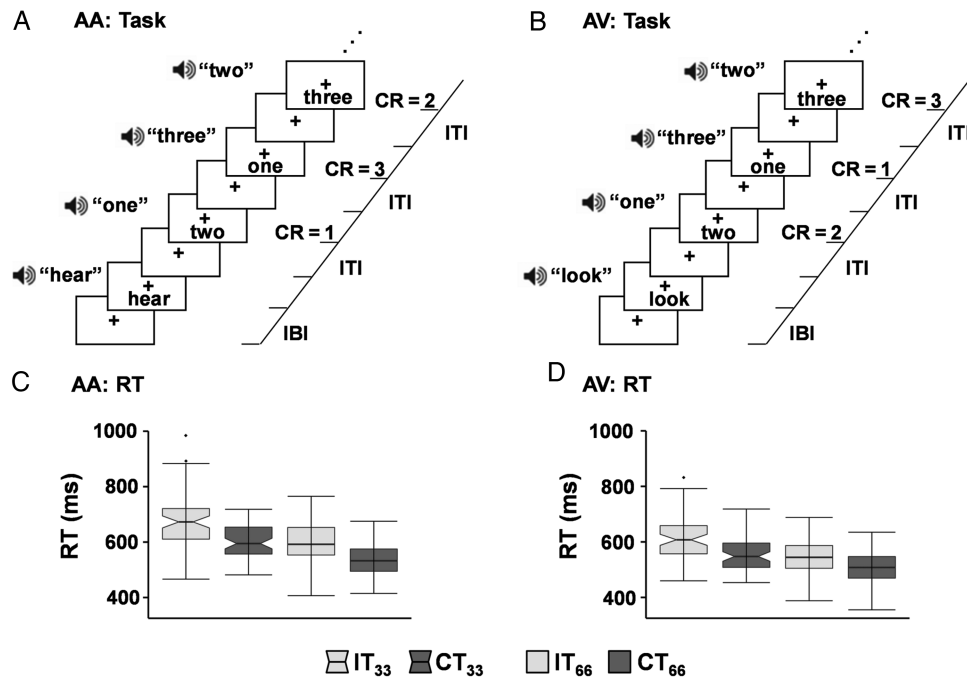
Figure 1. Diagrammatic representation of the task and behavioral results. Participants attended to target stimuli (numbers: one, two or three) in either the auditory (attend-auditory: AA) (A) or visual (attend-visual: AV) (B) modality while ignoring incongruent (IT) or congruent (CT) distractor stimuli (numbers: one, two or three) in the opposite modality. Correct responses (CR) are indicated on right side of panels. Multimodal cue words indicate the modality for focused attention. The interblock interval (IBI) varied randomly, with the intertrial interval (ITI) determined by the rate of stimulus presentation (0.33 Hz or 0.66 Hz). Box-and-whisker plots depict reaction times (RT) for both AA (C) and AV (D) conditions for low ($_{33}$) and high ($_{66}$) frequency trials.

The interblock interval varied between 8, 10, and 12 s to decrease temporal expectations and permit modeling of the baseline response (visual fixation plus baseline gradient noise). Response time data were analyzed with a $2 \times 2 \times 2$ (Modality [AA vs. AV] × Congruency [Congruent vs. Incongruent] × Frequency [0.33 vs. 0.66 Hz]) repeated-measures analysis of variance (ANOVA). For the resting-state scan (Experiments 2 and 3), participants were instructed to stare at a centrally presented white fixation cross on a black background for approximately 5 min.

MR Imaging and Statistical Analyses

High resolution T_1 (1.0 mm^3) and echo-planar ($3.55 \times 3.55 \times 4.75$ mm) images were collected on a 3 T scanner. The first 3 images of each run were eliminated to account for T_1 equilibrium effects, resulting in 966 images for the multisensory task and 149 images for resting-state data. Both task- and resting-state data were despiked, motion corrected, slice-time corrected, normalized and blurred with a 6 mm kernel (Supplementary Material). Deconvolution generated a single hemodynamic response function for each trial-type relative to the baseline state (visual fixation plus gradient noise) based on the first 22-s poststimulus onset including separate regressors for error trials (Mayer et al. 2011). A $2 \times 2 \times 2$ (Modality [AA vs. AV] × Congruency [Congruent vs. Incongruent] × Frequency [0.33 vs. 0.66 Hz]) repeated-measures ANOVA was used to examine a priori hypotheses. All voxelwise results (task and resting-state) were individually corrected for false positives at $P < 0.05$ using both parametric and cluster-based thresholds ($P < 0.005$ and minimum cluster size = 1088 μL) based on 10 000 Monte Carlo simulations (AFNI's 3dClustSim program).

fcMRI maps were calculated by first regressing motion parameters, their first-order derivatives, and estimates of physiological noise (derived from white matter and cerebral spinal fluid masks), followed by band-pass filtering (0.01–0.10 Hz). Seeds were empirically derived from group comparisons of task data, and used as an independent assessment of modality-specific connectivity across both cohorts of participants (Experiments 2 and 3).

Results

Behavioral Data from Experiment 1

Reaction time data indicated significant main effects of Congruency ($F_{1,63} = 161.67$, $P < 0.001$), Modality ($F_{1,63} = 118.14$, $P < 0.001$) and Frequency ($F_{1,63} = 237.03$, $P < 0.001$), as well as significant 2-way interactions for all effects (all P's < 0.05). Follow-up analyses indicated that the Congruency × Modality interaction resulted from a greater difference ($t_{63} = 2.34$, $P = 0.02$) between incongruent and congruent trials during attend-auditory (66.69 ± 52.32 ms) relative to attend-visual (49.02 ± 41.75 ms) trials. A significantly greater difference ($t_{63} = 2.51$, $P = 0.01$) between the incongruent and congruent trials was also present for low (65.51 ± 45.35 ms) relative to high (50.20 ± 41.78 ms) frequency trials (Congruency × Frequency interaction). Finally, the Frequency × Modality interaction was driven by a greater difference ($t_{63} = -3.94$, $P < 0.001$) between auditory and visual stimuli in the low (56.21 ± 45.46 ms) relative to high (34.95 ± 33.42 ms) frequency trials. Accuracy data approached ceiling, and are presented in Supplementary Material.

Experiment 1 Functional Task Results

A voxelwise, $2 \times 2 \times 2$ (Modality × Congruency × Frequency) repeated-measures ANOVA was performed on multisensory task data. Incongruent trials (Fig. 2) resulted in increased activation
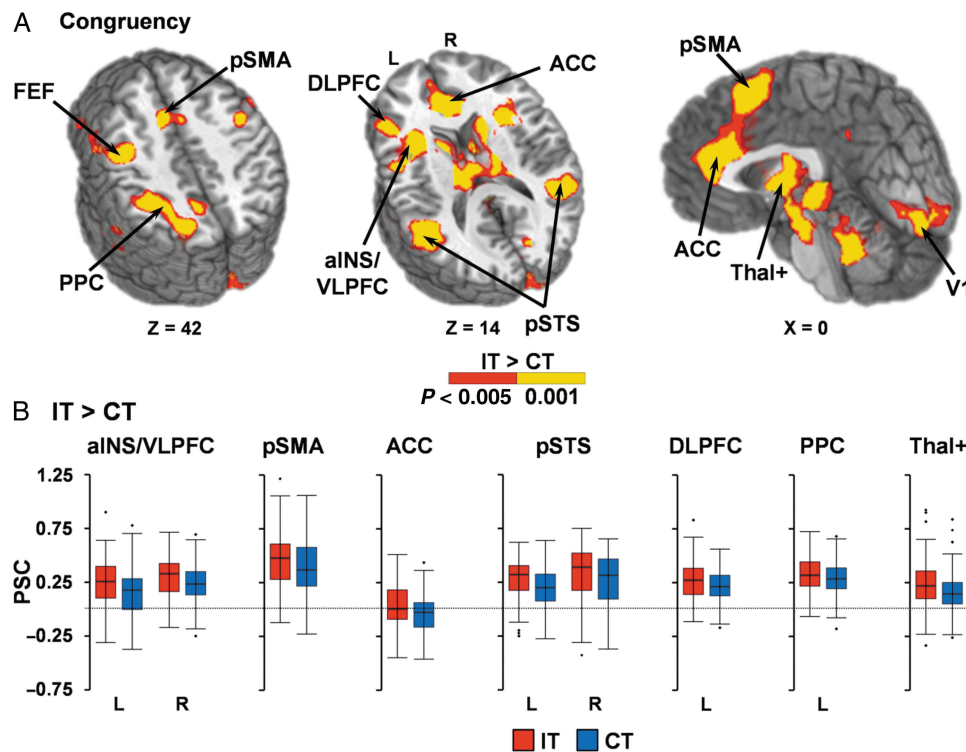
Figure 2. (A) presents significant regions (minimum $P < 0.005$ and minimum cluster size = 1088 µL) exhibiting greater activation ($P < 0.005$: red; $P < 0.001$: yellow) for incongruent (IT) relative to congruent (CT) trials. Selected regions included the bilateral anterior insular/ventrolateral prefrontal cortex (aINS/VLPFC), dorsolateral prefrontal cortex (DLPFC), frontal eye fields (FEF), presupplementary motor area (pre-SMA), anterior cingulate gyrus (ACC), posterior superior temporal sulcus (pSTS)/ middle temporal gyrus, left posterior parietal cortex (PPC), primary visual cortex (V1), thalami and subthalamic nuclei (Thal+). Locations of the sagittal (X) and axial (Z) slices are given according to the Talairach atlas for the left (L) and right (R) hemispheres. Percent signal change (PSC) values are presented in (B) for IT (red) and CT (blue) trials.

within the right (BAs 13/44/45/47; µL = 6286) and left (BAs 13/22/ 44/45/47; µL = 10 795) anterior insula/ventrolateral prefrontal cortex (VLPFC); left DLPFC extending into FEF/premotor cortex (BAs 6/9/46; µL = 21 400); right FEF/premotor extending into motor cortex (BAs 4/6/9; µL = 4294); bilateral presupplementary motor area (pre-SMA) extending into dorsal anterior cingulate gyrus (dACC; BAs 6/8/24/32; µL = 7879); right (BAs 13/21/22/39/40; µL = 11 001) and left (BAs 13/21/22/37/39/40; µL = 15 476) posterior superior temporal sulcus (pSTS)/middle temporal gyrus; left posterior parietal cortex (BAs 7/19/40; µL = 8657); primary visual cortex (BAs 17/ 18/23/30; µL = 7693); bilateral striatum/thalami/subthalamic nuclei (µL = 33 089) and bilateral (µL = 9890) and left (µL = 2505) cerebellum relative to congruent trials. In contrast, the bilateral anterior cingulate gyrus (ACC; BAs 9/24/32; µL = 12 503) and the left precuneus (BAs 7/31; µL = 1760) exhibited greater deactivation during congruent relative to incongruent trials.

Attend-visual trials (Fig. 3A,B) resulted in increased activation within right (BAs 17/18/19/20/36/37; µL = 16 382) and left (BAs 17/18/19/37; µL = 23 385) V1/ventral visual streams, as well as within right (BAs 7/19/39; µL = 9303) and left (BAs 7/18/19/31; µL = 3056) dorsal visual streams. In addition, the right (BAs 9/45/ 46; µL = 4130) and left caudal DLPFC (BAs 9/45/46; µL = 3271); right (BAs 6/9; µL = 3099) and left (BAs 6/9; µL = 4529) inferior FEF (iFEF) extending into premotor cortex; left mid-insula (BA 13; µL = 1183); bilateral thalami (including the lateral geniculate)/subthalamic nuclei (µL = 4399); and bilateral cerebellum (µL = 4790) also exhibited increased activation during attend-visual trials. Greater deactivation during attend-visual compared with attend-auditory trials was also observed in the right (BAs 18/19/30; µL = 5450) and left (BAs 18/19/30; µL = 2318) lingual

gyrus extending into the bilateral cuneus (BAs 7/18/19/31; µL = 15 330).

In contrast, activation was greater for attend-auditory trials (Fig. 3A,C) in the left primary and secondary auditory cortex (BAs 13/40/41/42; µL = 1102) relative to attend-visual trials. Additional areas of increased activation during attend-auditory trials included the right (BAs 8/9; µL = 3005) and left (BAs 8/9/10; µL = 5754) superior frontal gyrus (SFG) extending into rostral DLPFC; dACC (BAs 6/9/32; µL = 2293); right premotor cortex (BA 6; µL = 2700); and bilateral precuneus (BA 7; µL = 5110). Finally, attend-auditory trials also resulted in greater deactivation for several regions of the default mode network (DMN), including the bilateral medial frontal (BAs 6/8/9; µL = 6937) and posterior cingulate (BAs 7/23/ 29/30/31; µL = 3057) gyri, as well as right (BAs 28/35/36; µL = 1959) and left (BAs 27/28/35/36; µL = 2022) hippocampi.

The Congruency × Modality interaction was significant within the medial aspects of the right sensorimotor cortex extending into parietal lobule (BAs 3/4/5/7/40; µL = 1930). Simple effects indicated that this was due to greater deactivation in attend-auditory relative to attend-visual for incongruent trials ($P < 0.05$) with no differences between modalities on congruent trials ($P > 0.10$). The Frequency × Modality interaction was significant in the right (BAs 17/18/19/37; µL = 5033) and left (BAs 17/18/19; µL = 3980) V1 and extrastriate cortex. Simple effects indicated that there were increased ARM between high and low frequency trials during the attend-visual relative to attend-auditory conditions ($P < 0.05$).

The Modality × Congruency × Frequency interaction was not significant. Results for the main effect of Frequency and the Congruency × Frequency interaction are presented in Supplementary Material.
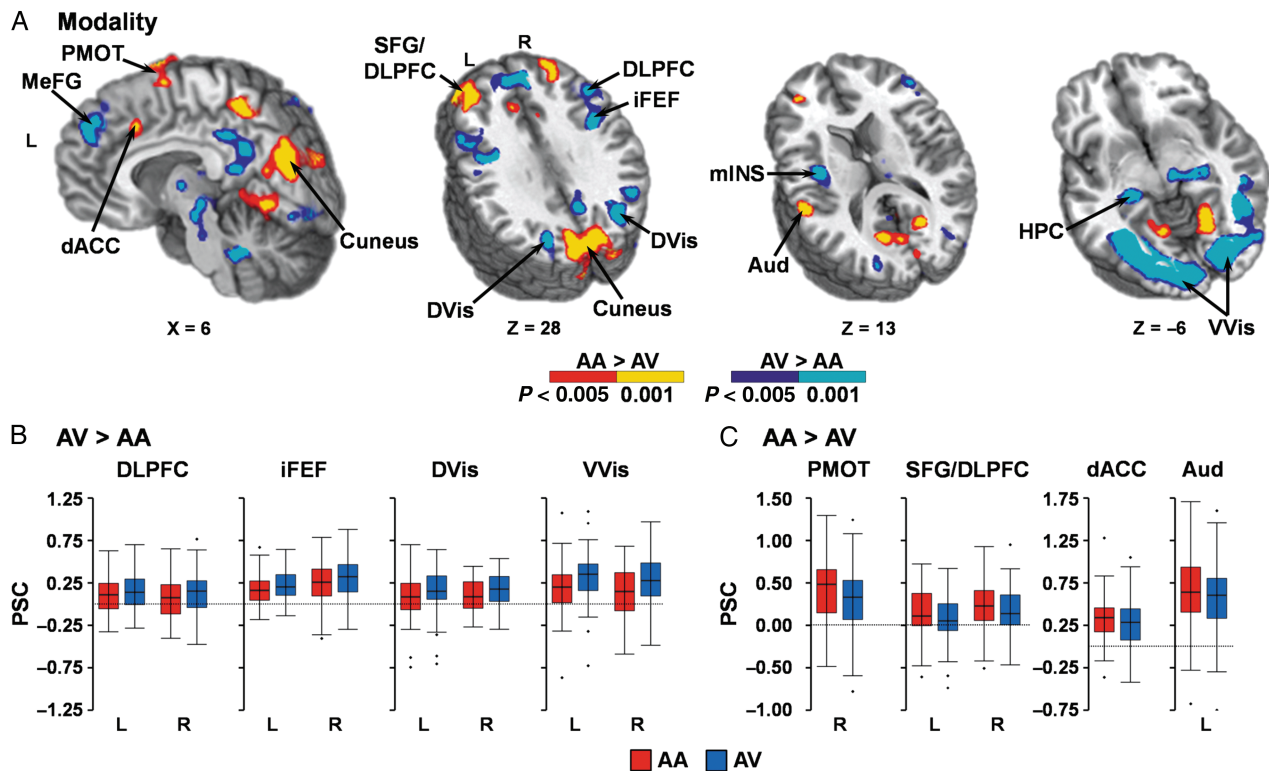
Figure 3. (A) presents significant regions (minimum P < 0.005 and minimum cluster size = 1088 µL) exhibiting greater activation for either attend-auditory (AA) relative to attend-visual (AV) trials (P < 0.005: red; P < 0.001: yellow) or for AV relative to AA trials (P < 0.005: blue; P < 0.001: cyan). Regions exhibiting increased activation during AV trials (B) included the medial frontal gyrus (MeFG), dorsolateral prefrontal cortex (DLPFC), inferior frontal eye fields (iFEF), mid-insula (mINS), hippocampi (HPC), primary visual cortex extending into ventral visual stream (VVis), and dorsal visual stream (DVis). Regions exhibiting increased activation during AA trials (C) included the superior frontal gyrus (SFG) extending into the DLPFC, dorsal anterior cingulate cortex (dACC), left primary and secondary auditory cortex (Aud), right premotor cortex (PMOT) and cuneus. Locations of the sagittal (X) and axial (Z) slices are given according to the Talairach atlas for the left (L) and right (R) hemispheres. Percent signal change (PSC) values for selected regions are presented in B,C (AA = red; AV = blue).

Reproducibility of Modality Effects

Jackknifing was used to examine the reproducibility of the significant modality effects observed during Experiment 1 within frontal cortical regions for either AV or AA trials at smaller sample sizes more typically used in neuroimaging experiments. Participants were randomly selected without replacement into subsamples ranging from 20 to 55 participants (in intervals of 5). Paired t-tests examined whether the main effect of modality was significant (P < 0.005) within each of the frontal regions and in the same direction (AV > AA or AA > AV) as observed with the full sample. This procedure was repeated 1000 times at each sample size, with Figure 4 depicting the percentage of times each region surpassed the designated P-value. Results indicate that larger sample sizes (N = 40–50) were necessary for robustly (>90% of iterations) identifying modality-specific patterns of activity within a majority (7/9) of prefrontal cortex regions.

Finally, the relationship between response time data and prefrontal cortical regions demonstrating modality-specific specialization was evaluated with a series of multiple regressions. The first regression examined whether differences in prefrontal activation (attend-visual minus attend-auditory trials) would be associated with response time differences (attend-visual minus attend-auditory trials) on a global level (i.e., across all prefrontal regions listed in Table 1). Results from this global model were not significant. Two separate multiple regressions were therefore conducted next to determine whether a more specialized relationship existed for prefrontal cortex within each of the sensory

modalities (e.g., regressing auditory prefrontal regions only on attend-auditory response time data). However, neither the attend-auditory or attend-visual model was significant following Bonferroni correction for multiple comparisons.
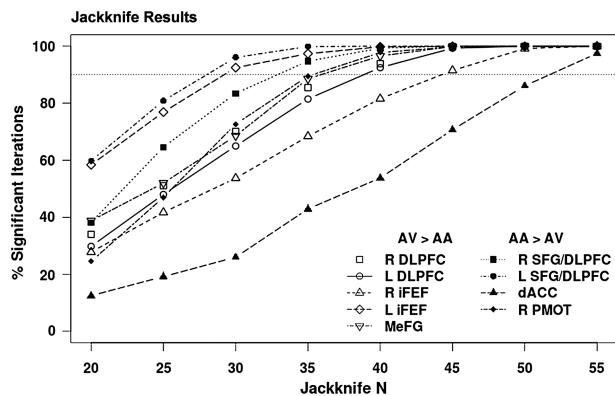
Experiments 2 and 3: Connectivity Results

Experiment 2 (N = 64) examined whether prefrontal regions that exhibited modality-specific effects (AA > AV or AV > AA) during the multisensory task (Table 1) would show a similar directional bias in terms of resting-state functional connectivity with seeds from either the auditory or visual cortex. Specifically, the unisensory cortical regions that exhibited ARM during the main effect of Modality (left auditory cortex, right and left primary/ventral visual cortex) were used as empirically derived seeds. Paired-samples t-tests then evaluated whether connectivity was greater for left auditory cortex relative to right or left visual seeds. Regions demonstrating both significant fcMRI and significant Modality effects (i.e., from Experiment 1) were determined by conducting an additional small-volume correction. Specifically, an additional overlap criterion (8 native voxels; 512 µL) was imposed on the individually corrected results from both the main effect of Modality and the connectivity results (corrected fcMRI ∩ corrected main effect of Modality). The thresholding criteria were based on 10 000 Monte Carlo simulations. To ensure reproducibility of findings, connectivity analyses were replicated on an independent cohort (Experiment 3; N = 42) of healthy controls who had completed different functional tasks (i.e., nonmultisensory) prior to the collection of resting-state data (see Supplementary Material).

Results (Fig. 5; Table 1) indicated a clear intersection of task Modality effects (AA > AV) and fcMRI results (auditory seed > visual seed) for prefrontal regions that were preferentially activated during the attend-auditory condition. Specifically, increased connectivity was observed between the auditory seed and the dACC, left SFG/rostral DLPFC, and right premotor cortex across both Experiments 2 and 3 (Supplementary Fig. 1) as well as seed placements (right versus left visual seed; Supplementary Fig. 2). The right SFG/rostral DLPFC was the only auditory prefrontal area that failed to

Figure 4. Jackknife analyses examined the reproducibility of modality effects within frontal cortical regions for attend-visual (AV) and attend-auditory (AA) trials during Experiment 1. Significant frontal regions exhibiting a main effect of modality included the dorsolateral prefrontal cortex (DLPFC), inferior frontal eye fields (iFEF), medial frontal gyrus (MeFG), superior frontal gyrus/DLPFC (SFG/DLPFC), dorsal anterior cingulate gyrus (dACC), and premotor cortex (PMOT) within the right (R) and left (L) hemispheres. Jackknifed sample size is presented on the X axis (Jackknife N), whereas the percentage of iterations (total = 1000) where significant effects were reproduced (% significant iterations) is displayed on the Y axis. Results indicate that larger sample sizes were necessary for robustly (90% denoted with dotted line) identifying the differential patterns of modality-specific activity in a majority of regions.

replicate across the combinations of cohorts and methodologies. Interestingly, the region of the precuneus/cuneus that exhibited greater deactivation during attend-visual trials also showed increased fcMRI for auditory relative to visual seeds.

In contrast, regions within the prefrontal cortex that showed increased activity during the attend-visual condition either showed no (right DLPFC and iFEF) or inconsistent (left DLPFC and iFEF) preferential connectivity to auditory and visual cortex across different cohorts and methods. Rather, results indicated increased fcMRI for visual relative to auditory seeds within the most lateral aspects of the medial, superior, and middle frontal gyri (Supplementary Fig. 1), regions which did not spatially overlap with the prefrontal regions exhibiting attentional modulations.

The pattern of increased connectivity with auditory cortex in prefrontal regions that showed increased activation while attending to auditory stimuli was generally reproducible (Supplementary Results) regardless of visual seed volume (Supplementary Fig. 3) or whether seeds were selected using individually defined anatomical labels from FreeSurfer (Supplementary Figs 4 and 5).

Discussion

The current study investigated whether the LPFC is specialized for directing attention to a particular sensory modality or whether prefrontal cortex operates in a supramodal capacity-dependent on task demands (e.g., here conflict resolution). Task (Experiment 1) and connectivity (Experiments 2 and 3) results indicated specialization for the rostral LPFC (SFG and rostral DLPFC) for auditory input, corroborating nonhuman primate research which demonstrate the existence of auditory fields in this region (Barbas and Mesulam 1985; Medalla and Barbas 2014). In contrast, both medial (ACC and pre-SMA) and caudal LPFC (DLPFC and FEF) were activated in a supramodal fashion based on task demands, suggesting more flexible roles during top-down multisensory cognitive control. Finally, behavioral and functional results were generally supportive of the Visual Dominance Theory/Colavita

Table 1. Intersection of resting-state connectivity results in the study (Experiment 2) and replication (Experiment 3) cohorts with prefrontal regions (Experiment 1) exhibiting a main effect of modality (excluding DMN)

Modality effect	Center of mass			Left AS versus right VS		Left AS versus left VS	
	X	Y	Z	Study	Rpl	Study	Rpl
AV > AA							
R. DLPFC							
BAs 9/45/46 (4130 μL)	49	27	21	n.s.	n.s.	n.s.	n.s.
L. DLPFC							
BAs 9/45/46 (3271 μL)	−46	20	24	AS > VS (1012 μL)	n.s.	n.s.	n.s.
R. FEF							
BAs 6/9 (3099 μL)	43	4	35	n.s.	n.s.	n.s.	n.s.
L. FEF							
BAs 6/9 (4529 μL)	−40	2	37	AS > VS (1951 μL)	AS > VS (533 μL)	AS > VS (1317 μL)	n.s.
AA > AV							
R. SFG/DLPFC							
BAs 8/9 (3005 μL)	31	43	35	AS > VS (883 μL)	n.s.	AS > VS (823 μL)	VS > AS (578 μL)
L. SFG/DLPFC							
BAs 8/9/10 (5754 μL)	−32	44	30	AS > VS (1411 μL)	AS > VS (1103 μL)	AS > VS (1294 μL)	AS > VS (899 μL)
dACC							
BAs 6/9/32 (2293 μL)	−2	17	37	AS > VS (2266 μL)	AS > VS (1839 μL)	AS > VS (2126 μL)	AS > VS (1302 μL)
R. Premotor Cortex							
BA 6 (2700 μL)	17	−3	64	AS > VS (1600 μL)	AS > VS (1578 μL)	AS > VS (1489 μL)	AS > VS (960 μL)

Note: AS, Auditory seed in left primary/secondary auditory cortex; VS, visual seed in right primary/secondary and ventral visual cortex; AV, attend-visual; AA, attend-auditory; Rpl, replication; BAs, Brodmann areas; μL, microliters; n.s., not significant; DLPFC, dorsolateral prefrontal cortex; FEF, frontal eye fields; SFG, superior frontal gyrus; dACC, dorsal anterior cingulate gyrus; R., right; L., left.

A fcMRI and Modality
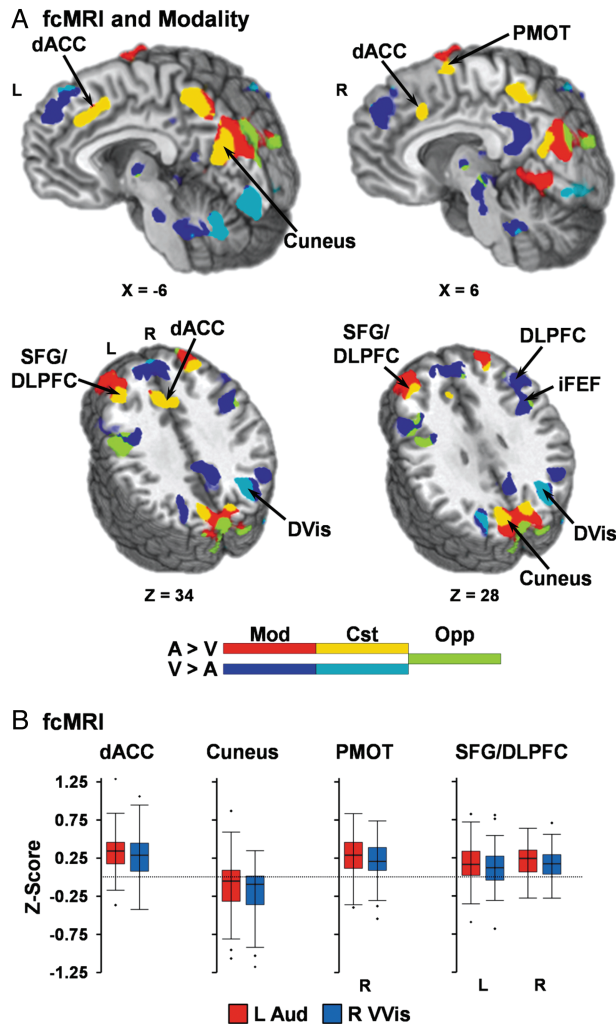
Figure 5. (A) compares significant (minimum P < 0.005 and minimum cluster size = 1088 µL) connectivity (fcMRI) results from the left auditory (L Aud) versus right primary/secondary/ventral visual (R VVis) seeds with results from the main effect of modality (Mod). Locations of sagittal (X) and axial (Z) slices are given according to the Talairach atlas for the left (L) and right (R) hemispheres. Regions that exhibited activation only during the main effect of modality are depicted in red (attend-auditory (AA) > attend-visual (AV)) or dark blue (AV > AA). Regions that exhibited consistent (Cst) task/fcMRI effects are depicted in yellow for auditory (A) inputs (AA > AV and L Aud > R VVis) or cyan for visual (V) inputs (AV > AA and R VVis > L Aud). Finally, regions demonstrating opposing (Opp) modality-specific and fcMRI activation (e.g., AV > AA and L Aud > R VVis) are depicted in chartreuse. (B) depicts fcMRI strength (Fisher Z-scores) for auditory (red) or visual (blue) seeds in regions exhibiting consistent effects for auditory inputs (A) including bilateral superior frontal gyrus (SFG)/dorsolateral prefrontal cortex (DLPFC), dorsal anterior cingulate (dACC), right premotor cortex, and cuneus.

effect, with evidence of both increased interference from visual relative to auditory distractors and greater magnitude/volume of ARM within visual cortex.

The majority of previous studies on conflict resolution have typically focused on tasks with unisensory stimulation (Ridderinkhof et al. 2004; Roberts and Hall 2008). Current results indicate that the CCN functions in a supramodal fashion to resolve conflict, with incongruent trials resulting in both increased reaction times and increased activation within medial (ACC/pre-SMA) and lateral (anterior insula/VLPFC, DLPFC extending into FEF) prefrontal cortex, posterior parietal cortex, striatum and thalamus

regardless of the sensory modality for directed attention. In contrast to previous unisensory studies, the pSTS also showed increased activation during incongruent trials. Located between the auditory and ventral visual streams, the pSTS is topographically organized (Dahl et al. 2009) and plays a critical role in integrating audio-visual information during basic (Beauchamp et al. 2004; Calvert and Thesen 2004) and higher-order (Deen et al. 2015) cognitive tasks. Current and previous findings suggest that the pSTS also represents a unique node of the multisensory CCN (Mayer et al. 2009, 2011; Chen and Zhou 2013), although future studies are needed to determine whether the pSTS directly modulates unisensory cortex or is modulated by other cortical regions.

In the current task, multisensory stimuli were simultaneously presented (Donohue et al. 2011) and spatial/temporal task contexts minimized to reduce known biases toward visual or auditory information (Weissman et al. 2004; Mayer et al. 2009). However, faster reaction times, fewer errors and reduced behavioral markers of interference (incongruent–congruent trials) were still observed for attend-visual relative to attend-auditory trials. Reduced DMN deactivation and transcallosal inhibition of the right sensorimotor cortex were also observed during the attend-visual trials. These neuronal markers of inhibition have been previously associated with task difficulty (McKiernan et al. 2003; McGregor et al. 2014), and collectively provide further support for the dominance of visual information during typical multisensory scenarios (Colavita 1974; Posner et al. 1976; Donohue et al. 2013) as well as for reduced neuronal competition within the visual system from auditory distractors (Schmid et al. 2011).

Similarly, ARM (i.e., positive modulations) were more robustly present in bilateral visual (primary, secondary, ventral, and dorsal visual streams) relative to auditory cortices. Moreover, similar to our previous results (Mayer et al. 2009), only visual cortex exhibited ARM that were differentially activated by increased perceptual/cognitive load (Frequency × Modality interaction). An unanticipated finding were the negative ARM (i.e., deactivation) observed within the cuneus and superior aspects of lingual gyrus during attend-visual relative to the attend-auditory trials. Blood–oxygen level-dependent (BOLD) deactivations in visual cortex have been previously observed during unisensory visual stimulation (Shmuel et al. 2002), and could represent a similar inhibitory phenomenon during multisensory stimulation. Alternatively, BOLD deactivation could also represent direct cross-modal input from auditory cortex (Kayser and Logothetis 2007; Driver and Noesselt 2008). Additional support for this hypothesis comes from fcMRI analyses, as connectivity in this region was greater for auditory relative to visual cortex for task-based seeds (Supplementary Figs 1–3).

Contrary to a priori predictions, activity in the prefrontal cortex was stratified dependent on the modality for focused attention. Similar to nonhuman primate studies (Barbas and Mesulam 1985; Medalla and Barbas 2014), the attend-auditory condition resulted in greater activation in bilateral SFG/rostral DLPFC, dACC and right premotor cortex whereas the attend-visual condition resulted in greater activation in the bilateral caudal DLPFC and iFEF. Connectivity analyses confirmed that prefrontal cortical regions preferentially modulated for attending to auditory stimuli also demonstrated higher connectivity for auditory relative to visual cortex across multiple samples and methodologies. In contrast, caudal lateral prefrontal regions (caudal DLPFC and iFEF) that were preferentially modulated during the attend-visual condition exhibited either no significant differences or opposing patterns of connectivity for visual versus auditory cortex.

As previously discussed, the caudal LPFC (i.e., DLPFC and FEF) has been strongly implicated in visual attention and modulation of visual cortex from both nonhuman primate and human neuroimaging/stimulation studies (Corbetta and Shulman 2002; Moore and Armstrong 2003; Egner and Hirsch 2005; Bressler et al. 2008; Tremblay et al. 2015). The role of the rostral LPFC (greater auditory modulation) is less understood, with current results demonstrating minimal overlap between rostral LPFC and the supramodal regions involved in the resolution of multisensory conflict (see Figs 2 and 3). Nonhuman primate studies indicate that the rostral LPFC is almost exclusively associated with auditory input whereas caudal LPFC areas have a "higher ratio" of visual relative to auditory connections (Barbas and Mesulam 1985; Medalla and Barbas 2014).

Thus, current and previous (Beauchamp et al. 2004; Braga et al. 2013) findings therefore suggest a more flexible, supramodal use of caudal LPFC across multiple cognitive functions and sensory domains. The majority of studies implicating the caudal LPFC in top-down modulation of sensory cortex focused exclusively on visual tasks/visual cortex (Braga et al. 2013) and did not necessarily consider other modalities of sensory input. In contrast, the rostral LPFC appears to be specialized for auditory input across both human and nonhuman primates (Medalla and Barbas 2014). Both jackknife results and replication of prefrontal stratification for the majority of auditory regions during fcMRI analyses (i.e., across independent samples and multiple methodologies) provides critical evidence for the generalizability of these findings. Importantly, replication of fcMRI results was not successful for all prefrontal regions (i.e., R SFG/DLPFC, L iFEF, and L DLPFC) across both independent samples and multiple methodologies, suggesting the need for other studies to rule-out any priming effects that were specific to multisensory stimulation.

It is also critical to note that differences in prefrontal cortex activation between attend-auditory and visual conditions only robustly emerged with larger sample sizes (Fig. 4). This likely explains why previous studies with similar designs but smaller number of subjects (Weissman et al. 2004), and previous studies comparing sensory input in other task such as working memory (Crottaz-Herbette et al. 2004), did not observe LPFC stratification. The only other study to report a similar pattern of LPFC stratification in humans relied on individual subject analyses in a relatively small (N = 10) participant cohort (Michalka et al. 2015). Findings from this study indicated interdigitating auditory and visual areas within human LPFC that are organized in a ventral-dorsal, as well as rostral-caudal, gradient and that respond supramodally based on task demands (Michalka et al. 2015). Thus, additional studies are needed to clarify how the prefrontal cortex is organized for modality-specific inputs as well as by the specific cognitive demands of tasks (Plakke and Romanski 2014; Bahlmann et al. 2015).

Several limitations of the current experiment should be noted. First, conflict occurs at the perceptual processing level, postperceptual representation level, and/or response selection level during multisensory tasks (Chen and Zhou 2013). Due to the limited temporal precision of the hemodynamic response (Smith et al. 2012), we cannot infer the exact temporal relations that exist between prefrontal regions and unisensory cortex (e.g., feed-forward versus feedback loops) as has previously been done using single cell recordings (Buffalo et al. 2010). Moreover, recent optogenetic evidence in rodents suggests that top-down control for conflicting multisensory stimuli is casually dependent on PFC interactions with the sensory aspects of the thalamus rather than the sensory cortex itself

(Wimmer et al. 2015). Thus, additional studies are needed to clarify the neuronal circuitry of multisensory cognitive control and how these circuits are influenced by task demands.

Second, although every effort was made to eliminate experimental bias, inherent bias toward more rapid/preferential processing for visual relative to auditory words may exist. This limitation is partially mitigated by connectivity findings, as the likelihood of experimental bias and priming effects are greatly reduced during resting-state analyses, especially given the similarity of connectivity findings across 2 independent cohorts who had performed different functional tasks. Finally, the intertrial intervals were purposely fixed within each block to parametrically vary the frequency of presentation (see Mayer et al. 2009). However, the fixed timing may have resulted in the development of expectancies about upcoming stimuli, which may have increased dACC (Shenhav et al. 2013) and DLPFC (Rahnev et al. 2011) activity across all trial types.

In summary, current results from behavioral, functional and ARM analyses provide further support for the Visual Dominance Theory during typical multisensory scenarios (Colavita 1974; Posner et al. 1976; Donohue et al. 2013). However, current results highlight the existence of a rostral-caudal gradient in the LPFC (superior frontal cortex-DLPFC-FEF) for directing attention to auditory versus visual input. The medial prefrontal cortex (dACC/pre-SMA) and more caudal regions of LPFC (caudal DLPFC/iFEF) were commonly activated across general task demands (i.e., resolving stimulus conflict) as well as preferentially attending to one sensory modality. In contrast, both task and connectivity results indicate the unique representations of auditory inputs in the SFG/rostral DPLFC. Collectively, these findings suggest that the modulation of the visual cortices relies on the flexible use of the supramodal CCN during multisensory processing, whereas the more specialized rostral lateral prefrontal cortex directs attention to the auditory modality (e.g., human auditory fields).

Supplementary material

Supplementary Material can be found at http://www.cercor. oxfordjournals.org/online.

Funding

Notes

References

Badre D, Wagner AD. 2004. Selection, integration, and conflict monitoring; assessing the nature and generality of prefrontal cognitive control mechanisms. Neuron. 41:473–487.

Bahlmann J, Blumenfeld RS, D'Esposito M. 2015. The rostro-caudal axis of frontal cortex is sensitive to the domain of stimulus information. Cereb Cortex. 25:1815–1826.

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

Barbas H, Mesulam MM. 1985. Cortical afferent input to the principalis region of the rhesus monkey. Neurosci. 15:619–637.

Beauchamp MS, Lee KE, Argall BD, Martin A. 2004. Integration of auditory and visual information about objects in superior temporal sulcus. Neuron. 41:809–823.

Botvinick MM, Braver TS, Barch DM, Carter CS, Cohen JD. 2001. Conflict monitoring and cognitive control. Psychol Rev. 108:624–652.

Braga RM, Wilson LR, Sharp DJ, Wise RJ, Leech R. 2013. Separable networks for top-down attention to auditory non-spatial and visuospatial modalities. Neuroimage. 74:77–86.

Braver TS. 2012. The variable nature of cognitive control: a dual mechanisms framework. Trends Cogn Sci. 16:106–113.

Bressler SL, Tang W, Sylvester CM, Shulman GL, Corbetta M. 2008. Top-down control of human visual cortex by frontal and parietal cortex in anticipatory visual spatial attention. J Neurosci. 28:10056–10061.

Brown JW, Braver TS. 2005. Learned predictions of error likelihood in the anterior cingulate cortex. Psychobiology. 307:1118–1121.

Buffalo EA, Fries P, Landman R, Liang H, Desimone R. 2010. A backward progression of attentional effects in the ventral stream. Proc Natl Acad Sci USA. 107:361–365.

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

Chen Q, Zhou X. 2013. Vision dominates at the preresponse level and audition dominates at the response level in cross-modal interaction: behavioral and neural evidence. J Neurosci. 33:7109–7121.

Colavita FB. 1974. Human sensory dominance. Percept Psychophys. 16:409–412.

Corbetta M, Shulman GL. 2002. Control of goal-directed and stimulus-driven attention in the brain. Nature Reviews. 3:201–215.

Crottaz-Herbette S, Anagnoson RT, Menon V. 2004. Modality effects in verbal working memory: differential prefrontal and parietal responses to auditory and visual stimuli. Neuroimage. 21:340–351.

Dahl CD, Logothetis NK, Kayser C. 2009. Spatial organization of multisensory responses in temporal association cortex. J Neurosci. 29:11924–11932.

Deen B, Koldewyn K, Kanwisher N, Saxe R. 2015. Functional organization of social perception and cognition in the superior temporal sulcus. Cereb Cortex. 25:4596–4609.

Desimone R, Duncan J. 1995. Neural mechanisms of selective visual attention. Annu Rev Neurosci. 18:193–222.

Donohue SE, Appelbaum LG, Park CJ, Roberts KC, Woldorff MG. 2013. Cross-modal stimulus conflict: the behavioral effects of stimulus input timing in a visual-auditory Stroop task. PLoS ONE. 8:e62802.

Donohue SE, Roberts KC, Grent-'t-Jong T, Woldorff MG. 2011. The cross-modal spread of attention reveals differential constraints for the temporal and spatial linking of visual and auditory stimulus events. J Neurosci. 31:7982–7990.

Driver J, Noesselt T. 2008. Multisensory interplay reveals cross-modal influences on 'sensory-specific' brain regions, neural responses, and judgments. Neuron. 57:11–23.

Egner T, Hirsch J. 2005. Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. Nat Neurosci. 8:1784–1790.

Ernst MO, Banks MS. 2002. Humans integrate visual and haptic information in a statistically optimal fashion. Nature. 415:429–433.

Ernst MO, Bulthoff HH. 2004. Merging the senses into a robust percept. Trends Cogn Sci. 8:162–169.

Fetsch CR, DeAngelis GC, Angelaki DE. 2013. Bridging the gap between theories of sensory cue integration and the physiology of multisensory neurons. Nat Rev Neurosci. 14:429–442.

Fias W, Reynvoet B, Brysbaert M. 2001. Are Arabic numerals processed as pictures in a Stroop interference task? Psychol Res. 65:242–249.

Hubel DH, Henson CO, Rupert A, Galambos R. 1959. "Attention" units in the auditory cortex. Psychobiology. 129:1279–1280.

Johnson JA, Zatorre RJ. 2005. Attention to simultaneous unrelated auditory and visual events: behavioral and neural correlates. Cereb Cortex. 15:1609–1620.

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

Lauritzen TZ, D'Esposito M, Heeger DJ, Silver MA. 2009. Top-down flow of visual spatial attention signals from parietal to occipital cortex. J Vis. 9:18–14.

MacDonald AW, Cohen JD, Stenger VA, Carter CS. 2000. Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. Psychobiology. 288:1835–1838.

Mayer AR, Franco AR, Canive J, Harrington DL. 2009. The effects of stimulus modality and frequency of stimulus presentation on cross-modal distraction. Cereb Cortex. 19:993–1007.

Mayer AR, Teshiba TM, Franco AR, Ling J, Shane MS, Stephen JM, Jung RE. 2011. Modeling conflict and error in the medial frontal cortex. Hum Brain Mapp. 33:2843–2855.

McGregor KM, Sudhyadhom A, Nocera J, Seff A, Crosson B, Butler AJ. 2014. Reliability of negative BOLD in ipsilateral sensorimotor areas during unimanual task activity. Brain Imaging Behav. 9:245–254.

McKiernan KA, Kaufman JN, Kucera-Thompson J, Binder JR. 2003. A parametric manipulation of factors affecting task-induced deactivation in functional neuroimaging. J Cogn Neurosci. 15:394–408.

Medalla M, Barbas H. 2014. Specialized prefrontal "auditory fields": organization of primate prefrontal-temporal pathways. Front Neurosci. 8:77.

Michalka SW, Kong L, Rosen ML, Shinn-Cunningham BG, Somers DC. 2015. Short-term memory for space and time flexibly recruit complementary sensory-biased frontal lobe attention networks. Neuron. 87:882–892.

Moore T, Armstrong KM. 2003. Selective gating of visual signals by microstimulation of frontal cortex. Nature. 421:370–373.

Motter BC. 1993. Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. J Neurophysiol. 70:909–919.

Plakke B, Romanski LM. 2014. Auditory connections and functions of prefrontal cortex. Front Neurosci. 8:199.

Posner MI, Nissen MJ, Klein RM. 1976. Visual dominance: an information-processing account of its origins and significance. Psychol Rev. 83:157–171.

Rahnev D, Lau H, de Lange FP. 2011. Prior expectation modulates the interaction between sensory and prefrontal regions in the human brain. J Neurosci. 31:10741–10748.

Ridderinkhof KR, van den Wildenberg WP, Segalowitz SJ, Carter CS. 2004. Neurocognitive mechanisms of cognitive control: the role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. Brain Cogn. 56:129–140.

Roberts KL, Hall DA. 2008. Examining a supramodal network for conflict processing: a systematic review and novel functional magnetic resonance imaging data for related visual and auditory stroop tasks. J Cogn Neurosci. 20:1063–1078.

Ruff CC, Bestmann S, Blankenburg F, Bjoertomt O, Josephs O, Weiskopf N, Deichmann R, Driver J. 2008. Distinct causal

influences of parietal versus frontal areas on human visual cortex: evidence from concurrent TMS-fMRI. Cereb Cortex. 18:817–827.

Ruff CC, Blankenburg F, Bjoertomt O, Bestmann S, Freeman E, Haynes JD, Rees G, Josephs O, Deichmann R, Driver J. 2006. Concurrent TMS-fMRI and psychophysics reveal frontal influences on human retinotopic visual cortex. Curr Biol. 16:1479–1488.

Schmid C, Buchel C, Rose M. 2011. The neural basis of visual dominance in the context of audio-visual object processing. Neuroimage. 55:304–311.

Shenhav A, Botvinick MM, Cohen JD. 2013. The expected value of control: an integrative theory of anterior cingulate cortex function. Neuron. 79:217–240.

Shmuel A, Yacoub E, Pfeuffer J, Van de Moortele PF, Adriany G, Hu X, Ugurbil K. 2002. Sustained negative BOLD, blood flow and oxygen consumption response and its coupling to the positive response in the human brain. Neuron. 36:1195–1210.

Smith SM, Bandettini PA, Miller KL, Behrens TE, Friston KJ, David O, Liu T, Woolrich MW, Nichols TE. 2012. The danger of systematic bias in group-level FMRI-lag-based causality estimation. Neuroimage. 59:1228–1229.

Spence C, Parise C, Chen YC. 2012. The Colavita Visual Dominance Effect. In: Murray MM, Wallace MT, editors. The Neural Bases of Multisensory Processes. Boca Raton (FL): CRC Press.

Talsma D. 2015. Predictive coding and multisensory integration: an attentional account of the multisensory mind. Front Integr Neurosci. 9:19.

Talsma D, Senkowski D, Soto-Faraco S, Woldorff MG. 2010. The multifaceted interplay between attention and multisensory integration. Trends Cogn Sci. 14:400–410.

Tremblay S, Pieper F, Sachs A, Martinez-Trujillo J. 2015. Attentional filtering of visual information by neuronal ensembles in the primate lateral prefrontal cortex. Neuron. 85:202–215.

van Atteveldt N, Murray MM, Thut G, Schroeder CE. 2014. Multisensory integration: flexible use of general operations. Neuron. 81:1240–1253.

Weissman DH, Warner LM, Woldorff MG. 2004. The neural mechanisms for minimizing cross-modal distraction. J Neurosci. 24:10941–10949.

Wimmer K, Compte A, Roxin A, Peixoto D, Renart A, de la Rocha J. 2015. Sensory integration dynamics in a hierarchical network explains choice probabilities in cortical area MT. Nat Commun. 6:6177.

Yuval-Greenberg S, Deouell LY. 2007. What you see is not (always) what you hear: induced gamma band responses reflect cross-modal interactions in familiar object recognition. J Neurosci. 27:1090–1096.