UC San Diego
UC San Diego Previously Published Works

Title

The Effects of Stimulus Modality and Frequency of Stimulus Presentation on Cross-modal Distraction

Permalink

Journal

ISSN

Authors

Mayer, AR
Franco, AR
Canive, J
et al.

Publication Date

DOI

Peer reviewed

FEATURE ARTICLE
The Effects of Stimulus Modality and Frequency of Stimulus Presentation on Cross-modal Distraction

A.R. Mayer[1,2], A.R. Franco[1,3], J. Canive[4,5] and D.L. Harrington[6,7]

[1]The Mind Research Network, Pete and Nancy Domenici Hall, 1101 Yale Boulevard NE, Albuquerque, NM 87131, USA, [2]Department of Neurology, University of New Mexico School of Medicine, Albuquerque, NM 87131, USA, [3]Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, NM 87131, USA, [4]New Mexico Veterans Affairs Health Care System, Research Service, Albuquerque, NM 87108, USA, [5]Department of Psychiatry, University of New Mexico School of Medicine, Albuquerque, NM 87131, USA, [6]Department of Radiology, University of California, San Diego, San Diego, CA 92122, USA and [7]San Diego Veterans Affairs Healthcare System, Research Service, San Diego, CA 92161, USA

Selective attention produces enhanced activity (attention-related modulations [ARMs]) in cortical regions corresponding to the attended modality and suppressed activity in cortical regions corresponding to the ignored modality. However, effects of behavioral context (e.g., temporal vs. spatial tasks) and basic stimulus properties (i.e., stimulus frequency) on ARMs are not fully understood. The current study used functional magnetic resonance imaging to investigate selectively attending and responding to either a visual or auditory metronome in the presence of asynchronous cross-modal distractors of 3 different frequencies (0.5, 1, and 2 Hz). Attending to auditory information while ignoring visual distractors was generally more efficient (i.e., required coordination of a smaller network) and less effortful (i.e., decreased interference and presence of ARMs) than attending to visual information while ignoring auditory distractors. However, these effects were modulated by stimulus frequency, as attempting to ignore auditory information resulted in the obligatory recruitment of auditory cortical areas during infrequent (0.5 Hz) stimulation. Robust ARMs were observed in both visual and auditory cortical areas at higher frequencies (2 Hz), indicating that participants effectively allocated attention to more rapidly presented targets. In summary, results provide neuroanatomical correlates for the dominance of the auditory modality in behavioral contexts that are highly dependent on temporal processing.

Keywords: auditory, FMRI, paced tapping, selective attention, temporal processing, visual

Introduction

Organisms are constantly bombarded with streams of information from multiple sensory modalities that must be rapidly processed to flexibly control behavior (Johnson and Zatorre 2005; Macaluso and Driver 2005). Cross-modal-selective attention enables salient or behaviorally relevant stimuli from one modality to be processed further and extraneous stimuli to be suppressed (Desimone and Duncan 1995; Driver 2001). When multisensory information is spatially or temporally congruent and can be integrated into a unitary concept, it can enhance task performance (Eimer et al. 2002; Spence et al. 2004; Dhamala et al. 2007) and even reduce visual-spatial neglect (Van Vleet and Robertson 2006). However, when multisensory streams contain unattended information that is temporally or spatially incongruent with task-relevant information (i.e., trying to ignore other numbers while adding) performance can be impaired (Cowan and Barron 1987; Driver

and Baylis 1993; Arnell and Duncan 2002; Mayer and Kosson 2004; Spence et al. 2004). The present study examined the neural modulation that results during a selective attention task requiring paced tapping to both auditory and visual stimuli in the presence of cross-modal distractors.

Hubel et al. (1959) were one of the first to suggest that selective attention-modulated neuronal responses occurred within unisensory cortical areas where bottom-up, perceptual processing occurs. Examples of attention-related modulations (ARMs) include the appearance of new waveforms and more synchronous neuronal spiking in electrophysiological recordings (Rif et al. 1991; Woldorff et al. 1993; Desimone and Duncan 1995; O'Craven et al. 1997; Reynolds et al. 1999). However, the most commonly reported ARM observed with noninvasive neuroimaging techniques is an enhanced neural response (i.e., upregulation) of primary and secondary sensory cortices that correspond to the attended stimulus modality (Talsma et al. 2006). This has been found in electrophysiological (Alho 1992; Woods et al. 1992) and functional magnetic resonance imaging (FMRI) studies of visual (Woodruff et al. 1996; Liu et al. 2003; Johnson and Zatorre 2005; Johnson and Zatorre 2006; Degerman et al. 2007), auditory (Woodruff et al. 1996; Grady et al. 1997; Alho et al. 1999; Jancke et al. 1999; Johnson and Zatorre 2005; Degerman et al. 2007), and olfactory (Zelano et al. 2005) attention. A less reliable ARM is an attenuated or depressed response (i.e., downregulation) of sensory areas that correspond to the unattended modality (Lewis et al. 2000; Laurienti et al. 2002; Shomstein and Yantis 2004; Johnson and Zatorre 2005; Johnson and Zatorre 2006; Talsma et al. 2006). Although the exact role of ARMs is unclear, selectively enhancing the neural response in the attended modality and suppressing the response in the unattended modality may minimize the contribution of cross-modal distractors (Weissman et al. 2004; Baier et al. 2006).

Several studies of attention to multisensory stimuli have measured attention during a passive condition (Laurienti et al. 2002) or during a memory task where performance was assessed after completion of the task (Johnson and Zatorre 2006). In both paradigms, the effects of unattended information on selective attention could not be directly ascertained because behavioral measures were not obtained during performance of the task. Other studies that employed online measurement of task performance have utilized complex auditory (melodies) and visual (abstract shapes) stimuli or required judgments about stimulus characteristics such as line orientation, stimulus duration, or pitch (Weissman et al. 2004;

Johnson and Zatorre 2005; Baier et al. 2006; Degerman et al. 2007). The latter paradigms may obscure the identification of ARMs because decision-making processes in these tasks typically activate extensive neural networks (Fink et al. 2000; Rao et al. 2001). Likewise, complex and fundamentally different stimuli (e.g., melodies vs. shapes) may place differential demands on the attentional load in individual sensory modalities, potentially confounding modality-specific neural activation with stimulus complexity (e.g., a person wondering if they have heard the melody before).

The ability to selectively attend to one modality and ignore information from another can also depend on the behavioral context (Ciaramitaro et al. 2007). When disparate information is presented simultaneously in different sensory modalities, visual information predominates in spatial tasks, as exemplified by the ventriloquist and McGurk effects (Busse et al. 2005). In contrast, auditory information tends to receive precedence when tasks involve a temporal component (Shams et al. 2002; Morein-Zamir et al. 2003). The temporal advantage of auditory information during rhythmic tapping is manifested by the higher accuracy and lower variability for auditory than visually paced movements (Repp and Penel 2004; Jantzen et al. 2005; Kato and Konishi 2006) and the greater interference of auditory distractors when tapping in synchrony to a visual signal relative to the reverse (Repp and Penel 2004; Kato and Konishi 2006). These contextual differences in information processing may be a result of physiological specialization such as the direct mapping of the retina on the visual cortex or the smaller epoch of time over which auditory hair cells average compared with photoreceptors (Witten and Knudsen 2005).

The present study investigated how the brain modulates selective attention to one signal modality while ignoring a cross-modal distractor in the context of a task that contained a significant temporal component. Subjects underwent FMRI as they tapped in synchrony to a constantly paced metronome in 3 different conditions. In the multisensory attention condition, subjects tapped to synchronous auditory and visual signals that occurred at the same frequency. In the attend-visual condition, subjects tapped in synchrony to a visual metronome and were instructed to ignore asynchronous auditory distractors. These conditions were then reversed in the attend-auditory condition. The sensory input, memory requirements, and motoric requirements could then be equated across all conditions.

We predicted that rhythmic tapping would be less variable for synchronous (i.e., multimodal) than asynchronous conditions (i.e., attend-auditory and visual conditions) but that variability would be greater in the attend-visual than the attend-auditory condition. We predicted that that these behavioral effects would manifest as greater enhancement and suppression effects (i.e., ARMs) relative to the multimodal condition in corresponding primary sensory cortical areas. Specifically, we predicted greater activation in visual cortical areas in the attend-visual than in both the attend-auditory and the multimodal conditions; less activation was expected in the auditory cortex in the attend-visual condition than in the multimodal condition. We predicted that a similar pattern would be present for the attend-auditory condition (enhancement effects in auditory cortex and suppression effects in visual cortex) but that the effects would be less robust as a result of the predicted decrease in task difficulty.

We also investigated whether parametric increases in stimulus frequency affected neural activation in the cross-modal distractor and multimodal conditions in the same way as they do for unimodal stimulation. This aspect of the study was intended to extend previous findings of monotonic increases in the amplitude of the hemodynamic response within sensory cortices with increasing rates of visual (Boynton et al. 1996; Ozus et al. 2001) and auditory (Binder et al. 1994; Rinne et al. 2005) signals. A similar response has been observed during synchronized tapping to tones in primary motor areas (Rao et al. 1996; Riecker et al. 2003) and in the cerebellum up to 3 Hz (Riecker et al. 2003). As such, we expected that activation would increase monotonically in the auditory, visual, and motor cortices in response to increasing stimulus rates in the *multisensory* condition. In the conditions containing a distractor stimulus (i.e., attend-visual and attend-auditory conditions), we predicted that response time would be more variable, and ARMs in primary and secondary sensory areas would be greater as the cross-modal distractors became more frequent.

Materials and Methods

Subjects

Twenty (4 females, 16 males) adult volunteers (mean age = 39.9 years; mean Edinburgh Handedness Quotient = 75) participated in the study. A history of participants' musical abilities was not assessed. Subjects with a history of neurological disease, major psychiatric disturbance, substance abuse, or psychoactive prescriptive medications were excluded. One female (excessive motion) and one male (poor behavioral data) subject were identified as outliers (above 3 standard deviations [SDs]) and excluded from further analyses. Therefore, a total of 18 subjects were included in the final analyses. Written informed consent was obtained from all participants prior to data collection, according to institutional guidelines at the University of New Mexico.

Task

Stimuli were presented in a blocked design. Each block began with a baseline period in which a white fixation cross (visual angle = 1.54°) was presented on a black background. Participants were instructed to maintain fixation on the cross throughout the course of the experiment. The duration of the baseline period was randomly varied between 10 and 14 s to prevent the development of temporal expectations and to allow for the best sampling of the hemodynamic response in the regression model (Burock et al. 1998). During the task, participants were instructed to bimanually tap their fingers (with the exception of the thumb) in synchrony with the onset of a reversing checkerboard (visual angle = 19.42° × 14.88° duration = 100 ms) and/or a pure tone (1000 Hz with a 10-ms linear rise and fall; duration = 100 ms) that were presented at standard intervals of 2000 (0.5 Hz), 1000 (1 Hz), or 500 ms (2 Hz).

In the *multimodal attention condition*, participants were instructed to attend to and tap in synchrony with both an auditory stimulus and a visual stimulus, which were simultaneously presented at the same frequency (0.5, 1, or 2 Hz). In the *attend-auditory* and *attend-visual conditions*, subjects were instructed to selectively attend to, and tap in synchrony with, either an auditory or a visual stimulus, respectively, while ignoring the stimulus in the other modality. As depicted in Table 1, in both the attend-auditory and attend-visual conditions, the stimulus in the ignored modality always occurred at a different frequency so that the ignored stimulus occurred both in and out of phase with the attended stimulus across the 8-s trial duration. Specifically, Table 1 (first column) shows that there were 2 trial types for each attended stimulus rate based on the frequency of the unattended modality (e.g., attended-auditory stimuli at 0.5 Hz were always paired with visual distractors occurring at either 1 or 2 Hz). The order of trials was pseudorandomized across all 6 functional neuroimaging runs.

Figure 1 specifically illustrates the timeline of each trial event. Each trial was preceded by a warning signal, which consisted of a visual icon (visual angle = 5.54° × 4.51°) and a 500-Hz tone, both lasting 1000 ms.

Table 1
Study design

Attention conditions	Stimulus types				
	Attended modality	Ignored modality	Rate (Hz) of attended/ ignored stimuli		
Multimodal attention[a]	Auditory and visual		0.5	1	2
Attend-auditory (A)	Auditory	Visual	0.5/1	1/0.5	2/0.5
Attend-auditory (B)	Auditory	Visual	0.5/2	1/2	2/1
Mean[b]	Auditory	Visual	0.5/1.5	1/1.25	2/0.75
Attend-visual (A)	Visual	Auditory	0.5/1	1/0.5	2/0.5
Attend-visual (B)	Visual	Auditory	0.5/2	1/2	2/1
Mean[b]	Visual	Auditory	0.5/1.5	1/1.25	2/0.75

[a]In the multimodal attention condition, auditory and visual stimuli were presented synchronously at the same rate (Hz).
[b]In the attend-auditory and attend-visual conditions, the mean rates of attended and ignored stimuli were averaged across 2 trials ([A + B]/2) based on the attended stimulus rate to form a single condition. Thus, the mean rate of attended stimuli remained at 0.5, 1, and 2 Hz, whereas the rate of the ignored stimuli were an average of the other 2 frequencies (e.g., for both attend 0.5 Hz conditions, the ignored stimuli occurred at an average rate of 1.5 Hz ([1 Hz + 2 Hz]/2).

The visual icon was a pictogram of an eye, an ear, or a hand that also cued the modality or modalities for focused attention. The cue for the multimodal attention condition was a hand. In the selective attention conditions, the pictogram of an eye cued the subject to tap in synchrony with the checkerboard and ignore the tone; the pictogram of the ear cued the subject to tap in synchrony with the tone and ignore the checkerboard. After the presentation of the signal cue, the baseline fixation stimulus was presented again for 1000 ms, followed by the start of the reversing checkerboard and target tones for a period of 8000 ms.

Subjects rested supine in the scanner with their head secured by chin and forehead straps, with additional foam padding to limit head motion within the head coil. Presentation software (Neurobehavioral Systems) was used for stimulus presentation, synchronization of stimulus events with the MRI scanner, and the collection of response time data for offline analyses. Due to software limitations with the Presentation program, only responses from the right index finger were recorded during the experiment. All participants were required to demonstrate competency on the task in a separate practice session before proceeding to the scanner environment. Specifically, participants first demonstrated that they were capable of tapping at the specified frequency based on the cue information (e.g., attend-auditory or attend-visual trials). Participants were then asked to repeat the practice session with an emphasis on reducing the trunk and head motion that can sometimes accompany rhythmic tapping.

MR Imaging
At the beginning of the scanning session, high resolution T1 (time echo [TE] = 4.76 ms, repetition time [TR] = 12 ms, 20° flip angle, number of excitations [NEX] = 1, slice thickness = 1.5 mm, field of view [FOV] = 256 mm, resolution = 256 × 256) and T2 (TE = 64 ms, TR = 9000 ms, 180° flip angle, NEX = 1, slice thickness = 1.8, FOV = 256 mm, resolution = 256 × 256) anatomic images were collected on a 1.5-Tesla Siemens Sonata scanner. For each of the 6 imaging series, 201 echo-planar images were collected using a single-shot, gradient-echo-planar pulse sequence (TR = 2000 ms, TE = 36 ms, flip angle = 90°, FOV = 256 mm, matrix size = 64 × 64). The first image of each run was eliminated to account for T1 equilibrium effects, leaving a total of 1200 images for the final analyses. Twenty-eight contiguous sagittal 5-mm thick slices were selected to provide whole-brain coverage (voxel size: 4 × 4 × 5 mm).

Image Processing and Statistical Analyses
Functional images were generated using Analysis of Functional Neuro-Images software package (Cox 1996). Time-series images were spatially registered in both 2- and 3-dimensional space to minimize effects of
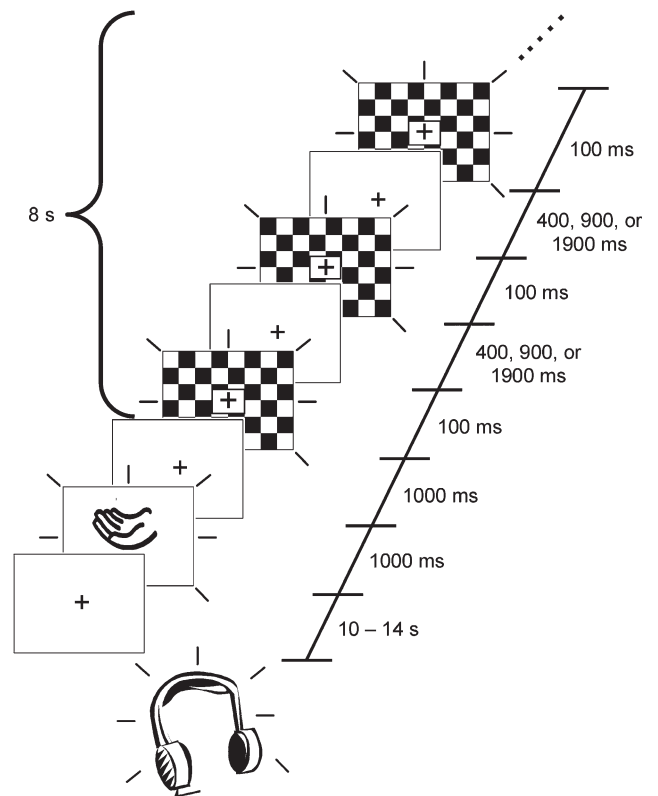

Figure 1. Diagrammatic representation of the multimodal condition. Each trial began with a variable fixation period, followed by the presentation of a cue (eye, ear, or hand) indicating the modality or modalities for focused attention. The auditory (tones) and visual (reversing checkerboard) stimuli were then presented for 8 s. The amount of time between successive target stimuli varied between 500 and 2000 ms, corresponding to stimulus presentation rates of 0.5, 1, or 2 Hz. Stimuli were presented at either different (the attend conditions) or the same (the multimodal condition) rate in the auditory and visual modalities. The background color of the screen and the fixation cross has been reversed in this cartoon to facilitate presentation.

head motion, temporally interpolated to correct for slice-time acquisition differences and despiked. A deconvolution analysis was used to generate one impulse response function (IRF) for each of the conditions on a voxel-wise basis. Each IRF was derived relative to the baseline state (fixation plus ambient noise) and based on the first 11 images (22 s) following the onset of the cue (total trial length varied from 20 to 24 s). An estimation of percent signal change (PSC) was then calculated by summing the coefficients for the images occurring 8–12 s poststimulus onset and dividing by the model intercept. The PSC maps were then converted to a 1-mm^3 standard stereotaxic coordinate space (Talairach and Tournoux 1988) and spatially blurred using a 4-mm Gaussian full-width half-maximum filter.

We first used a multiple regression analysis to identify areas that exhibited rate-dependent increases in activation related to the frequency of the multimodal auditory and visual signals and the motor response. Specifically, the *constant term* from the regression analysis identified areas that were activated during synchronized tapping compared with baseline, whereas the *linear effect* identified areas that also showed significant monotonic changes (increases or decreases) in response to changes in the frequency of sensory stimulation and the associated motor response. These analyses were conducted primarily to extend previous studies of stimulus frequency effects using unimodal visual and auditory stimuli (Binder et al. 1994; Boynton et al. 1996; Rao et al. 1996; Rinne et al. 2005).

In the second set of analyses, the multimodal and selective attention conditions were directly compared separately for when attention was allocated to the auditory or the visual modality. To set up the statistical

models to test our hypotheses, it was necessary to match both the frequency of the attended and ignored *auditory* stimulus rate and the attended and ignored *visual* stimulus rate (see Table 2). Two 3 × 3 repeated-measures analyses of variance (ANOVAs) were conducted to achieve this objective. In the first analysis, the within-subjects factors were coded based on the auditory modality (multimodal, attend-auditory, ignore-auditory) and attended-auditory stimulation rate (0.5, 1, or 2 Hz). In the second analysis, the within-subjects factors were coded based on the visual modality.

Table 2 demonstrates several important aspects regarding the main effects and the interaction from these 2 ANOVAs. First, Table 2 (footnotes b and c) shows that the main effect of attention condition was identical for both ANOVA models because the modality of the ignored and attended stimuli was interchanged across the 2 analyses (e.g., only cell order was changed). In addition, the average of the attended and ignored stimuli rates is equivalent across the 3 attention conditions (Table 2, column 5; footnote a) so that the main effect of condition (i.e., row means) was not confounded by differences in stimulus rate. Second, Table 2 demonstrates that main effects of rate (column means) should be present within cortical regions corresponding to the attended modality (Table 2; second to last row, footnote d), but not for cortical areas corresponding to the ignored modality where stimulus frequency was equivalent (Table 2; last row, footnote e). For example, in the ANOVA testing the main effect of rate for the visual stimulus, significant rate effects are expected in visual cortex (due to increasing rates of the average visual stimulation) but not in auditory cortex (due to constant rates of the average auditory stimulation). Finally, the interaction (condition × rate) was necessarily confounded by the study design with the exception of the cortical areas corresponding to the attended stimulus modality (i.e., auditory cortex in ANOVA testing auditory stimulus rate; visual cortex in ANOVA testing visual stimulus rate). Specifically, for all other brain regions, the attended stimulus rates in the "ignore" condition differed from attended stimulus rates in the "attend" condition. For example, while auditory stimuli were held at a constant rate across the 3 conditions in the auditory ANOVA (bolded for effect in column 4 of Table 2), the visual stimulus rate was varied. This would likely result in "variable" activation (i.e., a significant interaction effect) in the visual cortex as a result of the different physical properties of the stimulus (i.e., visual stimuli were presented at either 0.5 [multimodal] or either 1 or 2 Hz when auditory stimuli were matched at the 0.5-Hz frequency) rather than as a result of a true interaction effect.

To minimize false positives, a parametric voxel-wise threshold corresponding to $P < 0.005$ and a minimum cluster size of 480 μL were adopted for the regression and the 2 ANOVAs (Forman et al. 1995). These thresholds were derived from 10 000 Monte Carlo simulations, which demonstrated that the chance of probability of obtaining a significant activation cluster for an entire volume (Type I error) was less than $P < 0.05$.

Results

Behavioral Results

Response times were continuously recorded from the onset of the reversing checkerboard and target tones. To ensure that participants performed the task accurately, we calculated the amount of time that elapsed between successive behavioral response times (i.e., the intertap interval [ITI]). ITIs that were greater or less than one standard interval of the target response (correct response = 0–4000 ms ITI for 0.5 Hz, 0–2000 ms ITI for 1 Hz, 0–1000 ms ITI for 2 Hz) were classified as errors and excluded from further processing. Less than 0.5% of the trials in each condition were classified as an error based on this criterion, indicating that performance was highly accurate. Therefore, this behavioral measure was not analyzed further. The remaining responses were then used to calculate the mean ITI, SD of the ITI, and coefficient of variation (COV) (COV = [SD of ITI]/standard interval) for each subject and each condition. The ITI provided a measure for assessing whether participants correctly performed the task. The SD of the ITI was divided by its respective standard interval to compute the COV, which is a measure of processing efficiency (Spencer and Ivry 2005). In the present study, the COV provided a measure of the impact of cross-modal distractors on processing efficiency (Repp and Penel 2004), such that larger a COV was expected when there was a distractor than when there was not (i.e., multimodal condition).

Two 3 × 3 repeated-measures ANOVAs with attention condition (multimodal, attend-auditory, attend-visual) and rate of attended modality (0.5, 1, and 2 Hz) as the within-subjects factor were conducted on the ITI and the COV data. Examination of Figure 2A indicates that the attention manipulation was successful, as the ITIs for all the conditions were centered around the frequency of the target interval (i.e., ITI for 0.5 Hz was ≈2000 ms, 1 Hz ≈ 1000 ms, and 2 Hz ≈ 500 ms). For ITI, significant effects were found for attention condition ($F_{2,34} = 5.1$, $P < 0.01$), stimulus rate ($F_{2,34} = 7837.9$, $P < 0.001$),

Table 2

Coding scheme for separately testing the effects of attending to either auditory or visual signals while ignoring the other modality

Conditions for testing visual stimulus rate and ARMs

Attention condition	Attended modality	Ignored modality	Rate (Hz) of attended/ignored stimuli			Mean rate (Hz) attended/ignored for condition[a]
Multimodal	Auditory and visual		**0.5**	**1**	**2**	1.16
Attend-auditory [b]	Auditory	Visual	**0.5**/1.5	**1**/1.25	**2**/0.75	1.16/1.16
Ignore-auditory [c]	Visual	Auditory	1.5/**0.5**	1.25/**1**	0.75/**2**	1.16/1.16
	Mean rate (Hz) of auditory stimuli[d]		**0.5**	**1**	**2**	
	Mean rate (Hz) of visual stimuli[e]		1.16	1.16	1.16	

[a]The mean Hz of multimodal, attended, and ignored stimuli are equivalent when averaged across the 3 different frequency rates, thereby ensuring that main effects of condition are not confounded by differential effects of rate (rates bolded in table).

[b]In both the attend-auditory and attend-visual conditions, the rates of attended stimuli are based on the attended stimulus rate (see Table 1). The rate of the attended versus ignored stimuli are separated by a "/" for the fourth and fifth columns of the table.

[c]In both the ignore-auditory and ignore-visual conditions, the 2 trial types were averaged based on frequency (0.5, 1, or 2 Hz) of the ignored stimuli. Thus, the mean rate of ignored stimuli remains fixed at 0.5, 1, and 2 Hz, and these conditions can be directly compared with the attended modality conditions, as the stimulus frequencies are matched (e.g., auditory cortex during attend-auditory condition).

[d]The average rate of attended and ignored stimuli was 0.5, 1, or 2 Hz for visual stimuli in the visual ANOVA and auditory stimuli in the auditory ANOVA (see footnote c). Therefore, a main effect of rate should be present in the visual, but not auditory, cortex for the visual ANOVA.

[e]In contrast, the average rate of the ignored stimuli are equivalent when averaged across the 3 different attend conditions, suggesting that a main effect of rate should not be present for cortical regions corresponding to the unattended modality (e.g., auditory cortex in visual ANOVA).
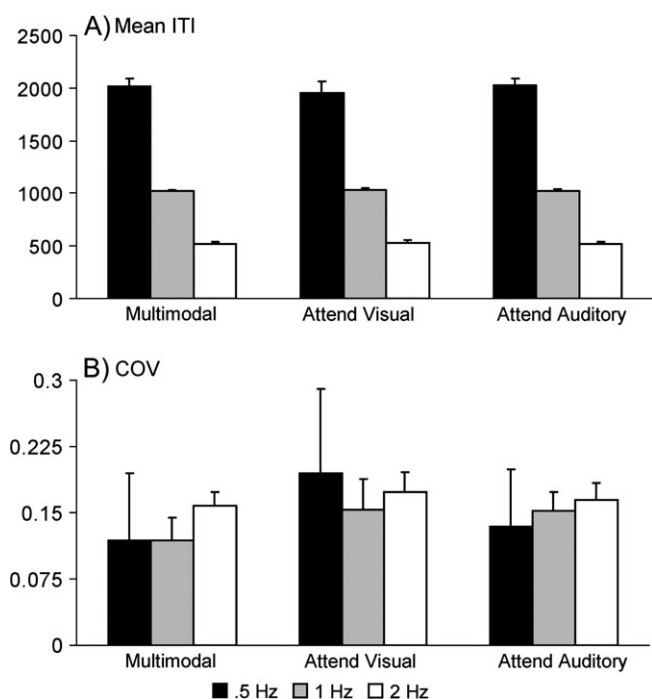
Figure 2. Behavioral data from the current experiment. Panel (*A*) presents the mean ITI and SD (error bars) for the multimodal, attend-visual, and attend-auditory conditions at 0.5 (black bar), 1 (gray bar), and 2 (white bar) Hz. Examination of the mean ITI indicates that the attentional manipulation was successful as participants in all conditions tapped at the frequency specified by the cue. Panel (*B*) exhibits the mean COV and SD (error bars) for all conditions.

and the interaction ($F_{4,68} = 12.2$, $P < 0.001$). To understand the source of the condition × rate interaction, simple effect tests were conducted comparing the attention conditions at each stimulus rate. A significant condition effect was present only at 0.5 Hz ($F_{2,16} = 7.46$, $P < 0.005$), and follow-up post-hoc *t*-tests indicated that the mean ITI in the attend-visual (mean = 1952.4 ms) was significantly shorter than the mean ITIs of both the attend-auditory (mean = 2027.0 ms; $t_{17} = -3.5$, $P < 0.005$) and multimodal (mean = 2017.4 ms; $t_{17} = -3.7$, $p < 0.005$) conditions following correction for family-wise error.

For the analyses of COV, significant effects were found for attention condition ($F_{2,34} = 23.7$, $P < 0.001$) and the interaction ($F_{4,68} = 5.3$, $P < 0.001$). A nonsignificant trend was also present for the effect of rate ($F_{2,34} = 3.1$, $P = 0.06$). Paired *t*-tests were first conducted on the marginal means of the attentional conditions to examine our a priori hypothesis of greater interference in the attend-visual than in the attend-auditory condition. The results indicated that processing efficiency was worse in the attend-visual (mean = 0.175) than in both the attend-auditory (mean = 0.154; $t_{17} = 3.3$, $P < 0.005$) and the multimodal (mean = 0.134; $t_{17} = 7.4$, $P < 0.001$) conditions. Processing efficiency was also worse in the attend-auditory than in the multimodal ($t_{17} = 3.4$, $P < 0.005$) condition. Simple effect tests were then conducted comparing the effect of attention condition at each frequency to identify the source of the interaction. Significant main effects were followed by post-hoc *t*-tests (corrected for family-wise error at $P < 0.005$). At 0.5 Hz, there was a main effect of condition ($F_{2,16} = 9.2$, $P < 0.001$), and paired *t*-tests indicated that the COV was greater in the attend-visual than in the multimodal ($t_{17} = 4.4$, $P < 0.001$) and the attend-auditory ($t_{17} = 3.4$, $P < 0.005$) conditions. There

was no difference between the attend-auditory and the multimodal conditions. At 1 Hz, the main effect of attention condition ($F_{2,16} = 22.3$, $P < 0.001$) was due to a larger COV in the attend-auditory ($t_{17} = 6.2$, $P < 0.001$) and the attend-visual ($t_{17} = 4.8$, $P < 0.001$) conditions when compared with the multimodal condition. There was no difference between the attend-auditory and attend-visual conditions. At 2 Hz, the main effect of attention condition ($F_{2,16} = 10.3$, $P < 0.005$) was due to a greater COV in the attend-visual than in the multimodal ($t_{17} = 4.7$, $P < 0.001$) condition, with a significant trend emerging in the attend-visual versus attend-auditory comparison ($t_{17} = 1.9$, $P = 0.075$).

Functional Results

Multimodal Condition

A multiple regression analysis was conducted on the PSC data from the multimodal attention condition to identify the regions activated during paced tapping to temporally congruent audio-visual stimuli and to verify that primary sensori-motor regions would be sensitive to increasing rates of multimodal stimulus presentation (i.e., linear effect of rate). Table 3 and Figure 3*A* show all regions that were activated by the task but that did not demonstrate a significant linear effect of rate. In addition to sensory and motor areas, activation was also observed in the supplementary motor area (SMA) (Brodmann area [BA] 6) extending into the anterior cingulate gyrus (BA 24), bilateral middle frontal gyrus (BA 9), bilateral inferior frontal and precentral gyrus (BA 9/44) extending into the anterior aspects of the insula (BA 13), the right middle temporal gyrus (BA 37), bilateral inferior parietal lobule (BA 40), the left putamen, and several clusters within the cerebellum. Areas of deactivation (Table 3 and Fig. 3*B*) were observed in the bilateral pre-SMA (BA 9) extending into the anterior cingulate gyrus (BA 32), left medial and superior frontal gyrus (BA 6/8), and the anterior and posterior aspects of the left middle temporal gyrus (BA 21).

Table 4 and Figure 3*C* describe all regions that exhibited significant linear increases in activation as the rate of the stimulus and response increased. These regions included the bilateral motor cortex (BAs 4/3), bilateral auditory cortex (BAs 13/41), the bilateral visual cortex and cuneus (BA 18), and the right fusiform and parahippocampal gyri (BA 19/37). Additional activations included the bilateral subthalamic nuclei and the right culmen (lobules IV and V) of the cerebellum.

Multimodal and Selective Attention Effects

Next, we conducted 2 separate 3 × 3 (attention condition [Attend, Ignore, Multimodal] × attended rate [0.5, 1, or 2 Hz]) repeated-measures ANOVAs to investigate areas that exhibited ARMs when attention was directed to the auditory or visual modality. As previously noted, the main effect of condition was identical across both ANOVAs (see Table 2). ARMs were classified as "strong positive" if activity was greater in the attend condition than both the multimodal and ignore conditions (e.g., attend-visual > attend-auditory, and attend-visual > multimodal), "positive" if activation in the attend condition was greater than either the multimodal or the ignore conditions (e.g., attend-visual > attend-auditory, or attend-visual > multimodal), "strong negative" if suppression was greater in both attend conditions than in the multimodal condition (i.e., multimodal > attend-visual, and multimodal > attend-auditory), or "negative" if suppression was only present for one of the attend conditions

Table 3

Regions showing significant activation or deactivation (constant term) during paced tapping in the multimodal attention condition

Region	Side	Activation					Deactivation				
		BA	x	y	z	Volume (mL)	BA	x	y	z	Volume (mL)
Frontal lobe											
SMA and anterior cingulate gyrus	M	6/24	4	−3	51	6.416					
SMA and anterior cingulate gyrus	M						9/32	−1	43	13	10.049
Medial and superior frontal gyrus	L						6/8	−16	28	42	6.375
Middle frontal gyrus	R	9	33	32	28	2.705					
	L	9	−34	32	27	1.645					
Inferior frontal gyrus, precentral gyrus and insula	R	9/44/13	44	6	13	6.496					
	L	9/44/13	−42	3	9	5.946					
Pre- and postcentral gyrus	R	4/3	39	−13	49	2.092					
	L	4/3	−37	−24	52	1.873					
Temporal lobe											
Middle/superior temporal gyrus and insula	R	22/13	53	−35	14	4.877					
Middle temporal gyrus	R	37	54	−53	−3	0.989					
	L							−45	−71	22	0.604
	L						21	−51	−8	−13	0.622
Parietal lobe											
Inferior parietal lobule and supramarginal gyrus	R	40	38	−45	41	3.595					
Inferior parietal lobule	L	40	−41	−41	42	1.541					
Temporal parietal juncture	L	40/22	−57	−33	23	0.648					
Precuneus	L		−28	−50	33	0.620					
Occipital lobe											
Lingual gyrus and cuneus	R	17/18	8	−81	6	9.828					
	L	17/18	−12	−76	−1	4.341					
Lingual gyrus	R	19	21	−63	5	0.569					
Subcortical											
Putamen	L		−23	−5	6	4.058					
Subthalamic nuclei	L		−5	−21	−7	0.925					
Cerebellum											
Culmen (IV–VI) and fusiform gyrus	R		16	−52	−14	7.286					
	L		−18	−57	−16	9.070					

Note: Side refers to the hemisphere showing activation where M = midline, L = left hemisphere, and R = right hemisphere. The Brodmann area (BA), the center of mass in Talairach coordinates (x, y, z) and volume are specified for each area of activation.

(i.e., multimodal > attend-auditory, or multimodal > attend-visual).

Main effects of condition were observed in widespread cortical and subcortical networks (see Fig. 4A and B; Table 5). In these regions, follow-up pair-wise t-tests were then conducted contrasting the PSC between the attention conditions to identify the source of the main effects. Table 5 shows that the majority of regions exhibited greater activation during the attend-visual condition. Specifically, *strong positive visual ARMs* (attend-visual > multimodal and attend-visual > attend-auditory; represented by 2 Δ's in Table 5) and a condition-specific deactivation or *negative ARM* (multimodal > attend-auditory) were observed in the bilateral visual cortex (BAs 19/37). Strong positive visual ARMs were also observed within the medial prefrontal cortex (BAs 6/32/24), bilateral frontal gyri (BAs 6/9), left insula, bilateral posterior parietal lobes (BAs 7/40), right thalamus, and several clusters within the cerebellum. In addition, several other cortical areas and thalamic nuclei exhibited *positive visual ARMs* (denoted by a P in Table 5).

Strong negative ARMs (multimodal > attend-visual and multimodal > attend-auditory; represented by 2 X's in Table 5) were observed in the bilateral medial frontal and anterior cingulate gyrus (BAs 11/32), right insula (BA 13), right precentral gyrus (BA 4), and the bilateral posterior cingulate gyrus (BAs 23/31). However, the PSC for the anterior and posterior medial wall areas indicated that these regions were actually deactivated less in the multimodal than the attend conditions (see Figure 4.B). In addition, *condition-specific negative ARMs* (multimodal > attend-auditory; denoted by N in Table 5) were also observed in the left motor cortex (BAs 4/3),
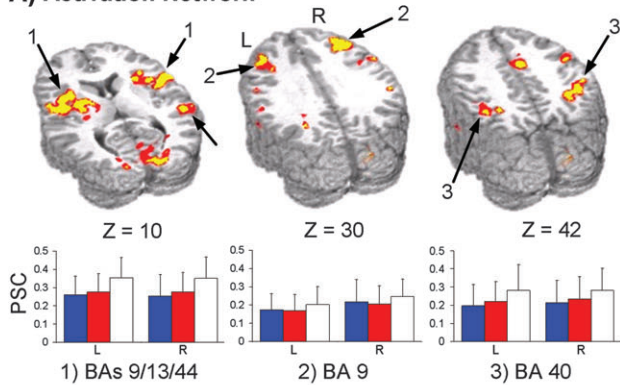
the right middle and superior temporal gyrus (BA 13), the right lingual gyrus and cuneus (BAs 17/18), the bilateral pulvinar, and the bilateral pons.

Table 6 lists all the regions that exhibited a main effect of rate in the separate ANOVAs coded for whether attention was predominantly directed to the auditory or visual modality (see Table 2). In both ANOVAs, stimulus rate had a significant effect on brain activation in several common regions (Fig. 5; yellow coloring) including bilateral SMA proper (BAs 6/31), bilateral sensory-motor cortex (BAs 4/3/2), bilateral thalamus, and the bilateral cerebellum. Large clusters of common activation were also observed bilaterally in auditory cortical areas (BAs 13/41) extending into the inferior parietal lobule (BA 40). In contrast, only a small cluster of common activation was observed in the left middle occipital gyrus and cuneus (BAs 18/19). Common rate effects in these traditionally unisensory structures were somewhat surprising given that stimulus for the unattended modality was equivalent in both analyses (see Table 2, footnote e).
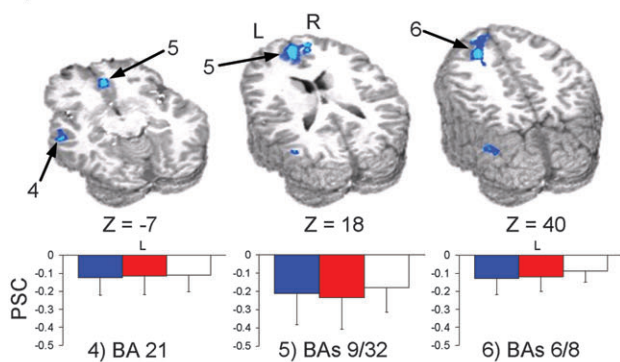
The only area exhibiting a unique rate effects in the auditory ANOVA (Fig. 5; blue coloring) was the bilateral primary and secondary auditory cortex (BA 13/22/41). In contrast, the main effect of rate in the visual ANOVA (Fig. 5; red coloring) resulted not only in widespread activation of the bilateral visual cortex (BAs 17/18/19/36/37) but also in several other structures including pre-SMA and cingulate gyrus (BAs 6/24/32), bilateral frontal gyri (BAs 6/9), bilateral anterior insula (BA 13), the bilateral posterior parietal lobule (BAs 7/40), and bilateral cerebellum.

Figure 6 and Table 7 describe regions in the primary and secondary cortical areas that showed significant attention

A) Activation Network

Z = 10   Z = 30   Z = 42

1) BAs 9/13/44   2) BA 9   3) BA 40

B) Deactivation Network

Z = -7   Z = 18   Z = 40

4) BA 21   5) BAs 9/32   6) BAs 6/8

C) Stimulus Rate Network

Z = -8   Z = 6   Z = 52

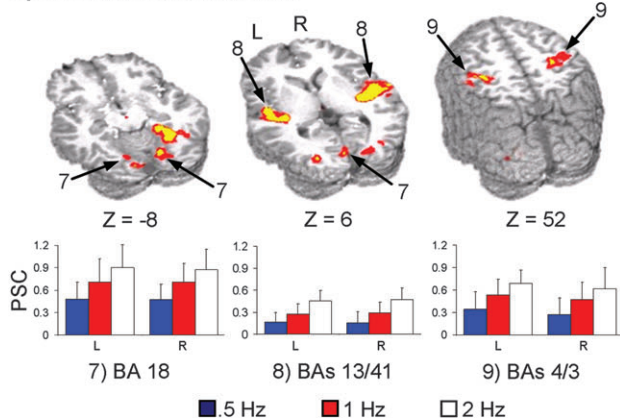7) BA 18   8) BAs 13/41   9) BAs 4/3

■ .5 Hz   ■ 1 Hz   □ 2 Hz

Figure 3. Three networks that were activated during the multimodal attention condition. Panel (*A*) displays and graphs the PSC for selected regions that demonstrated task-related activation, but that did not exhibit significant effects of stimulus rate. These regions included 1) bilateral insula and prefrontal cortex (BAs 9/13/44), 2) bilateral dorsolateral prefrontal cortex (BA 9), and 3) bilateral inferior parietal lobe (BA 40). The additional arrow (z = 10) indicates the location of activation within the posterior aspects of the right middle and superior temporal gyrus (area not graphed). Panel (*B*) displays and graphs the PSC for selected regions that exhibited negative activation during the multimodal task including the 4) left middle temporal gyrus (BA 21), 5) bilateral supplementary motor area and cingulate gyrus (BAs 9/32), and 6) left medial and superior frontal gyrus (BAs 6/8). Panel (*C*) displays regions that exhibited significant linear effects in response to increasing rates of stimulus presentation. These regions included 7) bilateral visual cortex (BA 18), 8) bilateral auditory cortex (BAs 13/41), and 9) bilateral motor cortex (BAs 3/4). For all panels, the locations of axial slices (z) are given according to the Talairach atlas. An estimate of the PSC is presented in the graphs at the bottom of the figure for the .5 (blue bar), 1 (red bar) and 2 (white bar) Hz stimulation frequencies.

Table 4

Regions showing a linear effect of stimulus rate during paced tapping in the multimodal attention condition

Region	Side	SDA activation				
		BA	x	y	z	Volume (mL)
Frontal lobe						
Pre- and postcentral gyrus	R	4/3	31	−28	53	2.259
	L	4/3	−35	−28	53	1.590
Temporal lobe						
Insula, transverse, and superior	R	13/41	44	−20	8	7.927
temporal gyrus	L	13/41	−41	−22	8	6.360
Fusiform and parahippocampal gyrus	R	19/37	26	−50	−8	2.738
Occipital lobe						
Lingual gyri	R	18	11	−74	−5	1.579
	L	18	−10	−77	−2	1.780
Inferior/middle occipital gyrus	R	18	26	−82	6	2.367
and cuneus	L	18	−26	−83	9	1.970
Subcortical						
Subthalamic nuclei	M		−3	−22	−3	0.573
Cerebellum						
Culmen (IV–V)	R		16	−49	−17	0.736

Note: Side refers to the hemisphere showing activation where M = midline, L = left hemisphere, and R = right hemisphere. The Brodmann area (BA), the center of mass in Talairach coordinates (x, y, z) and volume are specified for each area demonstrating stimulus-dependent activations (SDA).

condition × rate interactions. In the visual ANOVA, significant interactions were found in an area encompassing the fusiform gyrus (BA 37), parahippocampal gyrus (BA 19), lingual gyrus (BA 18, 19), the inferior and middle occipital gyri (BA 19), middle temporal gyrus (BA 37), cuneus and precuneus (BA 31) in both hemispheres. Follow-up paired *t*-tests contrasting the 3 attention conditions (i.e., multimodal, attend-visual, ignore-visual) indicated that there was significant enhancement of the hemodynamic response in the attend-visual relative to the ignore-visual conditions in all regions of interest (ROIs) of both hemispheres, but only at the 2-Hz frequency ($P < 0.005$; 480 μL). Greater activation in the attend-visual than the multimodal condition was also observed in all ROIs at 2 Hz, with the exception of the left middle temporal gyrus. In contrast, activation in the ignore-visual condition was reduced relative to the multimodal condition within the right precuneus, the bilateral middle occipital gyrus, and the right middle temporal gyrus.

For the auditory ANOVA, significant interactions were observed in the bilateral insula (BA 13), the transverse and superior temporal gyrus (BAs 22/41/42), and the temporoparietal juncture (BA 39/40). Similar to the visual modality, follow-up *t*-tests indicated that activation was greater for the attend-auditory than the ignore-auditory condition and activation was reduced for the ignore-auditory compared with the multimodal condition in all ROIs, but only at the 2-Hz frequency. However, no significant differences were found between the attend-auditory and the multimodal conditions. At the 0.5-Hz frequency, activation was greater for the ignore-auditory than both the attend-auditory and multimodal conditions bilaterally in the insula, the right superior temporal gyrus, and the right temporoparietal juncture.

Discussion

Our behavioral findings indicated that rhythmic tapping was better when auditory and visual information occurred at the

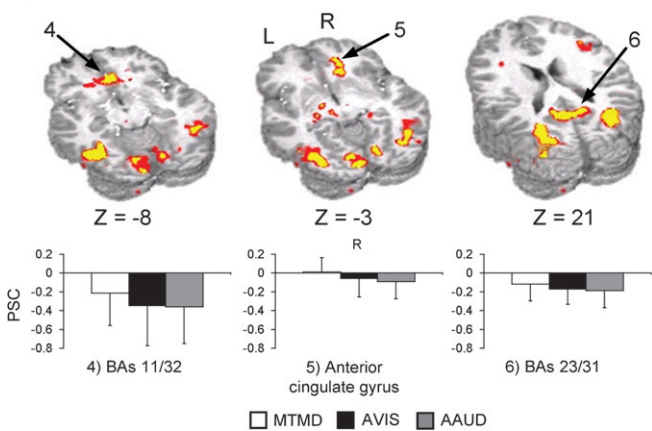## A) Positive Visual ARMs
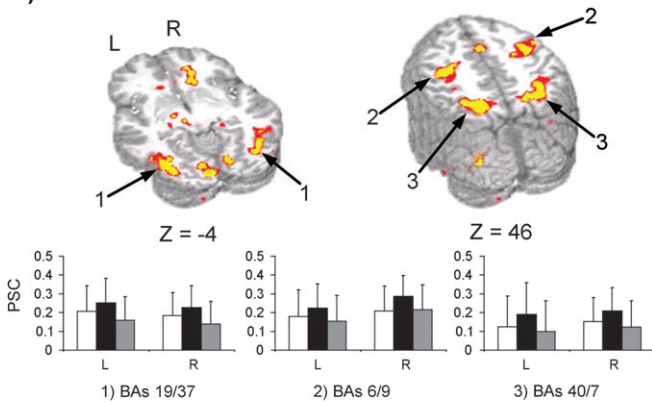
B) Negative ARMs

Figure 4. Regions that exhibited positive (panel *A*) or negative (panel *B*) ARMs as defined by the main effect of attention condition. A graph of the PSC is presented for the multimodal (MTMD; white bar), attend-visual (AVIS; black bar), or attend-auditory (AAUD; gray bar) conditions. Regions exhibiting significant positive visual ARMs (panel *A*) included 1) the bilateral ventral visual stream (BAs 19/37), 2) the bilateral frontal eye fields (BAs 6/9), and 3) the bilateral dorsal visual stream (BA 40/7). The bottom panel (panel *B*) presents ARMs that were associated with deactivation within the 4) bilateral anterior cingulate gyrus and medial frontal lobe (BAs 11/32), 5) right anterior cingulate gyrus, and 6) the posterior cingulate gyrus and precuneus (BAs 23/31). For all panels, the locations of axial slices (*z*) are given according to the Talairach atlas.

same frequency (multimodal attention condition) than when it conflicted; however, that interference was greater when subjects were selectively attending to visual signals while ignoring incongruent auditory information. These results are consistent with previous studies indicating that auditory distractors can be more difficult to ignore during synchronized tapping than visual distractors (Repp and Penel 2004; Kato and Konishi 2006). Our results further demonstrate that ignoring auditory distractors was consistently more difficult when synchronizing responses to a relatively slow-paced visual metronome (e.g., 0.5 Hz). The FMRI results paralleled the behavioral findings with evidence of extensive cortical and subcortical ARMs during the attend-visual but not the attend-auditory condition. However, contrary to our expectations, we found several traditional unisensory-auditory cortical areas that exhibited frequency effects in the *absence* of changing auditory stimulus rates (Table 6, common main effect of rate in both the auditory and visual ANOVAs) and enhanced ARMs

during the ignore-auditory condition, but only when the pacing of the attended visual signal was slowest (i.e., 0.5 Hz). We now turn to a discussion of these main findings, first considering effects of stimulus frequency during the multimodal and selective attention conditions and then discussing the implications of modality and stimulus frequency effects on ARMs.

Stimulus Frequency Effects on the Allocation of Attention to Multimodal Events

Paced tapping to synchronous multimodal stimuli was associated with a widespread pattern of activity that can generally be classified into 3 distinct networks. The first network consisted of traditional primary and secondary auditory (e.g., superior temporal gyrus), visual (e.g., lingual gyri, inferior and middle occipital cortex) and motor areas (i.e., pre- and postcentral gyrus, cerebellum), which showed linear increases in activation with stimulation frequency (i.e., Table 4). This finding is in accord with research demonstrating monotonic increases in visual (Boynton et al. 1996; Ozus et al. 2001) and in auditory (Binder et al. 1994; Rinne et al. 2005) cortices with increased unisensory stimulation rate and in the motor cortex (Rao et al. 1996; Riecker et al. 2003) as paced-synchronized tapping rate increased. Our study results generalize these findings to *multimodal* stimulation. The second network, including the SMA/anterior cingulate gyrus, the left medial and superior frontal gyrus, and the left middle temporal gyrus, demonstrated deactivation during multimodal attention and showed no effect of stimulus rate (i.e., Table 3). Several of these cortical regions have been previously identified as a default-mode network (Raichle et al. 2001; Shulman et al. 2002; Greicius et al. 2003), which is thought to mediate episodic memory processes common to passive mental activity.

A third network consisted of both cortical and subcortical structures that were engaged by tapping in time to synchronous multisensory information but did not exhibit a significant linear relationship with stimulus frequency (i.e., Table 3). This network included several heteromodal cortical areas that participate in timing and the integration of simultaneously occurring multisensory stimuli, including the bilateral middle frontal gyrus (i.e., dorsolateral prefrontal cortex [Petrides and Pandya 1999], anterior aspects of the bilateral insula and the bilateral inferior parietal lobule [Harrington et al. 2004], the putamen [Harrington et al. 2004], and the posterior aspect of the right middle and superior temporal gyrus [Calvert et al. 2001; Beauchamp et al. 2004b; Kayser and Logothetis 2007; Naghavi et al. 2007]).

Our second set of analyses (auditory and visual ANOVAs) examined the main effects of stimulus rate when attention was differentially allocated across the 2 sensory modalities (multimodal, attend, and ignore conditions). As expected, attending and rhythmically tapping to increasing frequencies of either visual or auditory metronomes resulted in common activation of the premotor cortex (SMA proper), cerebellar lobules IV and V, and primary motor cortex in both the visual and auditory ANOVA. All these regions were likely linked to the motoric requirements of the task. In addition, significant effects of stimulus rate were also observed bilaterally within the insula, superior temporal gyrus and inferior parietal lobule (superior to Heschl's gyrus centered on the lateral sulcus), and within a small volume of the left middle occipital gyrus and cuneus in both the auditory and visual ANOVA. The common activation of

Table 5

Regions showing a main effect of attention condition.

Region	Side	Activation					ARMs			
							Positive		Negative	
		BA	x	y	z	Volume (mL)	V > B	V > A	B > A	B > V
Frontal lobe										
Pre-SMA and cingulate gyrus	M	6/32/24	5	11	43	3.059	Δ	Δ		
Medial frontal and anterior cingulate gyrus	M	11/32	−2	26	−9	1.408			X	X
Anterior cingulate gyrus	R		16	30	−4	0.901			X	X
Inferior frontal, middle frontal, and precentral gyrus	R	6/9	35	−3	45	7.621	Δ	Δ		
	L	6/9	−34	−10	43	3.969	Δ	Δ		
Insula	R	13	34	17	10	1.138	P			
	R	13	41	−7	12	0.507			X	X
	L	13	−30	18	11	0.792	Δ	Δ		
Middle frontal gyrus	R	9	40	18	25	1.602	P			
Precentral gyrus	R	4	29	−22	53	0.719			X	X
Pre- and postcentral gyrus	L	4/3	−26	−31	56	1.876		P	N	
Temporal lobe										
Middle and superior temporal gyrus	R	13	48	−42	12	0.993	P		N	
Parietal lobe										
Precuneus and posterior parietal lobule	R	40/7	26	−61	37	10.423	Δ	Δ		
	L	40/7	−25	−61	40	7.818	Δ	Δ		
Posterior cingulate gyrus and precuneus	M	23/31	−2	−55	24	4.165			X	X
Occipital lobe										
Lingual gyrus and cuneus	R	17/18	12	−83	4	3.626		P	N	
	L	17/18	−10	−84	2	5.002	Δ	Δ		
Ventral visual stream	R	19/37	37	−68	7	8.544	Δ	Δ	N	
	L	19/37	−34	−70	4	10.026	Δ	Δ	N	
Subcortical										
Ventral lateral nucleus of the thalamus	R		9	−14	7	1.113	Δ	Δ		
	L		−10	−15	0	0.750		P		
Pulvinar nucleus of the thalamus	R		−17	−29	6	0.685		P	N	
	L		−21	−27	3	1.290		P	N	
Cerebellum										
Declive (VI)	R		13	−68	−15	3.193		P		
	L		−17	−70	−17	4.432	Δ	Δ		
Culmen (IV–V)	R		16	−40	−17	0.594		P		
Culmen (VI) and fusiform gyrus	R		26	−45	−23	1.500	Δ	Δ		
Vermis	M		0	−53	−26	2.176	Δ	Δ		
Pons	M		1	−23	−19	0.501		P	N	
Inferior semilunar lobule (Crus 2)	R		30	−63	−44	0.534	Δ	Δ		

Note: Side refers to the hemisphere showing activation where M = midline, L = left hemisphere, and R = right hemisphere. The Brodmann area (BA), the center of mass in Talairach coordinates (x, y, z) and volume are specified for each area of activation that was significant for the main effect of condition. The second set of columns describes whether the ROIs exhibited greater ARMs during the attend-visual (V), attend-auditory (A), or multimodal (B) attention conditions with triangles (Δ) indicating strong positive ARMs and 2 X's indicating strong negative ARMs. A single "P" indicated a positive ARM whereas a single "N" indicated a negative ARM.

these traditionally unisensory areas was unexpected given that the frequency of stimulation was held constant in these comparisons (see Table 2; footnote d) for the mostly unattended modality, suggesting that the significant rate effect may be reflective of a cognitive rather than sensory process. Moreover, it is unlikely that the activation within the superior temporal gyrus was the result of multisensory integration as the cluster location was both anterior and superior to the posterior superior temporal sulcus sensory integration region that has previously been identified in other studies (Calvert et al. 2001; Beauchamp et al. 2004a; Calvert and Thesen 2004; Macaluso and Driver 2005) and was found to be active in the right hemisphere during the multimodal condition in the present study (see Table 3). Instead, the activation of large regions of auditory cortical areas may provide a potential neuroanatomical correlate for the behavioral observation that auditory distractors are more difficult to ignore during synchronized tapping (Repp and Penel 2004; Kato and Konishi 2006). Namely, the temporal requirement of the task may have resulted in the obligatory recruitment of some auditory cortical areas regardless of whether the attention was allocated to the auditory or visual modality.

When attention was primarily focused on tapping to an auditory metronome (auditory ANOVA), main effects of rate were uniquely observed in the bilateral primary auditory cortex within Heschl's gyrus and more inferior regions of secondary auditory cortex along the superior temporal gyrus. In contrast, when selectively attending to a visual metronome (visual ANOVA), main effects of rate were uniquely found in the striate and extrastriate cortices. The unique activation of primary and secondary cortical areas corresponding to the attended modality is consistent with previous studies comparing synchronized tapping to unimodal auditory or visual stimuli for a single frequency (Jancke et al. 2000; Jantzen et al. 2005). Of greater interest, however, was our finding of a significant effect of stimulus rate in the bilateral ventral and dorsal visual streams, the bilateral frontal eye fields (Paus 1996), pre-SMA and anterior cingulate gyrus, bilateral inferior frontal and precentral gyrus, and bilateral tonsil of the cerebellum in the visual ANOVA (Fig. 5; red activation). These results are consistent with previous research demonstrating unique activation of the dorsal visual stream and ventral premotor areas during paced tapping to a unimodal visual but not an auditory metronome (Jantzen et al. 2005). Collectively, current and previous results suggest that a more distributed frontoparietal association network is engaged by attention and paced tapping to visual than to auditory metronomes,

Table 6
Regions exhibiting a main effect of stimulus rate (Hz) in the ANOVAs of the multimodal and the attend-auditory conditions and the multimodal and the attend-visual conditions

Region	Side	Activation				
		BA	x	y	z	Volume (mL)
Common areas						
Frontal lobe						
SMA and paracentral lobule	M	6/31	1	−15	52	2.580
Pre- and postcentral gyrus	R	4/3/2	35	−24	50	13.769
	L	4/3/2	−38	−26	51	10.584
Temporal lobe						
Insula, superior temporal gyrus, and inferior parietal lobule	R	13/41/40	43	−22	14	5.242
	L	13/41/40	−42	−28	18	3.499
Occipital lobe						
Middle occipital gyrus and cuneus	L	18/19	−29	−87	14	0.490
Subcortical						
Pulvinar and ventral posterior medial nucleus of the thalamus	R		13	−22	3	1.440
	L		−14	−21	2	1.073
Cerebellum						
Culmen (IV–V)	R		13	−50	−17	5.045
	L		−14	−51	−16	3.984
Rate effects for auditory analysis						
Temporal lobe						
Insula, superior and transverse temporal gyrus	R	13/22/41	48	−15	4	11.296
	L	13/22/41	−47	−18	6	12.498
Rate effects for visual analysis						
Frontal lobe						
SMA, superior frontal, and cingulate gyrus	M	6/24/32	2	16	35	4.669
Precentral and inferior frontal gyrus	R	6/9	45	4	27	4.189
	L	6/9	−45	2	27	6.210
Middle frontal gyrus	R	6	25	−1	54	1.369
	L	6	−23	−4	54	2.291
Insula	L	13	−37	11	6	1.432
Insula and inferior frontal gyrus	R	13/45/47	35	15	3	4.916
Parietal lobe						
Posterior parietal lobule and precuneus	R	7/40	19	−64	42	15.367
	L	7/40	−23	−60	43	10.850
Occipital lobe						
Visual cortex, ventral visual stream and cuneus	R	17/18/19/36/37	22	−70	6	69.392
	L	17/18/19/36/37	−22	−73	6	62.502
Cerebellum						
Tonsil (VIII and IX)	R		15	−56	−35	5.152
	L		−15	−52	−41	4.632

Note: The table first lists common regions, which were those that showed a main effect of rate in both the auditory and visual modality analyses, followed by regions that exhibited unique activation for only the auditory or visual modality. Side refers to the hemisphere showing activation where M = midline, L = left hemisphere, and R = right hemisphere. The Brodmann area (BA), the center of mass in Talairach coordinates (x, y, z) and volume are specified for each area of activation.

irrespective of motor requirements or the presence of cross-modal distractors.

### *Attention-Related Modulations*

Our ITI results demonstrated that participants accurately reproduced the experimenter-defined attended stimulus rate for all conditions (see Fig. 2A), thereby verifying that the attentional manipulation was successful (e.g., participants were rhythmically tapping at the specified frequency). In addition, paced tapping was most efficient (i.e., smallest COV) when auditory and visual signals coincided. Several of our functional results paralleled this behavioral finding. Specifically, activation was greater in the multimodal condition relative to both selective attention conditions in the right precentral gyrus (Table 5). This finding may suggest that focused attention to the motoric components of the tapping task was easiest in the absence of conflicting sensory information. The anterior aspect of the right insula was the only other area that exhibited greater activity in the multimodal than in both the selective attention conditions. This finding is likely due to the role of the insula in integrating multisensory information, which was only required in the multimodal condition (Naghavi et al. 2007). Second, several central nodes of the default-mode network

(Raichle et al. 2001; Shulman et al. 2002; Greicius et al. 2003), including the anterior and posterior cingulate gyrus, were deactivated (negative ARMs) more during the selective attention conditions, which likely engaged additional cognitive resources than during the multimodal condition (see Fig. 5). This result is consistent with reports that the level of deactivation in the default-mode network has been shown to depend on the cognitive load of the task (Esposito et al. 2006).

As in other studies (Repp and Penel 2004; Kato and Konishi 2006), our behavioral data (i.e., COV) also showed that processing efficiency was clearly the worst when attending and tapping to a visual metronome while ignoring auditory signals. These results are likely due to cognitive specialization of the auditory system for processing temporal information. Our functional data extend these behavioral results by demonstrating that when the rate of sensory stimulation was held constant (i.e., Table 5, main effect of attention condition), widespread ARMs were present for the attend-visual condition but were generally absent in the attend-auditory condition. These findings indicate that tapping in synchrony with a visual metronome while ignoring auditory distractors was more effortful and thus recruited additional neuronal resources, providing a neural correlate for the pattern of increased interference observed in the behavioral data. Specifically,
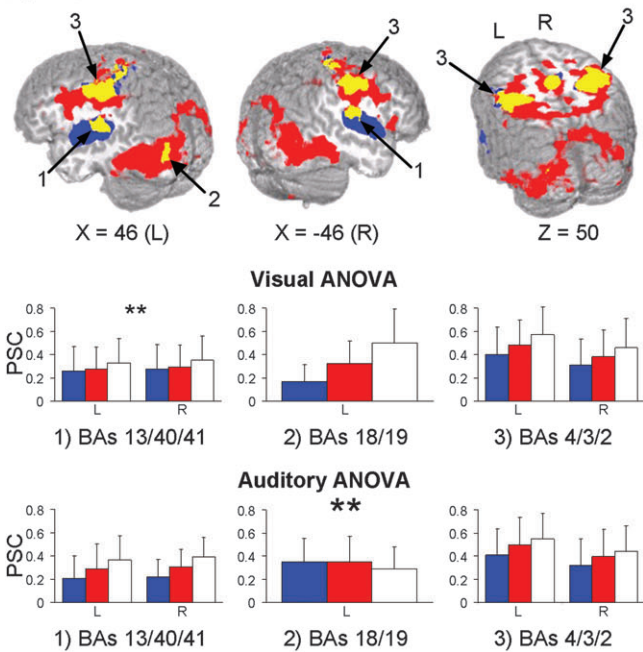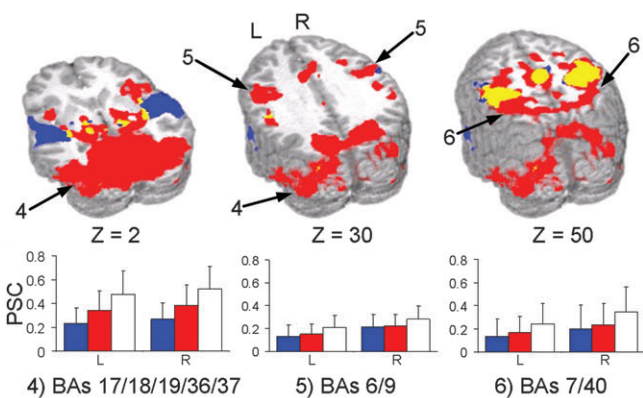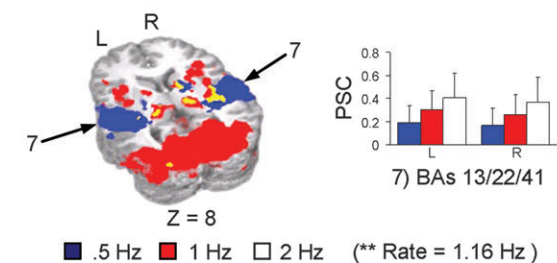
Figure 5. Regions that exhibited significant main effects of rate during both the auditory and visual ANOVAs (panel *A*; yellow coloring) or unique main effects of rate for either the visual (panel *B*; red coloring) or the auditory (panel *C*; blue coloring) ANOVA. In panel (*A*), the PSC is graphed for clusters that were activated both during the visual and auditory ANOVAs that included 1) the bilateral insula, superior temporal gyrus, and inferior parietal lobe (BAs 13/41/40); 2) the left middle occipital gyrus and cuneus (BA 18/19); and 3) the pre- and postcentral gyrus (BAs 4/3/2). Common main effects of rate for both the auditory and the visual ANOVAs outside of the motor circuit were unexpected given that stimulus frequency only varied for the attended stimuli (0.5 Hz, blue bar; 1 Hz, red bar; and 2 Hz, white bar), whereas stimulus frequency was constant (1.16 Hz; graphs denoted with **) for stimuli in the

strong positive visual ARMs (i.e., attend-visual > multimodal and attend-visual > attend-auditory) were observed bilaterally in primary and secondary visual cortex as well as in the ventral and dorsal visual streams.

Other strong positive ARMs during the attend-visual condition included the bilateral rostral and dorsal anterior cingulate gyrus and pre-SMA (Picard and Strick 1996); the inferior and middle frontal gyrus, including the FEFs; and the posterior parietal lobes. The posterior parietal lobes and fronto-oculomotor regions are key regions for both visual and spatial attention (Corbetta and Shulman 2002; Naghavi and Nyberg 2005), although they also modulate auditory attention (Zatorre et al. 2002; Mayer et al. 2006). Greater activation of the anterior cingulate gyrus and pre-SMA during the attend-visual than during the attend-auditory condition was not expected given that conflicting distractors were presented in both conditions. Moreover, these same medial frontal regions did not exhibit differential activation between the multimodal and attend-auditory condition even though conflicting information was present during the attend-auditory condition. The anterior cingulate gyrus and pre-SMA are thought to be crucial for selective attention, although controversy exists about whether these regions modulate attention to relevant stimuli or are involved in ignoring information that conflicts with intended behaviors (Botvinick et al. 1999; Banich et al. 2000; Botvinick et al. 2004; Weissman et al. 2004). Our results suggest that the mere presence of conflicting stimuli is not adequate for pre-SMA and anterior cingulate activity and potentially underscores the diminished relevance of the ignored visual stimuli when rhythmically tapping to an auditory metronome.

The present findings are generally consistent with a study of bimodal temporal discriminations, which reported greater activation when attending to visual than auditory targets in the frontal, parietal, and occipital cortices and greater activation when attending to auditory than visual targets in only the left frontal orbital cortex (Degerman et al. 2007). Because temporal discrimination and paced tapping both require processing of temporal information, the functional results from these studies may have been biased toward the auditory modality, resulting in less effortful processing and decreased ARMs compared with the attend-visual condition. Other studies of bimodal stimulation have reported both more robust ARMs in corresponding sensory cortices during the attend-visual condition compared with subthreshold effects during the attend-auditory conditions (Johnson and Zatorre 2005) as well as ARMs in both visual- and auditory-selective attention conditions (Johnson and Zatorre 2006). Clearly, more research is needed to understand whether the behavioral context (e.g., spatial vs. temporal performance demands) and task parameters (e.g., stimulus frequency; see discussion below) are important for determining the magnitude of ARMs, or if

predominantly ignored modality. Panel (*B*) displays and graphs the PSC in regions associated with a main effect of rate only for the visual ANOVA including 4) an extensive cluster from bilateral primary and secondary visual cortex (BAs 17/18/19/36/37), 5) bilateral precentral and inferior frontal gyrus (BA 6/9), and 6) the posterior parietal lobes (BAs 7/40). Panel (*C*) displays and graphs the PSC in the only areas that showed a main effect of rate in the auditory ANOVA, which was the primary and secondary auditory (BAs 13/22/41) cortex. For all panels, bars represent mean PSC values with 1 SD, and the locations of axial (*z*) and sagittal (*x*) slices are given according to the Talairach atlas.
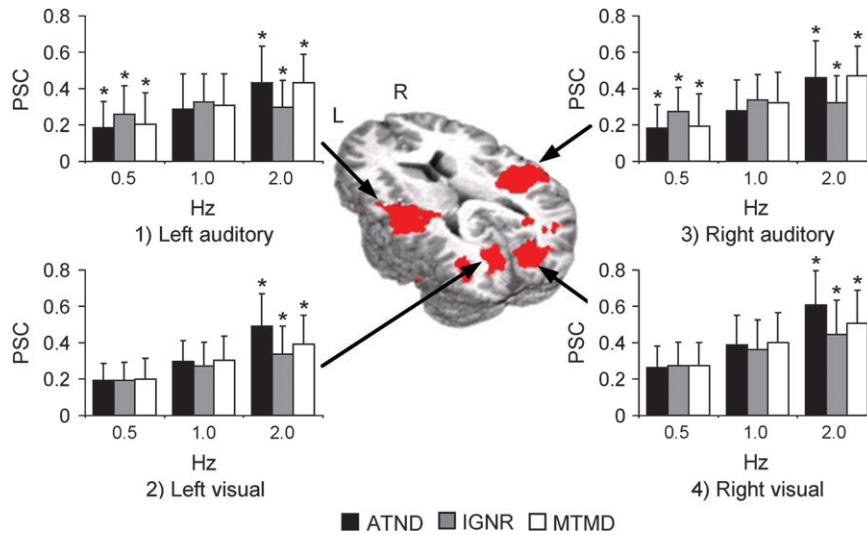
Figure 6. Primary and secondary auditory and visual cortex areas that exhibited a significant condition × rate interaction. A graph of the average PSC is presented for significantly activated voxels ($P < 0.005$) for all the ROIs in both the right and the left hemispheres. Graphs 1 and 2 plot the average PSC for auditory and visual cortices for left hemisphere ROIs at each frequency, and graphs 3 and 4 plot the same data for the right hemisphere. For all selective attention conditions, stimuli were grouped by both stimulus frequency and modality so that the basic stimulus properties were identical in the attend (ATND; black bar) and the ignore (IGNR; gray bar) conditions and were also directly comparable to the multimodal (MTMD; white bar) condition. Asterisks are used to denote significant differences between the conditions at each frequency.

Table 7

Follow-up analyses of the attention condition × rate interaction for auditory and visual cortical ROIs

Modality	Region	Side	0.5 Hz			2 Hz		
			ATND vs. IGNR	ATND vs. MTMD	IGNR vs. MTMD	ATND vs. IGNR	ATND vs. MTMD	IGNR vs. MTMD
Auditory ANOVA Auditory ROIs	Insula (BA 13)	L	IGNR		IGNR	ATND		MTMD
		R	IGNR		IGNR	ATND		MTMD
	Transverse and superior temporal gyrus (BA 22/41/42)	L			IGNR	ATND		MTMD
		R	IGNR		IGNR	ATND		MTMD
	Temporoparietal juncture (BA 39/40)	L				ATND		MTMD
		R	IGNR		IGNR	ATND		MTMD
Visual ANOVA Visual ROIs	Fusiform gyrus (BA 37)	L				ATND	ATND	
		R				ATND	ATND	
	Parahippocampal gyrus (BA 19)	L				ATND	ATND	
		R				ATND	ATND	
	Lingual gyrus (BA 18/19)	L				ATND	ATND	
		R				ATND	ATND	
	Inferior and middle occipital gyrus (BA 19)	L				ATND	ATND	MTMD
		R				ATND	ATND	MTMD
	Middle temporal gyrus (BA 37)	L				ATND		
		R				ATND	ATND	MTMD
	Cuneus and precuneus (BA 31)	L				ATND	ATND	
		R				ATND	ATND	MTMD

Note: Volume of activation was 10.452 mL for all ROIs in left auditory cortex, 12.726 mL for all ROIs in right auditory cortex, 16.916 mL for all ROIs in left visual cortex, and 21.162 mL for all ROIs in right visual cortex. Follow-up paired t-tests compared the attend (ATND), ignore (IGNR), and multimodal (MTMD) conditions; the significance of the tests are designated by the condition in which activation was greater. No significant findings were observed at the 1-Hz frequency.

they are generally more robust for attended visual than auditory stimuli.

Our findings of condition-specific negative ARMs (i.e., Table 5, multimodal > visual-attend) were also consistent with greater interference of ignored auditory signals in the visual-selective attention condition. Specifically, negative ARMs were observed within the bilateral ventral and dorsal visual streams, bilateral pulvinar nucleus and pons, and in the right middle and superior temporal gyrus. In these ROIs, activation was greater in the multimodal than in the auditory-selective attention but not the visual-selective attention condition, suggesting that it was less difficult to ignore visual distractors when tapping to auditory stimuli. In contrast, no regions exhibited *unique*

negative ARMs during multimodal attention when compared with the visual-attend condition. The absence of negative ARMs in the attend-visual condition may also provide indirect evidence for the obligatory recruitment of auditory networks in tasks with a significant temporal component. The automatic recruitment of auditory resources would serve to increase the magnitude of the response during the ignore condition, effectively eliminating negative ARMs. Our results are also consistent with reports that negative ARMs can be observed during bimodal stimulation (Johnson and Zatorre 2005) and are not limited to unisensory stimulation, as previously suggested (Laurienti et al. 2002). However, interference effects may be limited to experiments in which the multimodal stimuli

conflict (Baier et al. 2006) and appear to be more robust for auditory distractors during a rhythmic tapping task.

Finally, of particular interest was the finding that ARMs in primary and secondary visual and auditory areas were highly dependent on the frequency at which stimuli were presented (Table 7, condition × rate interaction). When stimuli were presented at a relatively fast pace of 2 Hz, both auditory and visual cortical areas showed enhancement (attend > ignore) and suppression (multimodal > ignore) effects. Likewise, there was a trend ($P = 0.06$) that the COV tended to be slightly higher at 2 Hz (0.173 ± 0.004) compared with both 1 Hz (0.144 ± 0.005) and 0.5 Hz (0.146 ± 0.014), suggesting that processing efficiency was somewhat reduced when the attended stimulus rate was the fastest. Although this finding might suggest that it was more difficult to ignore rapidly presented distractors, it more likely relates to processing efficiency of the attended stimulus because the COV tends to be larger for stimulus rates below 1 s (Gibbon et al. 1997). This latter interpretation is more consistent with the functional data, which suggested that at the 2-Hz stimulus rate, participants effectively allocated their attention to the target stimulus in both sensory modalities (i.e., robust ARMs in primary and secondary visual and auditory areas at 2 Hz). These findings contrasted with our behavioral and functional results at 0.5 Hz, both of which indicated that auditory distractors were particularly more salient when the attended stimulus rate was the slowest. Processing efficiency (COV) was the worst in the attend-visual compared with both the attend-auditory and the multimodal conditions at 0.5 Hz, and *ignoring* the auditory metronome (i.e., attend-visual condition) produced an *enhanced* neuronal response within the bilateral auditory cortex compared with the other 2 attention conditions (Table 7, auditory ANOVA). This finding was unexpected and suggests that ignored auditory tones were more salient than attended-auditory stimuli but only when the attended stimulus rate was relatively slow paced. This finding also corroborates our previous suggestion that the processing of auditory stimuli by auditory cortical areas may be obligatory during rhythmic tapping tasks, which then biases cognitions that are reliant on temporal processes toward the information contained in the auditory modality.

The interpretation of our findings could potentially be limited by several factors. First, we employed a conventional rather than a continuous (Seifritz et al. 2006) or sparse (Hall et al. 1999) echo-planar imaging (EPI) sequence, which may have increased interference and/or limited activation within auditory cortical areas through saturation effects (Bandettini et al. 1998). However, our behavioral results were generally similar to previous studies conducted outside of the scanner environment (Repp and Penel 2004; Kato and Konishi 2006), suggesting that this is not a compelling explanation. In addition, our findings of less robust activation in auditory cortical areas are generally consistent with a previous study that employed a sparse sampling procedure (Johnson and Zatorre 2005). A second potential limitation was our choice of a more complex visual stimulus (i.e., reversing checkerboard) compared with our auditory stimulus (pure tone), which may have produced additional functional activation within visual areas. However, behavioral studies using simple flashing lights that equate for stimulus saliency have also reported a dominance for paced tapping to auditory relative to visual metronomes (Repp and Penel 2004; Kato and Konishi 2006). More importantly, the basic properties of the visual stimuli were equated across conditions in our within-subjects design, suggesting that any functional differences in visual cortex

were likely the result of attentional rather than sensory processes. As a final comment, synchronized tapping is more variable to a visual than to an auditory metronome; as such, our findings may reflect the combined effects of more variable timing of the attended visual signals and greater interference from distracting auditory signals (Repp and Penel 2004; Kato and Konishi 2006). To delineate the unique contributions of each of these processes, future FMRI experiments will need to examine synchronous tapping to unimodal metronomes both with and without cross-modal distractors.

In summary, our results revealed several potential neural mechanisms for the dominance of auditory signals over visual stimuli when selectively attending to temporal information in a multimodal context. When stimulus rate was equated, robust ARMs were evident within striate and extrastriate cortices, as well as within the frontal and parietal lobes during the attend-visual condition (Table 5). In contrast, there was no evidence of ARMs during the attend-auditory condition. However, in traditional unisensory cortical areas, the ARMs depended in part on the rate of stimulus information. Evidence for the obligatory recruitment of auditory cortex was found under conditions in which ignored auditory information was particularly detrimental to task performance. Specifically, there was greater activation in the auditory cortex during the ignore-auditory than during the attend-auditory condition (see Fig. 6) at 0.5 Hz, where there was a larger window of time for conflicting information to intrude upon selective attention processes. In contrast, positive and negative ARMs in primary and secondary cortical areas were present for both the attend-visual and the attend-auditory conditions but only at 2 Hz. Here, the tendency for lower processing efficiency at 2 Hz (i.e., higher COV) was likely due to selectively attending and responding to a fast-paced target stimulus. This interpretation was consistent with the pattern of ARMs (Table 7, Attend > Ignore and Multimodal > Ignore), which indicated that participants effectively allocated attention to the target stimulus, irrespective of sensory modality. Our results and those of others support the recruitment of a more extensive cortical network during rhythmic tapping to visual cues with, or without (Jantzen et al. 2005), auditory distractors. Collectively, these results indicate that the processing of auditory information is more efficient (i.e., requires coordination of a smaller network), less effortful (i.e., reduced variability and fewer ARMs), and obligatory (i.e., automatic recruitment of auditory cortex during both attend and ignore conditions) in behavioral contexts that have a significant temporal component.

Address correspondence to Andrew Mayer, PhD, The Mind Research Network, Pete and Nancy Domenici Hall, 1101 Yale Boulevard NE, Albuquerque, NM 87131, USA. Email: amayer@mrn.org.

References

Alho K. 1992. Selective attention in auditory processing as reflected by event related brain potentials. Psychophysiology. 29:247–263.

Alho K, Medvedev SV, Pakhomov SV, Roudas MS, Tervaniemi M, Reinikainen K, Zeffiro T, Naatanen R. 1999. Selective tuning of the left and right auditory cortices during spatially directed attention. Brain Res Cogn Brain Res. 7:335-341.

Arnell KM, Duncan J. 2002. Separate and shared sources of dual-task cost in stimulus identification and response selection. Cognit Psychol. 44:105-147.

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.

Bandettini PA, Jesmanowicz A, Van Kylen J, Birn RM, Hyde JS. 1998. Functional MRI of brain activation induced by scanner acoustic noise. Magn Reson Med. 39:410-416.

Banich MT, Milham MP, Atchley RA, Cohen NJ, Webb A, Wszalek T, Kramer AF, Liang Z, Barad V, Gullett D, et al. 2000. Prefrontal regions play a predominant role in imposing an attentional 'set': evidence from fMRI. Brain Res Cogn Brain Res. 10:1-9.

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

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

Binder JR, Rao SM, Hammeke TA, Frost JA, Bandettini PA, Hyde JS. 1994. Effects of stimulus rate on signal response during functional magnetic resonance imaging of auditory cortex. Cogn Brain Res. 2:31-38.

Botvinick M, Nystrom LE, Fissell K, Carter CS, Cohen JD. 1999. Conflict monitoring versus selection-for-action in anterior cingulate cortex. Nature. 402:179-181.

Botvinick MM, Cohen JD, Carter CS. 2004. Conflict monitoring and anterior cingulate cortex: an update. Trends Cogn Sci. 8:539-546.

Boynton GM, Engel SA, Glover GH, Heeger DJ. 1996. Linear systems analysis of functional magnetic resonance imaging in human V1. J Neurosci. 16:4207-4221.

Burock MA, Buckner RL, Woldorff MG, Rosen BR, Dale AM. 1998. Randomized event-related experimental designs allow for extremely rapid presentation rates using functional MRI. NeuroReport. 9:3735-3739.

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 USA. 102:18751-18756.

Calvert GA, Hansen PC, Iversen SD, Brammer MJ. 2001. Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the BOLD effect. Neuroimage. 14:427-438.

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

Ciaramitaro VM, Buracas GT, Boynton GM. 2007. Spatial and cross-modal attention alter responses to unattended sensory information in early visual and auditory human cortex. J Neurophysiol. 98:2399-2413.

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

Cowan N, Barron A. 1987. Cross-modal, auditory-visual Stroop interference and possible implications for speech memory. Percept Psychophys. 41:393-401.

Cox RW. 1996. AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. Computers and Biomedical research. 29:162-173.

Degerman A, Rinne T, Pekkola J, Autti T, Jaaskelainen IP, Sams M, Alho K. 2007. Human brain activity associated with audiovisual perception and attention. Neuroimage. 34:1683-1691.

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

Dhamala M, Assisi CG, Jirsa VK, Steinberg FL, Kelso JA. 2007. Multisensory integration for timing engages different brain networks. Neuroimage. 34:764-773.

Driver J. 2001. A selective review of selective attention research from the past century. Br J Psychol. 92:53-78.

Driver J, Baylis G. 1993. Cross-modal negative priming and interference in selective attention. Bull Psychon Soc. 31:45-48.

Eimer M, van Velzen J, Driver J. 2002. Cross-modal interactions between audition, touch, and vision in endogenous spatial attention: eRP evidence on preparatory states and sensory modulations. J Cogn Neurosci. 14:254-271.

Esposito F, Bertolino A, Scarabino T, Latorre V, Blasi G, Popolizio T, Tedeschi G, Cirillo S, Goebel R, Di Salle F. 2006. Independent component model of the default-mode brain function: assessing the impact of active thinking. Brain Res Bull. 70:263-269.

Fink GR, Marshall JC, Shah NJ, Weiss PH, Halligan PW, Grosse-Ruyken M, Ziemons K, Zilles K, Freund HJ. 2000. Line bisection judgments implicate right parietal cortex and cerebellum as assessed by fMRI. Neurology. 54:1324-1331.

Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC. 1995. Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. Magn Reson Med. 33:636-647.

Gibbon J, Malapani C, Dale CL, Gallistel C. 1997. Toward a neurobiology of temporal cognition: advances and challenges. Curr Opin Neurobiol. 7:170-184.

Grady CL, Van Meter JW, Maisog JM, Pietrini P, Krasuski J, Rauschecker JP. 1997. Attention-related modulation of activity in primary and secondary auditory cortex. NeuroReport. 8:2511-2516.

Greicius MD, Krasnow B, Reiss AL, Menon V. 2003. Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. Proc Natl Acad Sci USA. 100:253-258.

Hall DA, Haggard MP, Akeroyd MA, Palmer AR, Summerfield AQ, Elliott MR, Gurney EM, Bowtell RW. 1999. Sparse temporal sampling in auditory fMRI. Hum Brain Mapp. 7:213-223.

Harrington DL, Boyd LA, Mayer AR, Sheltraw DM, Lee RR, Huang M, Rao SM. 2004. Neural representation of interval encoding and decision making. Brain Res Cogn Brain Res. 21:193-205.

Hubel DH, Henson CO, Rupert A, Galambos R. 1959. "Attention" Units in the Auditory Cortex. Psychobiology. 129:1279-1280.

Jancke L, Loose R, Lutz K, Specht K, Shah NJ. 2000. Cortical activations during paced finger-tapping applying visual and auditory pacing stimuli. Brain Res Cogn Brain Res. 10:51-66.

Jancke L, Mirzazade S, Shah NJ. 1999. Attention modulates activity in the primary and the secondary auditory cortex: a functional magnetic resonance imaging study in human subjects. Neurosci Lett. 266:125-128.

Jantzen KJ, Steinberg FL, Kelso JA. 2005. Functional MRI reveals the existence of modality and coordination-dependent timing networks. Neuroimage. 25:1031-1042.

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

Johnson JA, Zatorre RJ. 2006. Neural substrates for dividing and focusing attention between simultaneous auditory and visual events. Neuroimage. 31:1673-1681.

Kato M, Konishi Y. 2006. Auditory dominance in the error correction process: a synchronized tapping study. Brain Res. 1084:115-122.

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

Laurienti PJ, Burdette JH, Wallace MT, Yen YF, Field AS, Stein BE. 2002. Deactivation of sensory-specific cortex by cross-modal stimuli. J Cogn Neurosci. 14:420-429.

Lewis JW, Beauchamp MS, DeYoe EA. 2000. A comparison of visual and auditory motion processing in human cerebral cortex. Cereb Cortex. 10(9):873-888.

Liu T, Slotnick SD, Serences JT, Yantis S. 2003. Cortical mechanisms of feature-based attentional control. Cereb Cortex. 13:1334-1343.

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

Mayer AR, Harrington D, Adair JC, Lee R. 2006. The neural networks underlying endogenous auditory covert orienting and reorienting. Neuroimage. 30:938-949.

Mayer AR, Kosson DS. 2004. The effects of auditory and visual linguistic distracters on target localization. Neuropsychology. 15:248-257.

Morein-Zamir S, Soto-Faraco S, Kingstone A. 2003. Auditory capture of vision: examining temporal ventriloquism. Brain Res Cogn Brain Res. 17:154-163.

Naghavi HR, Eriksson J, Larsson A, Nyberg L. 2007. The claustrum/insula region integrates conceptually related sounds and pictures. Neurosci Lett. 422:77–80.

Naghavi HR, Nyberg L. 2005. Common fronto-parietal activity in attention, memory, and consciousness: shared demands on integration? Conscious Cogn. 14:390–425.

O'Craven KM, Rosen BR, Kwong KK, Treisman A, Savoy RL. 1997. Voluntary attention modulates fMRI activity in human MT-MST. Neuron. 18:591–598.

Ozus B, Liu HL, Chen L, Iyer MB, Fox PT, Gao JH. 2001. Rate dependence of human visual cortical response due to brief stimulation: an event-related fMRI study. Magn Reson Imaging. 19:21–25.

Paus T. 1996. Location and function of the human frontal eye-field: a selective review. Cognitive Neuropsychology. 34:475–483.

Petrides M, Pandya DN. 1999. Dorsolateral prefrontal cortex: comparative cytoarchitectonic analysis in the human and the macaque brain and corticocortical connection patterns. Eur J Neurosci. 11:1011–1036.

Picard N, Strick PL. 1996. Motor areas of the medial wall: a review of their location and functional activation. Cereb Cortex. 6:342–353.

Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL. 2001. A default mode of brain function. Proc Natl Acad Sci USA. 98:676–682.

Rao SM, Bandettini PA, Binder JR, Bobholz JA, Hammeke TA, Stein EA, Hyde JS. 1996. Relationship between finger movement rate and functional magnetic resonance signal change in human primary motor cortex. J Cereb Blood Flow Metab. 16:1250–1254.

Rao SM, Mayer AR, Harrington DL. 2001. The evolution of brain activation during temporal processing. Nat Neurosci. 4:317–323.

Repp BH, Penel A. 2004. Rhythmic movement is attracted more strongly to auditory than to visual rhythms. Psychol Res. 68:252–270.

Reynolds JH, Chelazzi L, Desimone R. 1999. Competitive mechanisms subserve attention in macaque areas V2 and V4. J Neurosci. 19:1736–1753.

Riecker A, Wildgruber D, Mathiak K, Grodd W, Ackermann H. 2003. Parametric analysis of rate-dependent hemodynamic response functions of cortical and subcortical brain structures during auditorily cued finger tapping: a fMRI study. Neuroimage. 18:731–739.

Rif J, Hari R, Hamalainen MS, Sams M. 1991. Auditory attention affects two different areas in the human supratemporal cortex. Electroencephalogr Clin Neurophysiol. 79:464–472.

Rinne T, Pekkola J, Degerman A, Autti T, Jaaskelainen IP, Sams M, Alho K. 2005. Modulation of auditory cortex activation by sound presentation rate and attention. Hum Brain Mapp. 26:94–99.

Seifritz E, Di Salle F, Esposito F, Herdener M, Neuhoff JG, Scheffler K. 2006. Enhancing BOLD response in the auditory system by neurophysiologically tuned fMRI sequence. Neuroimage. 29:1013–1022.

Shams L, Kamitani Y, Shimojo S. 2002. Visual illusion induced by sound. Brain Res Cogn Brain Res. 14:147–152.

Shomstein S, Yantis S. 2004. Control of attention shifts between vision and audition in human cortex. J Neurosci. 24:10702–10706.

Shulman GL, Tansy AP, Kincade M, Petersen SE, McAvoy MP, Corbetta M. 2002. Reactivation of networks involved in preparatory states. Cereb Cortex. 12:590–600.

Spence C, Pavani F, Driver J. 2004. Spatial constraints on visual-tactile cross-modal distractor congruency effects. Cogn Affect Behav Neurosci. 4:148–169.

Spencer RM, Ivry RB. 2005. Comparison of patients with Parkinson's disease or cerebellar lesions in the production of periodic movements involving event-based or emergent timing. Brain Cogn. 58:84–93.

Talairach J, Tournoux P. 1988. Co-planar stereotaxic atlas of the human brain. New York: Thieme.

Talsma D, Kok A, Ridderinkhof KR. 2006. Selective attention to spatial and non-spatial visual stimuli is affected differentially by age: effects on event-related brain potentials and performance data. Int J Psychophysiol. 62:249–261.

Van Vleet TM, Robertson LC. 2006. Cross-modal interactions in time and space: auditory influence on visual attention in hemispatial neglect. J Cogn Neurosci. 18:1368–1379.

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

Witten IB, Knudsen EI. 2005. Why seeing is believing: merging auditory and visual worlds. Neuron. 48:489–496.

Woldorff MG, Gallen CC, Hampson SA, Hillyard SA, Pantev C, Sobel D, Bloom FE. 1993. Modulation of early sensory processing in human auditory cortex during auditory selective attention. Proc Natl Acad Sci USA. 90:8722–8726.

Woodruff PW, Benson RR, Bandettini PA, Kwong KK, Howard RJ, Talavage T, Belliveau J, Rosen BR. 1996. Modulation of auditory and visual cortex by selective attention is modality-dependent. NeuroReport. 7:1909–1913.

Woods DL, Alho K, Algazi A. 1992. Intermodal selective attention. I. Effects on event-related potentials to lateralized auditory and visual stimuli. Electroencephalogr Clin Neurophysiol. 82:341–355.

Zatorre RJ, Bouffard M, Ahad P, Belin P. 2002. Where is 'where' in the human auditory cortex? Nat Neurosci. 5:905–909.

Zelano C, Bensafi M, Porter J, Mainland J, Johnson B, Bremner E, Telles C, Khan R, Sobel N. 2005. Attentional modulation in human primary olfactory cortex. Nat Neurosci. 8:114–120.