Effect of Attention on Early Cortical Processes Associated with the Sound-induced Extra Flash Illusion

Jyoti Mishra[1], Antigona Martínez[1,2], and Steven A. Hillyard[1]

Abstract

■ When a single flash of light is presented interposed between two brief auditory stimuli separated by 60–100 msec, subjects typically report perceiving two flashes [Shams, L., Kamitani, Y., & Shimojo, S. Visual illusion induced by sound. *Brain Research, Cognitive Brain Research, 14,* 147–152, 2002; Shams, L., Kamitani, Y., & Shimojo, S. Illusions. What you see is what you hear. *Nature, 408,* 788, 2000]. Using ERP recordings, we previously found that perception of the illusory extra flash was accompanied by a rapid dynamic interplay between auditory and visual cortical areas that was triggered by the second sound [Mishra, J., Martínez, A., Sejnowski, T. J., & Hillyard, S. A. Early cross-modal interactions in auditory and visual cortex underlie a sound-induced visual illusion. *Journal of Neuroscience, 27,* 4120–4131, 2007]. In the current study, we investigated the effect of attention on the ERP components associated with the illusory extra flash in 15 individuals who perceived this cross-modal illusion frequently. All early ERP components in the cross-modal difference wave asso-ciated with the extra flash illusion were significantly enhanced by selective spatial attention. The earliest attention-related modulation was an amplitude increase of the positive-going PD110/PD120 component, which was previously shown to be correlated with an individual's propensity to perceive the illusory second flash [Mishra, J., Martínez, A., Sejnowski, T. J., & Hillyard, S. A. Early cross-modal interactions in auditory and visual cortex underlie a sound-induced visual illusion. *Journal of Neuroscience, 27,* 4120–4131, 2007]. The polarity of the early PD110/PD120 component did not differ as a function of the visual field (upper vs. lower) of stimulus presentation. This, along with the source localization of the component, suggested that its principal generator lies in extrastriate visual cortex. These results indicate that neural processes previously shown to be associated with the extra flash illusion can be modulated by attention, and thus are not the result of a wholly automatic cross-modal integration process. ■

INTRODUCTION

Events in the natural world are often multimodal, and intersensory interactions in the brain are critical to the generation of coherent percepts and the control of behavior (reviewed in Driver & Noesselt, 2008; Amedi, Von Kriegstein, Van Atteveldt, Beauchamp, & Naumer, 2005; Macaluso & Driver, 2005; Calvert, 2001; Stein & Meredith, 1993). Within the audiovisual domain, numerous behavioral studies have shown that simultaneous auditory and visual inputs interact such that visual perception can be altered by audition and vice versa. For example, the perceived location of sounds is robustly altered by concurrent visual stimuli at a nearby location, a phenomenon known as the ventriloquist illusion (Bonath et al., 2007; Vroomen & De Gelder, 2004; Hairston et al., 2003; Bertelson, 1999; Pick, Warren, & Hay, 1969). Conversely, the concurrent presentation of sounds can strikingly alter visual perception (McDonald, Teder-Sälejärvi, Di Russo, & Hillyard, 2003, 2005; Recanzone, 2003; Fendrich & Corballis, 2001; Sekuler, Sekuler, & Lau, 1997; Stein, London, Wilkinson, & Price, 1996). One of the most striking visual illusions induced by auditory stimulation is that introduced by Shams, Kamitani, and Shimojo (2000, 2002), wherein a single brief flash presented interposed between two pulsed sounds separated by 60–100 msec generates the percept of two distinct flashes, of which the second is illusory.

The neural basis of the illusory extra flash has been investigated in several physiological studies (Mishra, Martínez, Sejnowski, & Hillyard, 2007; Watkins, Shams, Tanaka, Haynes, & Rees, 2006; Shams, Iwaki, Chawla, & Bhattacharya, 2005; Arden, Wolf, & Messiter, 2003; Shams, Kamitani, Thompson, & Shimojo, 2001). Using ERP recordings, Mishra et al. (2007) found that the cross-modal interactions underlying the illusory flash phenomenon had a complex but distinct neural signature. An early positive component peaking at 120 msec after the onset of the first sound was identified in the cross-modal interaction waveforms associated with the illusion. This PD120 component was found to originate within extrastriate visual cortex, and its amplitude in individual subjects was predictive of the frequency with which they perceived the extra flash illusion. Indeed, the PD120 was completely absent in subjects who did not perceive the illusion. Interestingly, however, this component did not vary within individual subjects on trials when the illusion was seen versus not seen (Mishra et al., 2007). Accordingly, we suggested that

[1]University of California, San Diego, La Jolla, CA, [2]Nathan S. Kline Institute for Psychiatric Research, Orangeburg, NY

the PD120 reflects neural activity that is necessary but not sufficient to produce the illusion and characterizes individuals who are disposed to perceive the illusion. The objective of the present study was to investigate two factors that may modulate this early cross-modal interaction component as well as other neural processes associated with the sound-induced extra flash illusion. These factors were selective spatial attention and location of the stimuli within the visual field.

Attention is known to amplify and enhance the processing of external stimuli to which it is allocated. Perceptual thresholds are lowered, reaction times are speeded, and detection accuracy increases as a consequence of attention (e.g., Carrasco, 2006; Luck et al., 1994; Posner & Petersen, 1990). These behavioral effects have been linked with increases in the neural response to attended stimuli relative to unattended stimuli, ranging from increased firing rates at the single neuron level (Reynolds & Chelazzi, 2004) and enhancement of sensory ERPs (Hopfinger, Luck, & Hillyard, 2004; Hillyard & Anllo-Vento, 1998) to increased blood flow within sensory cortical areas (Kastner & Pinsk, 2004; Hopfinger, Buonocore, & Mangun, 2000; Kastner & Ungerleider, 2000). The role of attention in the context of cross-modal stimulation has also been studied extensively (Talsma, Kok, Slagter, & Cipriani, 2008; Molholm, Martínez, Shpaner, & Foxe, 2007; Talsma, Doty, & Woldorff, 2007; Busse, Roberts, Crist, Weissman, & Woldorff, 2005; Talsma & Woldorff, 2005; Eimer, van Velzen, & Driver, 2004; McDonald et al., 2003; Eimer & Driver, 2001; McDonald, Teder-Sälejärvi, Heraldez, & Hillyard, 2001; Eimer, 1999; Teder-Sälejärvi, Münte, Sperlich, & Hillyard, 1999; Eimer & Schroger, 1998; Hillyard, Simpson, Woods, Van Voorhis, & Münte, 1984). A general finding from these studies is that attention may enhance cross-modal interactions as early as 100 msec after stimulus onset.

The role of attention in modulating the neural interactions that underlie audiovisual illusions has been little investigated. Busse et al. (2005) found that attention enhanced cross-modal interactions in an experimental paradigm that simulated the ventriloquist illusion, but illusory perception was not measured as part of the study. The present study was designed to characterize the effect of attention on the neural interactions associated with the audiovisual extra flash illusion discovered by Shams et al. (2000). In particular, the aim was to find out whether the ERP components previously shown to be correlated with the perceptual illusion (Mishra et al., 2007) result from an automatic cross-modal interaction that is uninfluenced by attention. In the present design, stimuli were presented at two locations, one in the upper (UVF) and one in the lower visual field (LVF), while subjects focused attention on only one of the locations at a time. The UVF and LVF locations were chosen in order to help delineate the possible role of primary visual (striate) cortex in the generation of early cross-modal interaction components. It is well known that the earliest component of the visual-evoked potential, the so-called C1 elicited at 50–90 msec, inverts in polarity for stimuli presented in the upper versus the lower field, which supports the general consensus that the C1 originates in large part from striate cortex (Di Russo, Martínez, & Hillyard, 2003; Di Russo, Martínez, Sereno, Pitzalis, & Hillyard, 2002; Martínez et al., 2001; Clark, Fan, & Hillyard, 1995; Jeffreys, 1968). Following this logic, ERP components associated with the extra flash illusion that were primarily generated within striate cortex would also exhibit such polarity inversion.

METHODS

Task and Stimuli

Fifteen right-handed healthy adults (8 women, mean age = 21.4 years) participated in the study after giving written informed consent as approved by the University of California, San Diego Human Research Protections Program. Each participant had normal or corrected-to-normal vision and normal hearing. All subjects chosen for the experiment perceived the sound-induced extra flash illusion on 50% or more of the trials as tested in a short (5-min) screening session prior to the main experiment. Ten of the 15 participants were selected from the pool of subjects who perceived the extra flash illusion frequently during the experiment reported in Mishra et al. (2007).

The experiment was conducted in a sound-attenuated chamber having a background sound level of 32 dB and a background luminance of 2 cd/m^2. Subjects maintained fixation on a central cross positioned at a viewing distance of 120 cm. Auditory (A) and visual (V) stimuli were delivered from paired speakers and red LED, one pair in the UVF and another in the LVF. The speaker/LED pairs were positioned at 20° eccentricity to the left of fixation and at 30° polar angle above and below the horizontal meridian (Figure 1A). The eccentricity of the stimuli was the same as that used previously (Mishra et al., 2007). Large polar angles were chosen above and below the horizontal meridian so that the upper and lower auditory stimuli would be heard as spatially distinct. These particular polar angles were also chosen to investigate the possible involvement of neural generators in striate cortex, which reportedly produce maximal amplitudes for stimuli near 30° of stimulus elevation in the upper and lower fields (Clark et al., 1995). Each visual stimulus was a 5-msec 75-cd/m^2 flash, and each auditory stimulus was a 10-msec 76-dB noise burst.

Six different stimulus combinations were presented one at a time to either the UVF or the LVF in random order (Figure 1B). Both the order of the combinations and field of presentation were randomized within each block of trials. The combinations included unimodal auditory stimuli, occurring in pairs (A_1A_2) and unimodal visual stimuli occurring singly (V_1) or in pairs (V_1V_2). Bimodal stimulus combinations included $A_1V_1A_2$ and $A_1A_2V_1$. In this terminology, suffixes 1 or 2 denote the
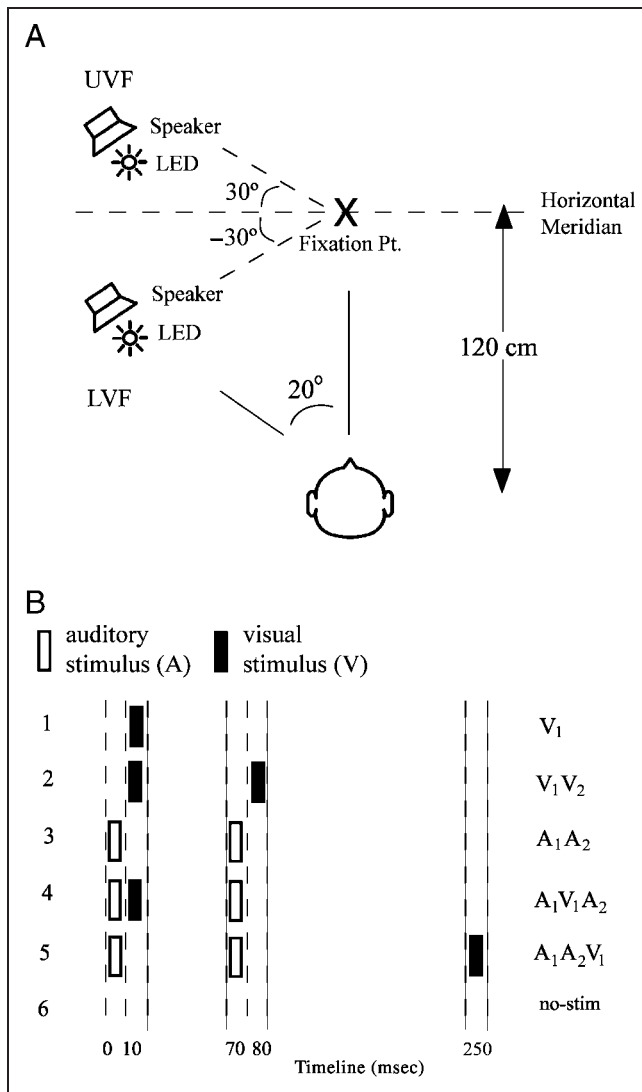
Figure 1. Overview of experimental design. (A) Schematic diagram of experimental setup. (B) Listing of the six different stimulus configurations, which were presented one at a time to either the UVF or LVF. Both the order of stimuli and the field of stimulation (upper or lower) were randomized within each block. Abscissa indicates times of occurrence of auditory (open bars) and visual (solid bars) stimuli. Auditory (A) and visual (V) stimuli are labeled 1 or 2 to designate their first or second occurrence in each configuration.

first or second occurrence of the auditory or visual component of each stimulus combination. Other bimodal combinations such as A_1V_1, $A_1V_1A_2V_2$ or $A_1V_1V_2$, which were previously studied by Mishra et al. (2007), could not be included in the present study as they would have significantly prolonged the experiment and produced subject fatigue and deteriorated performance. Finally, blank or no-stimulus (no-stim) trial ERPs were recorded over the same epochs as for actual stimuli but with no stimulus presented. The timing of the A and V components for each stimulus combination is shown in Figure 1B. The SOA between the two stimuli in the A_1A_2 and V_1V_2 pairs was 70 msec in every stimulus combination that included them. The SOA between A_1 and V_1 was

10 msec for $A_1V_1A_2$, and V_1 followed A_1 by 250 msec for $A_1A_2V_1$. The $A_1A_2V_1$ stimulus with the delayed flash did not produce an illusory second flash, and thus, served as a stimulus-matched behavioral control for the $A_1V_1A_2$ test stimulus that did produce the illusion, thereby ensuring that reports of the visual illusion were not based on simply counting the number of sounds.

Stimuli were presented in 16 blocks with 24 trials of each of the six stimulus combinations delivered in a randomized sequence (12 to the UVF and 12 to the LVF) on each block. All configurations occurred with equal probability and were presented at irregular intervals of 800–1200 msec. Within each block, a 5-sec break period of no stimulation was given every 30 sec. Subjects were instructed to attend to the visual stimuli in either UVF or LVF on each block and report the number of flashes perceived (1 or 2) after each stimulus combination occurring in the attended field that contained one or two flashes. Subjects were instructed to ignore all stimuli in the unattended visual field, and no responses were required to the unimodal auditory stimuli. The order of attended blocks was counterbalanced across subjects. Overall, 192 trials were recorded for each attended as well as unattended stimulus combination in each visual field.

Electrophysiological (ERP) Recordings

The EEG was recorded from 62 electrode sites using a modified 10–10 system montage (Teder-Sälejärvi, Di Russo, McDonald, & Hillyard, 2005). Horizontal and vertical electrooculograms (EOGs) were recorded by means of electrodes at the left and right external canthi and an electrode below the left eye, respectively. The importance of fixation was emphasized to subjects, and the experimenter continually monitored the EOG and verified fixation in all blocks. The large eccentricities and elevations of the auditory and visual stimuli also ensured that subjects did not deviate their gaze toward the stimulus positions, as such large ocular deviations would be easily detected by the EOG. All electrodes were referenced to the right mastoid electrode during recording. Electrode impedances were kept below 5 kΩ.

All signals were amplified with a gain of 10,000 and a band pass of 0.1–80 Hz (−12 dB/octave; 3 dB attenuation) and were digitized at 250 Hz. Automated artifact rejection was performed prior to averaging to discard trials with eye movements, blinks, or amplifier blocking. Signals were averaged in 500-msec epochs with a 100-msec prestimulus baseline. The averages were digitally low-pass filtered with a Gaussian finite impulse function (3 dB attenuation at 46 Hz) to remove high-frequency noise produced by muscle activity and external electrical sources. The filtered averages were digitally re-referenced to the average of the left and right mastoids.

The three-dimensional coordinates of each electrode and of three fiducial landmarks (the left and right preauricular points and the nasion) were determined by means of

a Polhemus spatial digitizer (Polhemus, Colchester, VT). The mean Cartesian coordinates for each site were averaged across all subjects and used for topographic mapping and source localization procedures.

Cross-modal interaction difference waves were calculated for the $A_1V_1A_2$ stimulus that generated the percept of the illusory extra flash by subtracting the ERPs elicited by the individual unimodal components of the bimodal configuration from the ERP elicited by the total configuration (Calvert, Stein, & Spence, 2004; Giard & Peronnet, 1999; Stein & Meredith, 1993). This difference wave was termed Ill_Diff as it reflected the auditory–visual interactions associated with the illusory second flash; the Ill_Diff was separately calculated for stimuli in the UVF and LVF and for both the attended and unattended conditions, as follows:

$$Ill\_Diff = [(A_1V_1A_2) + no\text{-}stim] - [A_1A_2 + V_1]$$

Four such difference waves were calculated: for attended stimuli in the UVF (Ill_Diff$_{ATT-UVF}$), for unattended stimuli in the UVF (Ill_Diff$_{UNATT-UVF}$), for attended stimuli in the LVF (Ill_Diff$_{ATT-LVF}$), and for unattended stimuli in the LVF (Ill_Diff$_{UNATT-LVF}$).

The blank or no-stimulus (no-stim) trials were included in the calculation of these cross-modal difference waves to balance any prestimulus activity (such as anticipatory contingent negative variation [CNV]) that may extend into the poststimulus period on all trials. If the no-stim trials were not included, such activity would be added once but subtracted twice in the difference wave, possibly introducing an early deflection that could be mistaken for a true cross-modal interaction (Gondan & Roder, 2006; Talsma & Woldorff, 2005; Teder-Sälejärvi, McDonald, Di Russo, & Hillyard, 2002).

As many authors have noted (e.g., Molholm et al., 2002; Giard & Peronnet, 1999), any departure from linear summation of the concurrent auditory- and visual-evoked activity will result in a deflection in the cross-modal interaction difference wave. In principle, such interaction effects could arise from (i) cross-modal modulation of neural activity that is normally evoked unimodally, (ii) recruitment of a new neural population not activated unimodally, or (iii) a shift in the latency of unimodally evoked activity. It is often difficult to specify which type of cross-modal interaction is reflected in such difference wave components, but they can give an indication of the timing and localization of the brain areas where the interactions are taking place.

Data Analysis

ERP components observed in each Ill_Diff difference wave were first tested for significance with respect to the pre-stimulus baseline by t tests over all subjects ($n = 15$). For all analyses, difference wave components were quantified as mean amplitudes within specific latency windows around the peak for each identified positive difference (PD) or negative difference (ND) component with respect to the mean voltage of a 100-msec prestimulus baseline. Components in the Ill_Diff$_{UVF}$ difference wave (both attended and unattended) were measured at 112–132 msec (PD120), 164–184 msec (PD180), and 240–260 msec (ND250). For the Ill_Diff$_{LVF}$ difference wave (again, both attended and unattended), components were measured at 104–124 msec (PD110), 164–184 msec (PD180), and 228–248 msec (ND240). Each of these components was measured as the mean voltage over a specific cluster of electrodes where its amplitude was maximal. The PD120 and PD110 components for stimuli in the UVF and LVF, respectively, were measured over 15 occipital electrode sites (6 in each hemisphere and 3 over midline); PD180 amplitudes (for both UVF and LVF) were measured over fronto-central electrode clusters (8 in each hemisphere and 4 over midline); and the ND250/ND240 components were measured over a similar set of central electrodes (8 in each hemisphere).

Scalp distributions of ERP components in the Ill_Diff difference waves were compared after normalizing their amplitudes prior to ANOVA according to the method described by McCarthy and Wood (1985). For posteriorly distributed components (PD120/PD110), comparisons were made over 18 occipital electrode sites (7 in each hemisphere and 4 over midline). For the other components (PD180 and ND250/ND240), comparisons were made over 38 electrodes spanning frontal, central, parietal, and occipital sites (15 in each hemisphere and 8 over midline). Differences in scalp distribution were reflected in significant Stimulus condition (ATT vs. UNATT or UVF vs. LVF) by Electrode interactions.

Modeling of ERP Sources

Source localization was carried out to estimate the intracranial generators of components in the grand-averaged difference waves within the same time intervals as those used for statistical testing. Source locations were estimated by dipole modeling using BESA (Brain Electrical Source Analysis 2000, version 5). The BESA algorithm estimates the location and the orientation of multiple equivalent dipolar sources by calculating the scalp distribution that would be obtained for a given dipole model (forward solution) and comparing it to the actual scalp-recorded ERP distribution (Scherg, 1990). The algorithm interactively adjusts (fits) the location and orientation of the dipole sources in order to minimize the relative variance between the model and the observed spatio-temporal ERP distribution. This analysis used the three-dimensional coordinates of each electrode site as recorded by a spatial digitizer. Only one symmetrical pair of dipoles was fit to each of the components of interest; the residual variance (RV) for each dipole pair model was minimized over the 20-msec latency range around the peak of the component. Dipole pairs
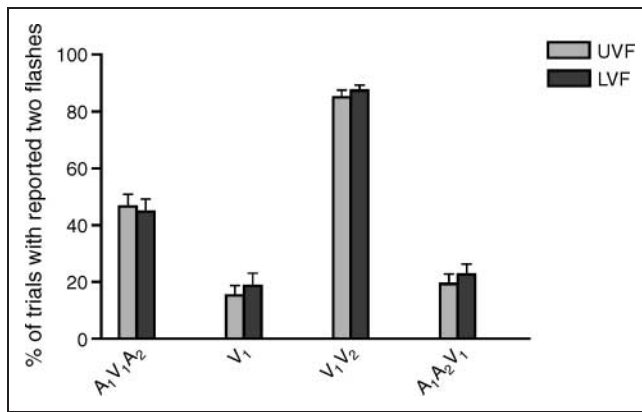
Figure 2. Comparisons of perceptual reports in the upper (UVF) and lower visual fields (LVF) for all experimental stimuli that contained flashes.

were constrained to be mirror-symmetrical with respect to location but were free to vary in orientation.

To visualize the anatomical brain regions giving rise to the different components, the locations of BESA source dipoles were transformed into the standardized coordinate system of Talairach and Tournoux (1988) and projected onto a structural brain image supplied by MRIcro (Rorden & Brett, 2000) using Analysis of Functional NeuroImaging (AFNI; Cox, 1996) software.

Response Contingent Analysis

In this analysis, the Ill_Diff$_{ATT-UVF}$ and the Ill_Diff$_{ATT-LVF}$ waveforms were calculated separately on trials where the extra flash illusion was perceived (SEE2 trials) and compared with the waveforms on trials where the illusory second flash was not seen (SEE1 trials). The main component in the SEE2 minus SEE1 difference waveforms was measured at 136–160 msec (ND150). This component was quantified as the mean voltage over the same fronto-central electrode clusters as those used to measure PD180 in the Ill_Diff waveforms (see Data Analysis section).

RESULTS

Behavioral Results

Subjects indicated by pressing one of two buttons the number of flashes perceived (1 or 2) in each stimulus combination in the attended field that contained flashes. Mean percentages of responses on which two flashes were reported over all 15 subjects are given in Figure 2 and corresponding reaction times in Table 1. Subjects reported perceiving an illusory second flash on an average of 47% and 45% of the $A_1V_1A_2$ attended trials in UVF and LVF, respectively. Subjects responded accurately to both unimodal visual stimuli (V_1 and V_1V_2) and to the bimodal control stimulus ($A_1A_2V_1$). Importantly, there was no correlation between subjects' illusory two-flash reports on the $A_1V_1A_2$ stimulus and incorrect two-flash responses on the $A_1A_2V_1$ control in UVF [$r(13) = -.25, p = ns$] or LVF [$r(13) = -.21, p = ns$], thereby showing that perception of the illusion was not due to a general tendency to report two flashes. For all stimuli, there were no significant differences in behavioral performance between stimuli presented in the UVF versus the LVF either for detection rates [UVF vs. LVF: $F(1, 14) = 1.79, p = ns$] or for reaction times [UVF vs. LVF: $F(1, 14) = 2.17, p = ns$].

Reaction times differed significantly across stimulus conditions [$F(3, 42) = 17.98, p < .0001$]. Reaction times on unimodal double-flash trials were found to be significantly faster than reaction times on single-flash trials in both visual fields [V_1V_2 vs. V_1: $F(1, 14) = 27.30, p < .0002$]. Reaction times on $A_1V_1A_2$ trials on which two flashes were perceived versus when only a single flash was seen did not differ significantly overall [$F(1, 14) = 0.55, p = ns$]. For $A_1V_1A_2$ stimuli in the lower field, however, a trend similar to unimodal flash trials was observed with faster reaction times on illusory trials (on which two flashes were seen) [Visual field × Illusory trials interaction: $F(1, 14) = 4.96, p < .05$].

ERP Results

The grand-averaged ERPs (over all 15 subjects) elicited by the illusion-inducing $A_1V_1A_2$ stimulus and by its unimodal components, V_1 and A_1A_2, are shown for attended

Table 1. Mean Reaction Times (*SEM*) for Reporting the Number of Flashes Seen (One or Two) for All Stimulus Combinations Containing One or Two Visual Stimuli Presented in the Upper (UVF) and Lower Visual Fields (LVF)

| Stimulus	*Mean RT (SEM), msec* *UVF [One/Two Flashes Perceived Trials]*	*Mean RT (SEM), msec* *LVF [One/Two Flashes Perceived Trials]*
$A_1V_1A_2$	655 (13)	647 (11)
	[656 (12)/655 (15)]	[654 (9)/639 (15)]
V_1	654 (11)	656 (12)
V_1V_2	610 (9)	600 (10)
$A_1A_2V_1$	612 (14)	602 (16)

and unattended presentations in the UVF and LVF in Figures 3 and 4, respectively. Unimodal visual ERPs to V_1 had characteristic P1 (120 msec), N1 (180 msec), and P2 (200 msec) components with maxima at posterior electrode sites, and an earlier N1 (165 msec) at anterior sites. The unimodal ERPs to A_1A_2 included auditory-evoked P1 (60 msec), N1 (105 msec), and P2 (180 msec) components with maxima at fronto-central electrode sites. The sharp positive-going deflection that peaks at around 20 msec in the A_1A_2 and $A_1V_1A_2$ waveforms was produced by the sound-evoked postauricular muscle reflex (P. A.; Picton, Hillyard, Krausz, & Galambos, 1974) recorded at the mastoid reference site.

For the unimodal stimuli, attention effects on ERP components were only found within the visual modality, with both the P1 and N1 components being enlarged over occipital and parieto-occipital electrodes to visual stimuli in the attended field (see Supplementary Table S1). These effects were consistent with visual attention effects found in previous studies (e.g., Martínez, Teder-Sälejärvi, & Hillyard,

2007; Martínez et al., 2001, 2006; Di Russo et al., 2003). For the bimodal $A_1V_1A_2$ stimuli, the attended ERPs showed a larger positivity relative to the unattended waveforms within the 120–150 msec time range (corresponding to the visual P1 component) over occipital electrode sites. The attention differences for the bimodal stimuli were not further characterized in these ERPs, however, as the effect of attention on the auditory and visual components of the configuration could not be separated from the attention effects on the cross-modal interaction of the unimodal components. Hence, in subsequent analyses, cross-modal difference waves were calculated (see Methods), and attention effects were analyzed on the cross-modal interaction components therein.

The cross-modal interaction difference waves for the $A_1V_1A_2$ stimuli were calculated for both UVF and LVF and for both attended and unattended conditions (Figure 5). For the attended difference waves, Ill_Diff$_{ATT–UVF}$ and Ill_Diff$_{ATT–LVF}$, the earliest significant components were prominent positivities at occipital sites that extended over

Figure 3. Grand-average ERPs elicited by attended and unattended bimodal ($A_1V_1A_2$) stimuli and by their unimodal constituents (V_1 and A_1A_2) presented in the UVF. (A) ERPs elicited by the illusion-inducing $A_1V_1A_2$ stimulus when attended (ATT) and unattended (UNATT). (B) Corresponding ERPs as in (A) elicited by the unimodal V_1 stimulus. (C) Corresponding ERPs as in (A) elicited by the unimodal A_1A_2 stimulus. Recordings are from left and right fronto-central (FC1, 2) and occipital (O1, 2) sites.
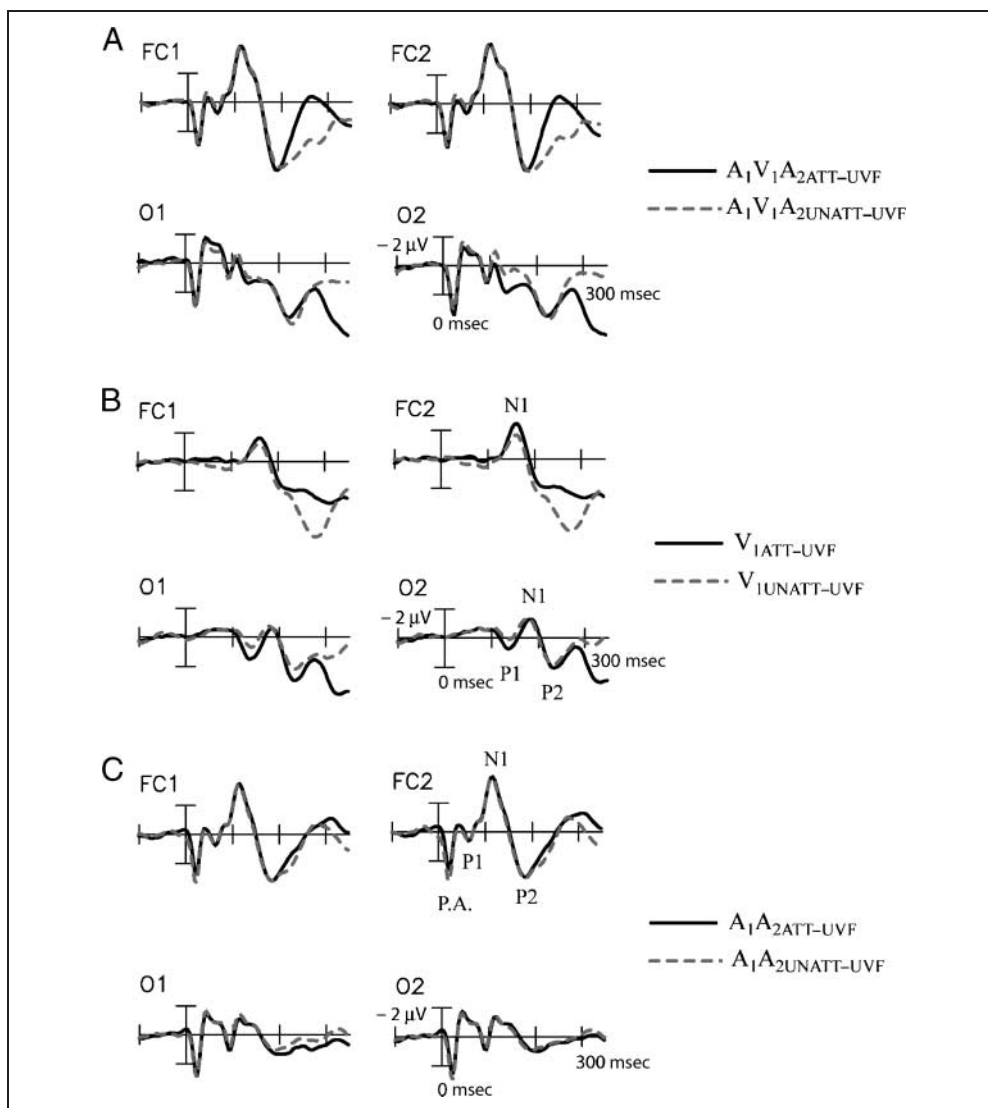
Figure 4. Same as Figure 3 for bimodal ($A_1V_1A_2$) and unimodal (V_1 and A_1A_2) stimuli presented in the LVF.
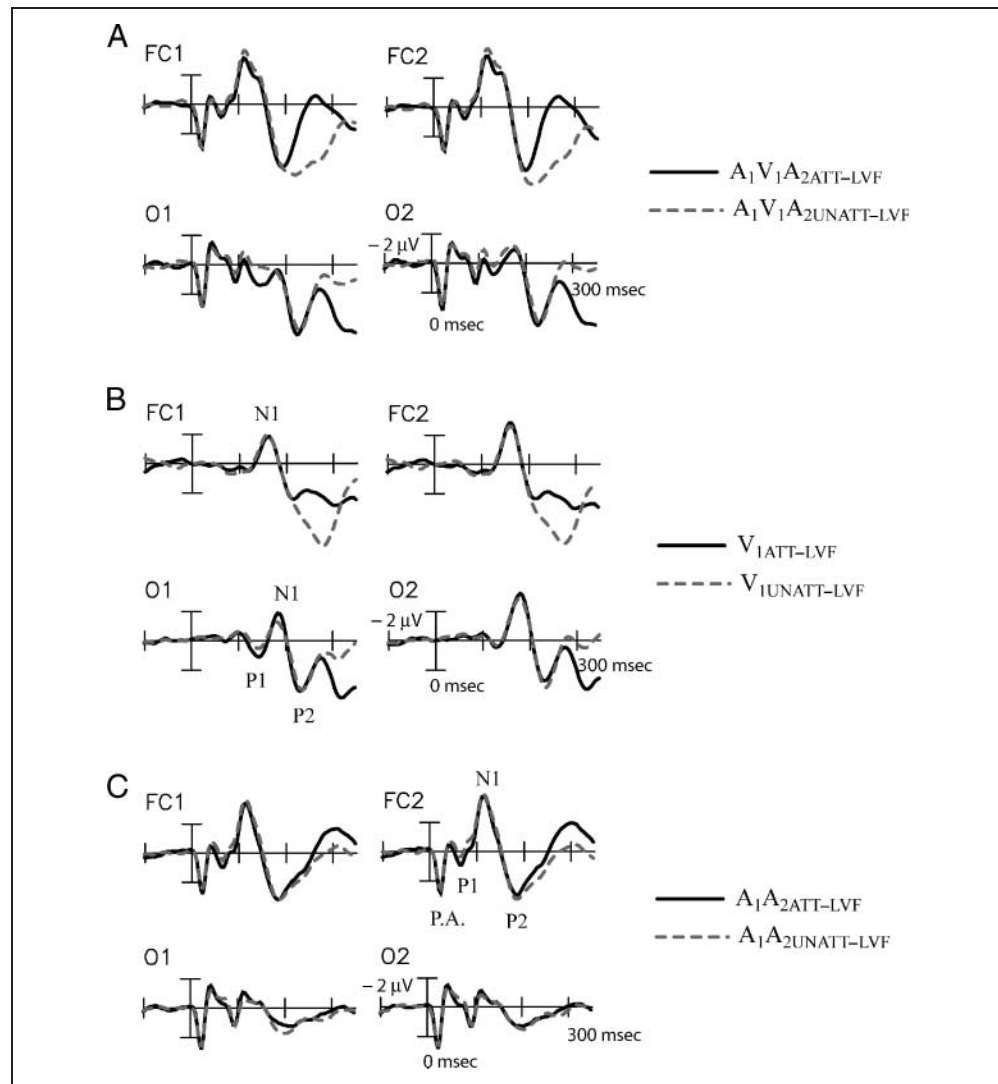
the interval 100–150 msec. These positivities were quantified around their early peaks, PD120 in the 112–132 msec time interval for UVF and PD110 in the 104–124 msec latency range for LVF, respectively. The PD120/PD110 deflections were followed by a larger positivity peaking at 180 msec over anterior sites in both attended UVF and LVF waveforms, termed PD180. The PD180 also extended posteriorly to the O1/O2 sites at reduced amplitude. The final components characterized within the attended Ill_Diff waves were negativities within the 240–260 msec interval (ND250) in UVF and the 228–248 msec interval (ND240) in LVF. The amplitudes and significance of these components with respect to the prestimulus baseline are given in Table 2. As in our previous study (Mishra et al., 2007), components occurring after 300 msec were not analyzed because of the likelihood that neural activity related to decision-making and response preparation would be confounded with activity related to cross-modal interaction and perceptual processing.
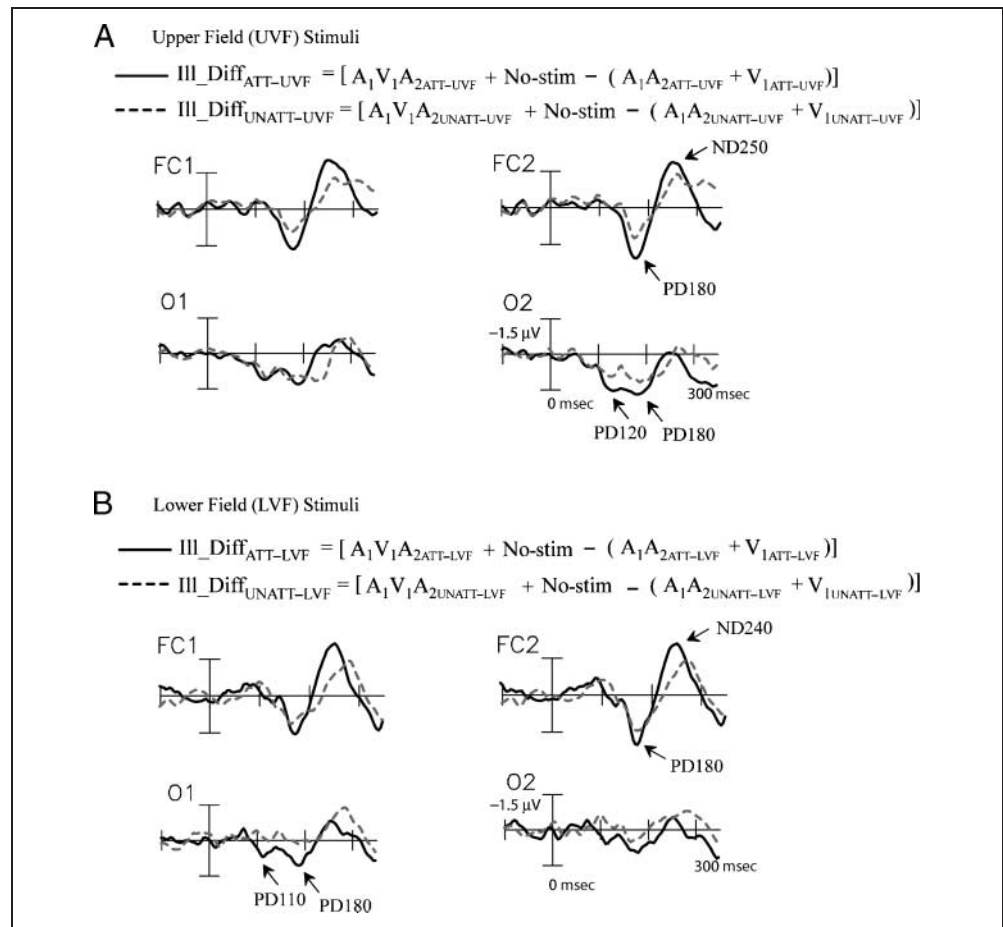
The amplitudes of these attended difference wave components in individual subjects did not correlate with the percentage of $A_1V_1A_2$ trials on which the subjects saw the illusion. Whereas our previous study found that subjects with larger PD120 components saw the illusion more frequently [$r(32) = .48, p < .005$; Mishra et al., 2007], such a relationship was not observed here for the PD120/PD110 [UVF: $r(13) = .11, p = ns$; LVF: $r(13) = .21, p = ns$], probably because subjects here were selected for perceiving the illusion more than half the time, and thus, had a narrower range of variation in their perceptual reports.

The difference wave components for the unattended stimuli, Ill_Diff$_{UNATT–UVF}$ and Ill_Diff$_{UNATT–LVF}$, were characterized in the same time intervals as the components in the attended difference waveforms (Table 2). The early PD120/PD110 components did not reach significance in the unattended waveforms, whereas the later PD180 and ND250/ND240 components were much reduced relative to their attended counterparts. Neither the PD180 in the Ill_Diff$_{UNATT–LVF}$ waveforms nor the ND250 in the Ill_Diff$_{UNATT–UVF}$ waves reached statistical significance.

The scalp voltage distributions of the attended and unattended Ill_Diff wave components in UVF and LVF are

Figure 5. Grand-average Ill_Diff difference waves that reflect the cross-modal neural interactions elicited by the illusion-inducing $A_1V_1A_2$ bimodal stimulus when attended (ATT) and unattended (UNATT). (A) Ill_Diff difference waves for attended and unattended stimuli in the UVF. (B) Corresponding difference waves as in (A) for stimuli in the LVF. Recordings are from left and right fronto-central (FC1, 2) and occipital (O1, 2) sites.

Table 2. Mean Amplitudes of ERP Components in the Difference Waves Associated with the Illusory Flash Generating $A_1V_1A_2$ Stimulus When Attended (ATT) or Unattended (UNATT) in the Upper (UVF) and Lower (LVF) Visual Fields

|	ERP Component	Amplitude (μV)	SEM (μV)	t(14)	p
Ill_Diff$_{ATT-UVF}$	PD120 (112–132 msec)	1.34	0.44	3.07	<.009
	PD180 (164–184 msec)	1.75	0.30	5.74	<.0001
	ND250 (240–260 msec)	−1.33	0.63	2.18	<.05
Ill_Diff$_{UNATT-UVF}$	(112–132 msec)	0.49	0.31	1.57	ns
	PD180 (164–184 msec)	0.97	0.37	2.64	<.02
	(240–260 msec)	−0.39	0.37	1.05	ns
Ill_Diff$_{ATT-LVF}$	PD110 (104–124 msec)	0.48	0.22	2.17	<.05
	PD180 (164–184 msec)	1.18	0.35	3.38	<.005
	ND240 (228–248 msec)	−1.36	0.40	3.43	<.005
Ill_Diff$_{UNATT-LVF}$	(104–124 msec)	−0.33	0.23	1.44	ns
	(164–184 msec)	0.32	0.37	0.88	ns
	ND240 (228–248 msec)	−0.47	0.21	2.29	<.04

Components were measured over scalp sites of maximal amplitude, as described in the Methods. Significance levels of component amplitudes were tested with respect to the 100-msec prestimulus baseline.

shown in Figure 6, along with statistical comparisons of the attended versus unattended amplitudes. Both PD120 and PD110 attended components in UVF and LVF, respectively, had occipital scalp distributions (Figure 6A). The right-hemispheric preponderance of PD120 in the Ill_Diff$_{ATT-UVF}$ wave did not reach significance [$F(1, 14) = 3.24$, $p = ns$], nor did the slight left laterality of PD110 in Ill_Diff$_{ATT-LVF}$ [$F(1, 14) = 3.36$, $p = ns$]. The topographies of the unattended components in the PD120/PD110 latency ranges are also shown, but their amplitudes were not statistically significant, as mentioned above. For the PD120/PD110 components, there was a significant effect of attention for both UVF and LVF stimuli (Figure 6A). The subsequent PD180 had a fronto-central distribution in the Ill_Diff$_{ATT-UVF}$, Ill_Diff$_{ATT-LVF}$, and Ill_Diff$_{UNATT-LVF}$ difference waves with a nonsignificant right-hemispheric preponderance. For the Ill_Diff$_{UNATT-UVF}$ waveform, however, the topography of PD180 was shifted posteriorly to centro-parietal sites. Although we have no explanation for this shift in scalp distribution, we note that a similar

shift in topography was also reported in a previous study comparing attended and unattended cross-modal difference waves (Talsma & Woldorff, 2005). The effect of attention on PD180 was significant for both UVF and LVF stimuli (Figure 6B). Lastly, the ND250 (UVF) and ND240 (LVF) components had prominent fronto-central distributions with significant attention effects in both visual fields (Figure 6C).

The effect of attention on the PD120/PD110 components was further characterized by subtracting the unattended from the attended Ill_Diff waveforms at each visual field location. Figure 7A shows the waveforms resulting from this subtraction at occipital electrodes, and Figure 7B shows the corresponding scalp distributions of these attention effects. The distributions of the attention effects in UVF versus LVF were compared following normalization according to the method of McCarthy and Wood (1985). The voltage topography of the PD120 attention effect (UVF) was found not to differ significantly from that of the PD110 attention effect (LVF) [Visual field × Electrode interaction: $F(17, 238) =$
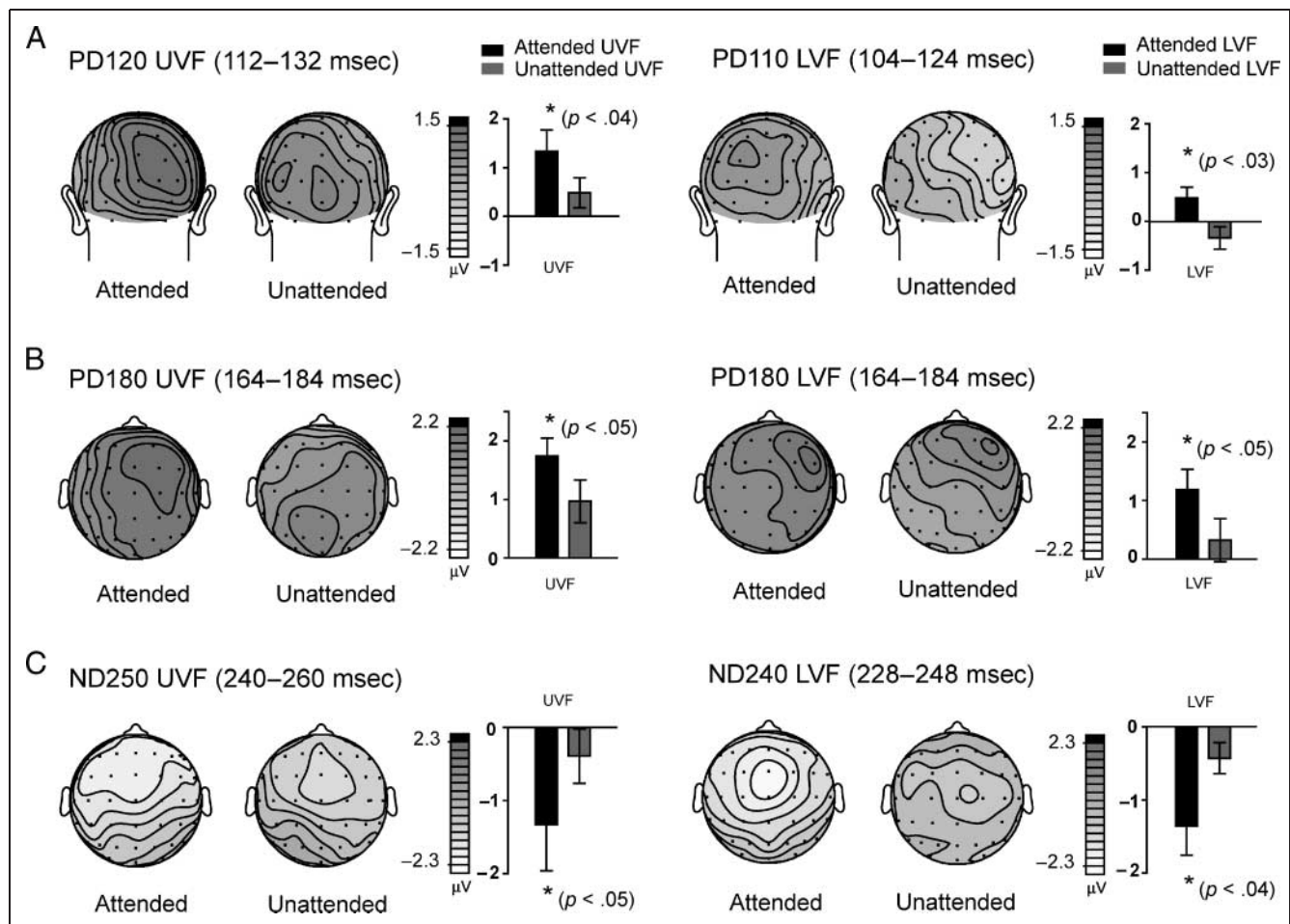
Figure 6. Topographical voltage maps of the three major components in the Ill_Diff$_{ATT}$ and Ill_Diff$_{UNATT}$ difference waves for UVF stimuli (left column) and LVF stimuli (right column). (A) PD120/PD110 component. (B) PD180 component. (C) ND250/ND240 component. Bar graphs next to the voltage maps depict the mean amplitude differences between the attended and unattended components measured at electrode sites where their amplitudes were maximal (see Methods).

Figure 7. Attention effect on the PD120/PD110 components. (A) Grand-average attention effect formed by subtracting the Ill_Diff$_{UNATT}$ difference wave from the Ill_Diff$_{ATT}$ difference wave at left and right occipital (O1, 2) sites for UVF and LVF stimuli. (B) Topographical voltage maps of the PD120/PD110 attention effects. (C) Estimated dipole sources modeled using BESA for the PD120/ PD110 components in the grand-average Ill_Diff attended waveforms (in black) and for their corresponding attention effects (in gray) for UVF stimuli in the left column and LVF stimuli in the right column. Results are shown on a standard fMRI rendered brain in Talairach space.
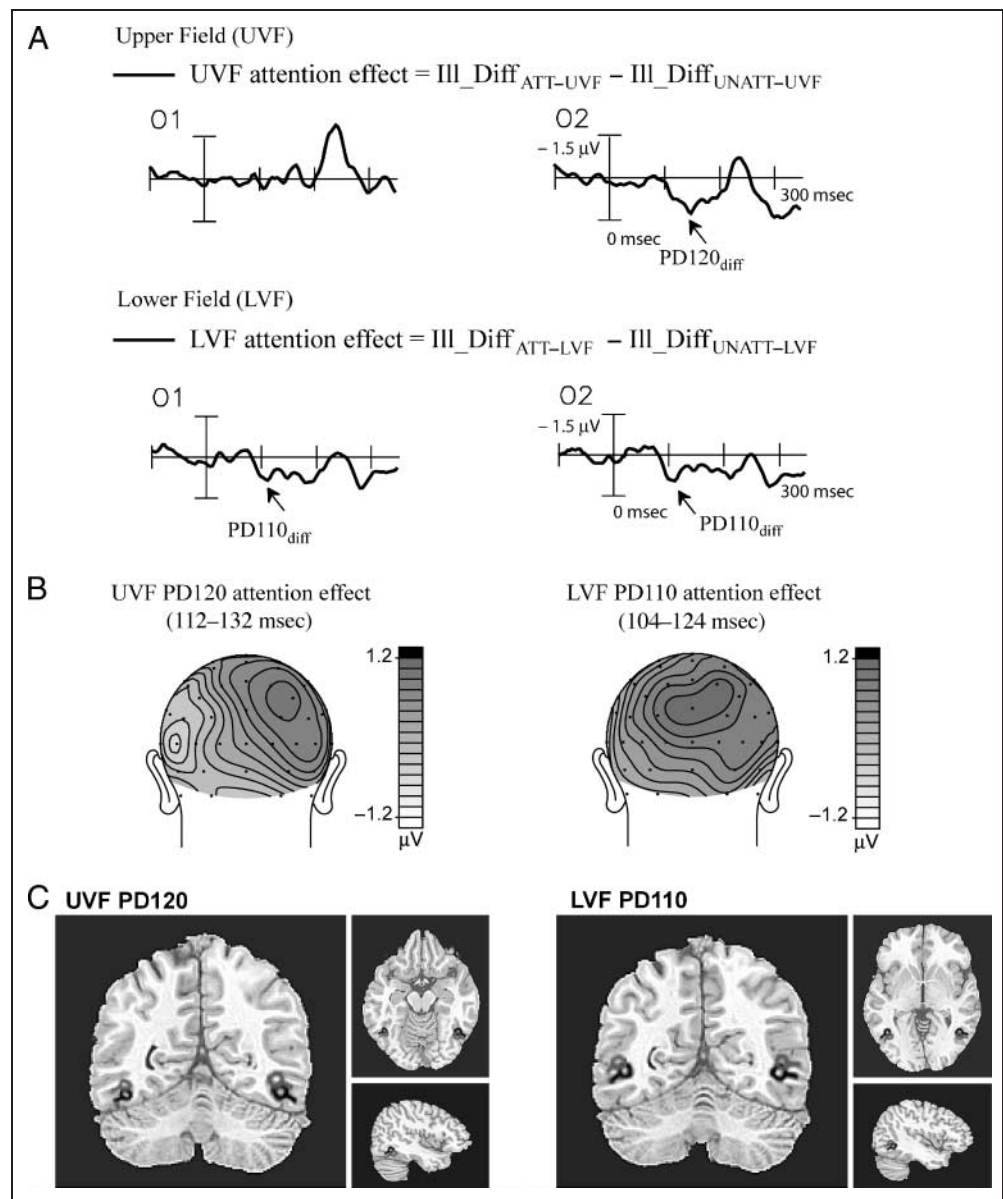
0.67, $p = ns$], although the UVF effect appears to be more lateralized.

Source Analysis

The neural generators of the significant components identified in the attended Ill_Diff difference waves as well as the PD120$_{diff}$/PD110$_{diff}$ attention effects in the UVF and LVF locations were modeled using dipole source localization. Pairs of dipoles were fit to the scalp topographies of the components using the BESA algorithm (Scherg, 1990). The location of the BESA dipoles was transformed into the standardized coordinate system of Talairach and Tournoux (1988) and superimposed on the rendered cortical surface of a single individual's brain. Talairach coordinates of the dipole pairs modeled to each component and an estimate

of their goodness of fit as reflected by residual variance are listed in Table 3.

In the UVF, the early PD120 in the Ill_Diff$_{ATT-UVF}$ difference wave and the PD120 attention effect were both localized to ventral–lateral extrastriate cortex in the region of fusiform gyrus (Figure 7C). The dipole in the right hemisphere accounted for greater component variance than the left hemisphere dipole. In the LVF, PD110 and its corresponding attention effect were also localized to lateral extrastriate visual cortex approximately 10–15 mm superior to the PD120 dipoles (Figure 7C). The left and right hemisphere dipoles accounted for equivalent variance in the case of the LVF dipole fits. The sources of the later components, both PD180 and ND250/ND240 in the attended Ill_Diff$_{ATT-UVF}$ and Ill_Diff$_{ATT-LVF}$ waveforms, were consistently localized to the region of superior temporal cortex (Table 3). There were no apparent differences

Table 3. Talairach Coordinates and Corresponding Brain Regions of the Dipole Fits as Modeled by BESA for the Significant Components in the Attended Ill_Diff Waveforms in UVF and LVF

	ERP Component	x (mm)	y (mm)	z (mm)	Brain Region	RV (%)
Ill_Diff$_{ATT-UVF}$	PD120	±43	−55	−13	inferior occipito-temporal cortex	8
	PD180	±47	−30	1	MTG/STG	9
	ND250	±53	−21	11	vicinity of STG	6
Ill_Diff$_{ATT-LVF}$	PD110	±42	−55	−1	inferior occipito-temporal cortex	13
	PD180	±40	−20	−1	MTG/STG	11
	ND240	±40	−29	4	vicinity of STG	6
UVF attention effect	PD120$_{diff}$	±40	−54	−11	inferior occipito-temporal cortex	14
LVF attention effect	PD110$_{diff}$	±46	−60	5	inferior occipito-temporal cortex	17
Ill_Diff$_{ATT-UVF/SEE2-SEE1}$	ND150	±46	−5	4	STG	10
Ill_Diff$_{ATT-LVF/SEE2-SEE1}$	ND150	±49	−23	11	STG	8

The residual variance (RV) values shown here are for models consisting of a single pair of dipoles fit to the indicated component. Coordinates of dipole fits for the PD120/PD110 attention effects, and for the ND150 component in the SEE2–SEE1 Ill_Diff waveforms in both UVF and LVF are also shown. MTG = medial temporal gyrus; STG = superior temporal gyrus.

in the dipole localizations for these later components between stimuli in the UVF versus LVF locations.

Response Contingent Analysis

The SEE2 and SEE1 difference waves in both UVF (Figure 8A) and LVF (Figure 8B), which distinguished the trials on which the illusion was seen versus not seen, differed significantly from each other over anterior scalp sites in the 136–160 msec latency range, as evident in an overall ANOVA over both VFs [$F(1, 14) = 9.19, p < .009$]. These differences were visualized in the SEE2 minus SEE1 trials difference waveforms as a negative component peaking at 150 msec (ND150), which was significant in UVF [$t(1, 14) = 2.31, p < .04$] as well as in the LVF [$t(1, 14) = 2.67, p < .02$]. The ND150 components in both the UVF and LVF had amplitude maxima over fronto-central sites, whereas no differences between trials were found to be significant over occipital electrodes. The scalp topographies of the ND150 components in the UVF and LVF are shown in Figure 9A. The generators of the ND150 in both UVF and LVF were localized using BESA dipole fits to the region of superior temporal gyrus (Table 3; Figure 9B).

DISCUSSION

Subjects reported perceiving an illusory second flash in the cross-modal $A_1V_1A_2$ stimulus on an average of 44–46% of attended trials in both the upper and lower peripheral visual field locations. The present study aimed to investigate the effect of attention on the neural interactions associated with the extra flash illusion, which were revealed in difference waves formed by subtracting the ERPs elicited by the unimodal components (V_1 and A_1A_2) from the ERPs

to the cross-modal combination ($A_1V_1A_2$) for each spatial location and attended state. These interaction difference waves associated with the illusion were termed Ill_Diff. For both UVF and LVF stimuli, the Ill_Diff waveforms included a PD120 (PD110 in the lower field) component, which was localized by dipole modeling to ventral occipito-temporal extrastriate visual cortex, and subsequent PD180 and ND250 (ND240 in lower field) components with generators estimated to lie in superior temporal cortex, a well-known region of polymodal interaction (Beauchamp, 2005; Calvert, 2001). These difference wave components were highly similar to those previously characterized in a study of the neural basis of the illusion (Mishra et al., 2007), and all three components were found to be enhanced by attention. These results demonstrate that cross-modal interactions associated with the extra flash illusion are strongly affected by the spatial allocation of attention and are not the result of wholly automatic integration processes.

In our previous study (Mishra et al., 2007), the amplitude of the early PD120 component was found to correlate positively with the proportion of trials on which an individual subject perceived the illusion. Because individuals who did not perceive the illusion were excluded from the present study, the PD120 amplitude did not fluctuate widely across subjects, and because of this range restriction, a significant correlation with behavior was not obtained. The present study revealed a new aspect of the PD120, however, namely that spatially directed attention modulates its amplitude. When the $A_1V_1A_2$ stimulus was actively ignored in either UVF or LVF, the PD120 component was reduced to nonsignificant levels in the interaction difference waves. This is the first report, to our knowledge, of an attention effect on cross-modal interactions in sensory specific cortex. In our previous study (Mishra et al., 2007), explicit instructions to attend were not provided, but we would

Figure 8. Comparison of attended Ill_Diff difference waves on trials when the extra flash illusion was seen (SEE2) versus not seen (SEE1) for stimuli in the UVF (A) and in the LVF (B). Recordings are from left and right fronto-central (FC1, 2) and occipital (O1, 2) sites.
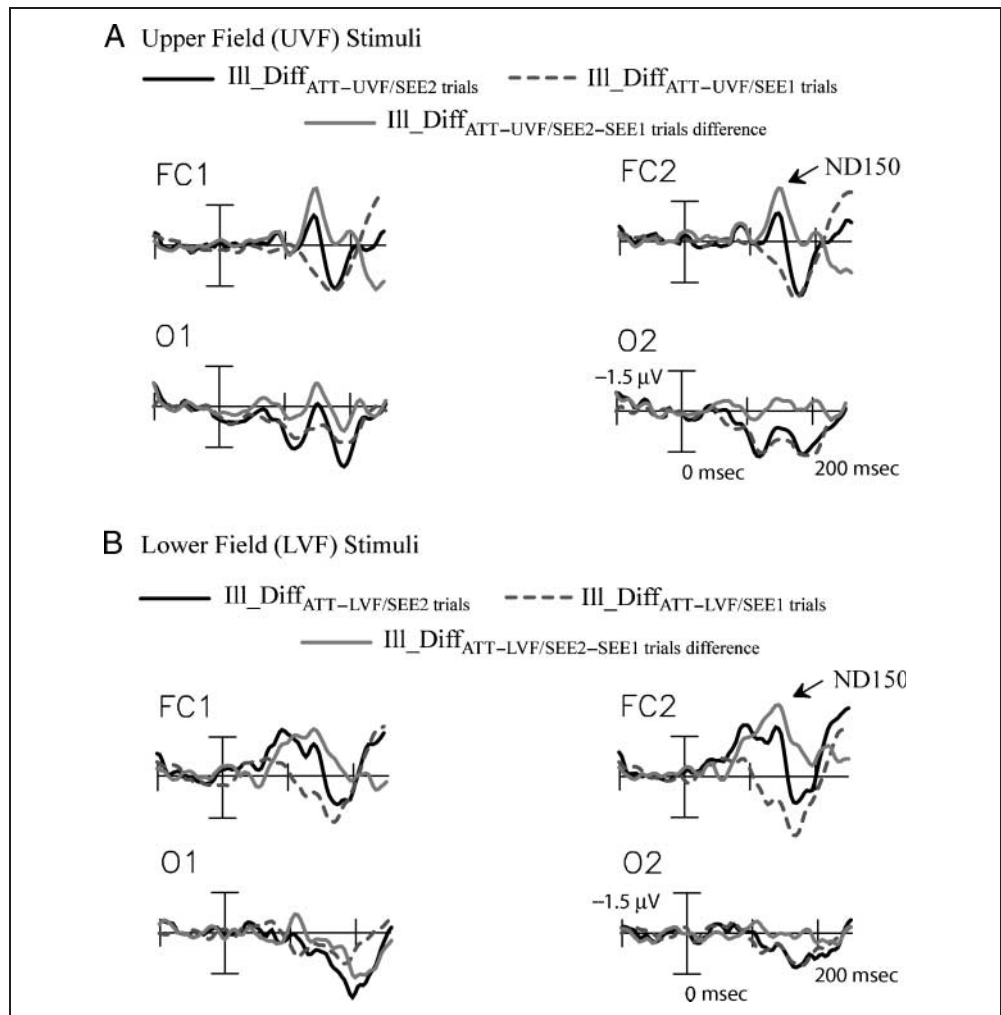
Figure 9. Topographical voltage maps and dipole sources of the ND150 component that was specifically elicited by the $A_1V_1A_2$ stimulus on SEE2 trials. Maps and sources are for the ND150 isolated by subtracting the Ill_Diff waveforms on SEE2 minus SEE1 trials for UVF stimuli (left column) and LVF stimuli (right column).
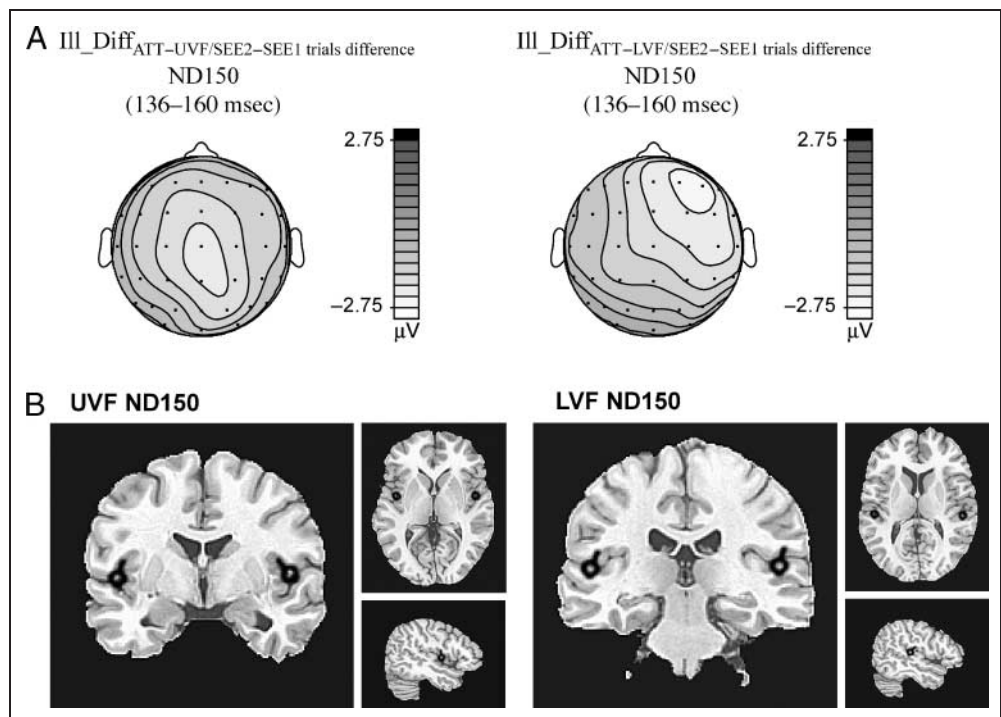
assume that subjects needed to maintain attention to all stimuli in order to make the assigned perceptual judgments. Taken together, our previous and current investigations suggest that the PD120 interaction component characterizes subjects who frequently perceive the extra flash illusion but is only elicited robustly when attention is directed toward the multimodal stimuli.

In order to confirm that the correlation between PD120 and illusory flash reports in our previous study (Mishra et al., 2007) was based on the subjects' perception of the illusion rather than on response bias, we reanalyzed data from that study in signal detection terms. This analysis showed that the PD120 was correlated across subjects with perceptual sensitivity to the illusion, d' [$r(32) = -.60, p < .0002$] and not with response bias, β [$r(32) = -.02, p = ns$]. These measures were calculated according to the method suggested by Watkins et al. (2006), with two-flash responses to V_1V_2 categorized as hits and illusory two-flash responses to $A_1V_1A_2$ categorized as false alarms. Moreover, the frequency of incorrect two-flash responses to the $A_1A_2V_1$ control stimulus, which serves as another index of response bias, did not correlate with PD120 amplitude [$r(32) = -.03, p = ns$] in our previous study (Mishra et al., 2007). This analysis provides evidence that PD120 is related more to the perceptual experience of the illusion than to response bias.

The neural generators of the PD120 component in UVF as well as LVF (PD110) were localized by dipole modeling to ventral–lateral extrastriate visual cortex in or near fusiform gyrus. The timing of the PD120/PD110 is similar to that of the early P1 component (80–120 msec) of the visual ERP (Hillyard, Vogel, & Luck, 1998), which is also strongly enhanced by spatial attention. Source localization of the PD120/PD110 further suggests that these components may arise from activity in neural populations in extrastriate visual cortex similar to those that give rise to the visual P1 (Martínez et al., 2001, 2006, 2007; Di Russo et al., 2002, 2003). The PD110 to lower field stimuli was localized about 10–15 mm superior to the PD120 generators for upper field stimuli. This difference might be accounted for by differential activation of retinotopic visual areas as a function of stimulus location, but it should be noted that such a small difference is difficult to confirm using inverse source modeling techniques. In any case, the results of this source localization analysis, together with the lack of polarity inversion of the PD120 component in the upper versus lower field comparison, provide strong evidence that its predominant generator site is situated in extrastriate areas outside primary visual cortex.

The PD120 (/PD110) emerges very rapidly, within 30–60 msec after onset of the second sound (A_2) within the $A_1V_1A_2$ stimulus. This suggests that direct connections between auditory and visual areas may be responsible for the generation of this component (Mishra et al., 2007). Such connections have been characterized in recent years in anatomical labeling studies in primates (Clavagnier, Falchier, & Kennedy, 2004; Rockland & Ojima, 2003; Falchier, Clavagnier,

Barone, & Kennedy, 2002) and have been shown to be denser in visual areas higher in the visual hierarchy than primary cortex. This anatomy is consistent with the localization of the PD120 to lateral extrastriate visual cortex. It is unlikely that the PD120 component could be driven via feedback from higher areas such as multisensory superior temporal cortex, given that feedback usually has a slower time course. Moreover, there was no ERP evidence to suggest modulation of polymodal cortex prior to the PD120/PD110. Although the exact pathways involved cannot be ascertained from the present evidence, the early cross-modal modulation observed here highlights the intimate link between processing in different sensory modalities that is being increasingly found in studies of multisensory integration (Driver & Noesselt, 2008; Schroeder & Foxe, 2005).

Following the PD120 component in the interaction difference waveforms were two large components, PD180 and ND250 (ND240 in the lower field), which were also observed in our previous study (Mishra et al., 2007). These late interaction components have been found to be elicited by cross-modal stimulus combinations in many previous multisensory ERP investigations (Talsma & Woldorff, 2005; Teder-Sälejärvi et al., 2002, 2005; Molholm et al., 2002). These results, along with the source localization of the later components to the superior temporal area, a well-known polymodal region (Calvert, 2001), support the hypothesis that PD180 and ND250 (/ND240) reflect general aspects of cross-modal interaction not specific to the extra flash illusion (Mishra et al., 2007). The present study found these later components to have reduced amplitudes in the unattended waveforms. This suggests that attention can significantly affect processes of multisensory integration in general and is in line with many previous investigations showing an influence of attention on cross-modal processing (Talsma et al., 2007; Busse et al., 2005; Talsma & Woldorff, 2005; McDonald et al., 2001, 2003; Talsma & Kok, 2002; Eimer & Driver, 2001; Teder-Sälejärvi et al., 1999; Hillyard et al., 1984).

In the response contingent analysis, the cross-modal interaction (Ill_Diff) waves revealed an enlarged negativity (ND150) on trials when the illusion was perceived versus not perceived. The ND150 was localized to superior temporal cortex for stimuli in both the UVF and LVF. This analysis did not reveal any occipital ERP differences associated with the perception of the illusion in the time range analyzed (0–300 msec). In our previous study (Mishra et al., 2007), similar negative difference wave components were found on trials where the illusion was perceived within a latency range of 100–150 msec (ND110 and ND130). These components also had neural generators localized to superior temporal gyrus in the vicinity of auditory cortex, and again no trial-specific ERP modulation of visual cortex was found in association with the illusion. Based on the similar ERP findings in these two studies, we would hypothesize that the illusion-trial-specific neural activity in temporal regions provides the proximal trigger for perceiving the illusion. These temporal lobe activations may

interact via connections to visual cortex with neural activity reflected in the PD120 component that is specific to individuals predisposed to see the illusion, thereby giving rise to the illusory percept. Thus, we propose that the PD120 reflects neural activity that is necessary for producing the illusion but is not sufficient, in the absence of the temporal cortex activation reflected in the ND110/130/150 that is specifically elicited on trials when the illusion is perceived.

Recently, response contingent activations associated with the double-flash illusion have been observed in primary visual cortex (V1) in fMRI investigations (Watkins, Shams, Josephs, & Rees, 2007; Watkins et al., 2006). Our results do not rule out the possibility of such a primary visual cortex effect, and if such neural activity was not well time-locked to the stimuli (such as induced oscillations, e.g., Mishra et al., 2007; Bhattacharya, Shams, & Shimojo, 2002), it might not be registered in the averaged ERP. It is also possible that the trial-specific involvement of primary visual cortex detected by fMRI does not occur in the initial 0–300 msec response phase analyzed here but is rather driven by feedback from higher polymodal areas at longer latencies. A distinction between these possibilities is not feasible using fMRI. As noted above, however, we did not observe any occipital modulations in our response contingent analyses, and thus, obtained no evidence for delayed feedback to V1 (at least prior to 300 msec poststimulus onset). It should also be noted that the illusion-trial-specific V1 activation reported in the fMRI studies (Watkins et al., 2006, 2007) was not correlated with an individual's frequency of perceiving the illusion, and thus, is not likely to be the hemodynamic counterpart of the neural activity giving rise to the PD120 component.

In summary, we found that multisensory integration processes previously shown to be closely linked to the auditory-induced extra flash illusion can be significantly enhanced by selective spatial attention. The earliest modulation by attention was found at 100–130 msec within ventral occipito-temporal extrastriate visual cortex on a component (PD120/PD110) that was previously shown to predict the frequency with which individuals perceived the illusion (Mishra et al., 2007). This critical component was found to be enhanced in amplitude by attention and did not reach significance when the stimuli were unattended. Following the PD120, there was a modulation of superior temporal activity (136–160 msec), with an enhanced negativity on trials where the illusion was perceived. Further research is needed to relate the attention-related modulations of these components with corresponding modulations of the illusory percept.

REFERENCES

Amedi, A., Von Kriegstein, K., Van Atteveldt, N. M., Beauchamp, M. S., & Naumer, M. J. (2005). Functional imaging of human crossmodal identification and object recognition. *Experimental Brain Research, 166,* 559–571.

Arden, G. B., Wolf, J. E., & Messiter, C. (2003). Electrical activity in visual cortex associated with combined auditory and visual stimulation in temporal sequences known to be associated with a visual illusion. *Vision Research, 43,* 2469–2478.

Beauchamp, M. S. (2005). See me, hear me, touch me: Multisensory integration in lateral occipital–temporal cortex. *Current Opinion in Neurobiology, 15,* 145–153.

Bertelson, P. (1999). Ventriloquism: A case of crossmodal perceptual grouping. *Advances in Psychology, 129,* 347–362.

Bhattacharya, J., Shams, L., & Shimojo, S. (2002). Sound-induced illusory flash perception: Role of gamma band responses. *NeuroReport, 13,* 1727–1730.

Bonath, B., Noesselt, T., Martínez, A., Mishra, J., Schwiecker, K., Heinze, H. J., et al. (2007). Neural basis of the ventriloquist illusion. *Current Biology, 17,* 1697–1703.

Busse, L., Roberts, K. C., Crist, R. E., Weissman, D. H., & Woldorff, M. G. (2005). The spread of attention across modalities and space in a multisensory object. *Proceedings of the National Academy of Sciences, U.S.A., 102,* 18751–18756.

Calvert, G. A. (2001). Crossmodal processing in the human brain: Insights from functional neuroimaging studies. *Cerebral Cortex, 11,* 1110–1123.

Calvert, G. A., Stein, B. E., & Spence, C. (2004). *The handbook of multisensory processing.* Cambridge, MA: MIT Press.

Carrasco, M. (2006). Covert attention increases contrast sensitivity: Psychophysical, neurophysiological and neuroimaging studies. *Progress in Brain Research, 154,* 33–70.

Clark, V. P., Fan, S., & Hillyard, S. A. (1995). Identification of early visual evoked potential generators by retinotopic and topographic analyses. *Human Brain Mapping, 2,* 170–187.

Clavagnier, S., Falchier, A., & Kennedy, H. (2004). Long-distance feedback projections to area V1: Implications for multisensory integration, spatial awareness, and visual consciousness. *Cognitive, Affective & Behavioral Neuroscience, 4,* 117–126.

Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers & Biomedical Research, 29,* 162–173.

Di Russo, F., Martínez, A., & Hillyard, S. A. (2003). Source analysis of event-related cortical activity during visuo-spatial attention. *Cerebral Cortex, 13,* 486–499.

Di Russo, F., Martínez, A., Sereno, M. I., Pitzalis, S., & Hillyard, S. A. (2002). Cortical sources of the early components of the visual evoked potential. *Human Brain Mapping, 15,* 95–111.

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

Eimer, M. (1999). Can attention be directed to opposite locations in different modalities? An ERP study. *Clinical Neurophysiology, 110,* 1252–1259.

Eimer, M., & Driver, J. (2001). Crossmodal links in endogenous and exogenous spatial attention: Evidence from event-related brain potential studies. *Neuroscience & Biobehavioral Reviews, 25,* 497–511.

Eimer, M., & Schroger, E. (1998). ERP effects of intermodal attention and cross-modal links in spatial attention. *Psychophysiology, 35,* 313–327.

Eimer, M., van Velzen, J., & Driver, J. (2004). ERP evidence for cross-modal audiovisual effects of endogenous spatial attention within hemifields. *Journal of Cognitive Neuroscience, 16,* 272–288.

Falchier, A., Clavagnier, S., Barone, P., & Kennedy, H. (2002). Anatomical evidence of multimodal integration in primate striate cortex. *Journal of Neuroscience, 22,* 5749–5759.

Fendrich, R., & Corballis, P. M. (2001). The temporal cross-capture of audition and vision. *Perception & Psychophysics, 63,* 719–725.

Giard, M. H., & Peronnet, F. (1999). Auditory–visual integration during multimodal object recognition in humans: A behavioral and electrophysiological study. *Journal of Cognitive Neuroscience, 11,* 473–490.

Gondan, M., & Roder, B. (2006). A new method for detecting interactions between the senses in event-related potentials. *Brain Research, 1073–1074,* 389–397.

Hairston, W. D., Wallace, M. T., Vaughan, J. W., Stein, B. E., Norris, J. L., & Schirillo, J. A. (2003). Visual localization ability influences cross-modal bias. *Journal of Cognitive Neuroscience, 15,* 20–29.

Hillyard, S. A., & Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proceedings of the National Academy of Sciences, U.S.A., 95,* 781–787.

Hillyard, S. A., Simpson, G. V., Woods, D. L., Van Voorhis, S., & Münte, T. F. (1984). Event-related brain potentials and selective attention to different modalities. In F. Renoso-Suarez & C. Ajmone-Marsan (Eds.), *Cortical integration* (pp. 395–414). New York: Raven Press.

Hillyard, S. A., Vogel, E. K., & Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: Electrophysiological and neuroimaging evidence. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 353,* 1257–1270.

Hopfinger, J. B., Buonocore, M. H., & Mangun, G. R. (2000). The neural mechanisms of top–down attentional control. *Nature Neuroscience, 3,* 284–291.

Hopfinger, J. B., Luck, S. J., & Hillyard, S. A. (2004). Selective attention: Electrophysiological and neuromagnetic studies. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 561–574). Cambridge, MA: MIT Press.

Jeffreys, D. A. (1968). Separable components of human evoked responses to spatially patterned visual fields. *Electroencephalography and Clinical Neurophysiology, 24,* 596.

Kastner, S., & Pinsk, M. A. (2004). Visual attention as a multilevel selection process. *Cognitive, Affective & Behavioral Neuroscience, 4,* 483–500.

Kastner, S., & Ungerleider, L. G. (2000). Mechanisms of visual attention in the human cortex. *Annual Review of Neuroscience, 23,* 315–341.

Luck, S. J., Hillyard, S. A., Mouloua, M., Woldorff, M. G., Clark, V. P., & Hawkins, H. L. (1994). Effects of spatial cuing on luminance detectability: Psychophysical and electrophysiological evidence for early selection. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 887–904.

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

Martínez, A., Di Russo, F., Anllo-Vento, L., Sereno, M. I., Buxton, R. B., & Hillyard, S. A. (2001). Putting spatial attention on the map: Timing and localization of stimulus selection processes in striate and extrastriate visual areas. *Vision Research, 41,* 1437–1457.

Martínez, A., Teder-Sälejärvi, W., & Hillyard, S. A. (2007). Spatial attention facilitates selection of illusory objects: Evidence from event-related brain potentials. *Brain Research, 1139,* 143–152.

Martínez, A., Teder-Sälejärvi, W., Vazquez, M., Molholm, S., Foxe, J. J., Javitt, D. C., et al. (2006). Objects are highlighted by spatial attention. *Journal of Cognitive Neuroscience, 18,* 298–310.

McCarthy, G., & Wood, C. C. (1985). Scalp distributions of event-related potentials: An ambiguity associated with analysis of variance models. *Electroencephalography and Clinical Neurophysiology, 62,* 203–208.

McDonald, J. J., Teder-Sälejärvi, W. A., Di Russo, F., & Hillyard, S. A. (2003). Neural substrates of perceptual enhancement by cross-modal spatial attention. *Journal of Cognitive Neuroscience, 15,* 10–19.

McDonald, J. J., Teder-Sälejärvi, W. A., Di Russo, F., & Hillyard, S. A. (2005). Neural basis of auditory-induced shifts in visual time–order perception. *Nature Neuroscience, 8,* 1197–1202.

McDonald, J. J., Teder-Sälejärvi, W. A., Heraldez, D., & Hillyard, S. A. (2001). Electrophysiological evidence for the "missing link" in crossmodal attention. *Canadian Journal of Experimental Psychology, 55,* 141–149.

Mishra, J., Martínez, A., Sejnowski, T. J., & Hillyard, S. A. (2007). Early cross-modal interactions in auditory and visual cortex underlie a sound-induced visual illusion. *Journal of Neuroscience, 27,* 4120–4131.

Molholm, S., Martínez, A., Shpaner, M., & Foxe, J. J. (2007). Object-based attention is multisensory: Co-activation of an object's representations in ignored sensory modalities. *European Journal of Neuroscience, 26,* 499–509.

Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). Multisensory auditory–visual interactions during early sensory processing in humans: A high-density electrical mapping study. *Brain Research, Cognitive Brain Research, 14,* 115–128.

Pick, H. L., Warren, D. H., & Hay, J. C. (1969). Sensory conflict in judgements of spatial direction. *Perception & Psychophysics, 6,* 203–205.

Picton, T. W., Hillyard, S. A., Krausz, H. I., & Galambos, R. (1974). Human auditory evoked potentials: I. Evaluation of components. *Electroencephalography and Clinical Neurophysiology, 36,* 179–190.

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13,* 25–42.

Recanzone, G. H. (2003). Auditory influences on visual temporal rate perception. *Journal of Neurophysiology, 89,* 1078–1093.

Reynolds, J. H., & Chelazzi, L. (2004). Attentional modulation of visual processing. *Annual Review of Neuroscience, 27,* 611–647.

Rockland, K. S., & Ojima, H. (2003). Multisensory convergence in calcarine visual areas in macaque monkey. *International Journal of Psychophysiology, 50,* 19–26.

Rorden, C., & Brett, M. (2000). Stereotaxic display of brain lesions. *Behavioral Neurology, 12,* 191–200.

Scherg, M. (1990). Fundamentals of dipole source analysis. In F. Grandori, M. Hoke, & G. L. Roman (Eds.), *Auditory evoked magnetic fields and electric potentials* (pp. 40–69). Basel: S. Karger.

Schroeder, C. E., & Foxe, J. (2005). Multisensory contributions to low-level, "unisensory" processing. *Current Opinion in Neurobiology, 15,* 454–458.

Sekuler, R., Sekuler, A. B., & Lau, R. (1997). Sound alters visual motion perception. *Nature, 385,* 308.

Shams, L., Iwaki, S., Chawla, A., & Bhattacharya, J. (2005). Early modulation of visual cortex by sound: An MEG study. *Neuroscience Letters, 378,* 76–81.

Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions. What you see is what you hear. *Nature, 408,* 788.

Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Brain Research, Cognitive Brain Research, 14,* 147–152.

Shams, L., Kamitani, Y., Thompson, S., & Shimojo, S. (2001). Sound alters visual evoked potentials in humans. *NeuroReport, 12,* 3849–3852.

Stein, B. E., London, R., Wilkinson, L. K., & Price, D. D. (1996). Enhancement of perceived visual intensity by auditory stimuli: A psychophysical analysis. *Journal of Cognitive Neuroscience, 8,* 497–506.

Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses.* Cambridge, MA: MIT Press.

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

Talsma, D., Doty, T. J., & Woldorff, M. G. (2007). Selective attention and audiovisual integration: Is attending to both modalities a prerequisite for early integration? *Cerebral Cortex, 17,* 679–690.

Talsma, D., & Kok, A. (2002). Intermodal spatial attention differs between vision and audition: An event-related potential analysis. *Psychophysiology, 39,* 689–706.

Talsma, D., Kok, A., Slagter, H. A., & Cipriani, G. (2008). Attentional orienting across the sensory modalities. *Brain & Cognition, 66,* 1–10.

Talsma, D., & Woldorff, M. G. (2005). Selective attention and multisensory integration: Multiple phases of effects on the evoked brain activity. *Journal of Cognitive Neuroscience, 17,* 1098–1114.

Teder-Sälejärvi, W. A., Di Russo, F., McDonald, J. J., & Hillyard, S. A. (2005). Effects of spatial congruity on audio-visual multimodal integration. *Journal of Cognitive Neuroscience, 17,* 1396–1409.

Teder-Sälejärvi, W. A., McDonald, J. J., Di Russo, F., & Hillyard, S. A. (2002). An analysis of audio-visual crossmodal integration by means of event-related potential (ERP) recordings. *Brain Research, Cognitive Brain Research, 14,* 106–114.

Teder-Sälejärvi, W. A., Münte, T. F., Sperlich, F., & Hillyard, S. A. (1999). Intra-modal and cross-modal spatial attention to auditory and visual stimuli. An event-related brain potential study. *Brain Research, Cognitive Brain Research, 8,* 327–343.

Vroomen, J., & De Gelder, B. (2004). Perceptual effects of cross-modal stimulation: Ventriloquism and the freezing phenomenon. In G. A. Calvert, C. Spence, & B. E. Stein (Eds.), *The handbook of multisensory processes* (pp. 141–150). Cambridge, MA: MIT Press.

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

Watkins, S., Shams, L., Tanaka, S., Haynes, J. D., & Rees, G. (2006). Sound alters activity in human V1 in association with illusory visual perception. *Neuroimage, 31,* 1247–1256.