The Rapid Distraction of Attentional Resources toward the Source of Incongruent Stimulus Input during Multisensory Conflict

Sarah E. Donohue, Alexandra E. Todisco, and Marty G. Woldorff

Abstract

■ Neuroimaging work on multisensory conflict suggests that the relevant modality receives enhanced processing in the face of incongruency. However, the degree of stimulus processing in the irrelevant modality and the temporal cascade of the attentional modulations in either the relevant or irrelevant modalities are unknown. Here, we employed an audiovisual conflict paradigm with a sensory probe in the task-irrelevant modality (vision) to gauge the attentional allocation to that modality. ERPs were recorded as participants attended to and discriminated spoken auditory letters while ignoring simultaneous bilateral visual letter stimuli that were either fully congruent, fully incongruent, or partially incongruent (one side incongruent, one congruent) with the auditory stimulation. Half of the audiovisual letter stimuli were followed 500–700 msec later by a bilateral visual probe stimulus. As expected, ERPs to the audiovisual stimuli showed an incongruency ERP effect (fully incongruent versus fully congruent) of an enhanced, centrally distributed, negative-polarity wave starting ~250 msec. More critically here, the sensory ERP components to the visual probes were larger when they followed fully incongruent versus fully congruent multisensory stimuli, with these enhancements greatest on fully incongruent trials with the slowest RTs. In addition, on the slowest-response partially incongruent trials, the P2 sensory component to the visual probes was larger contralateral to the preceding incongruent visual stimulus. These data suggest that, in response to conflicting multisensory stimulus input, the initial cognitive effect is a capture of attention by the incongruent irrelevant-modality input, pulling neural processing resources toward that modality, resulting in rapid enhancement, rather than rapid suppression, of that input. ■

INTRODUCTION

In our complex multisensory world, we are continually deluged by diverse sets of sensory input, only portions of which we can process well at any given moment. Fortunately, in many cases, there is redundancy between the input from the different sensory modalities (e.g., seeing and hearing a person speak), facilitating the linking of the multisensory stimuli into discrete objects (reviewed in Stein & Stanford, 2008). There are some instances, however, where spatially and/or temporally coincident multisensory stimuli do not contain the same information (e.g., hearing audio from a phone while driving) or where they even more directly conflict. Such conflicting input can lead to impaired behavioral outcomes (e.g., slowed RTs), impaired accuracy thereby potentially creating dangerous situations (Strayer & Johnston, 2001). One way to examine how multisensory conflict is influencing behavior is to study the dynamic modulation of the relative levels of attentional allocation to the relevant and irrelevant stimulus-input modalities in the face of such conflict.

Attention can be modulated under a wide variety of circumstances, and such modulations have been characterized by distinct neural hallmarks. Specifically, the allocation of attention to specific sensory input can be observed through increased firing rates of sensory neurons (Spitzer, Desimone, & Moran, 1988; Moran & Desimone, 1985), enhanced sensory-evoked potentials (Rugg, Milner, Lines, & Phalp, 1987; Vanvoorhis & Hillyard, 1977), and increased activity in sensory cortices as measured with functional neuroimaging (e.g., Mangun, Hopfinger, Kussmaul, Fletcher, & Heinze, 1997; Heinze et al., 1994). Importantly, in all of these cases, the enhanced stimulus processing as a function of attention is manifested by increased neural activity in the sensory cortices, or conversely, when increased neural activity is observed in those cortices during attentional manipulation, it can typically be inferred that more attention had been allocated to the processing of a given stimulus.

One major example where the allocation of attention seems to be important is when two stimulus inputs conflict, thereby necessitating the selection of that which is relevant for a behavioral goal. In traditional unimodal conflict tasks, such as the color-naming Stroop task (Stroop, 1935) or flanker task (Erikson & Erikson, 1974), the relevant dimension (e.g., the word font color in the Stroop task or the middle stimulus surrounded by distracters in the flanker task) must be selected for

Duke University

processing and response, whereas the irrelevant, distracting information must be ignored. The neural underpinnings of conflict processing is thought to involve a conflict-detection process on incongruent trials, occurring at least in part in the ACC (Fan, Hof, Guise, Fossella, & Posner, 2008; MacDonald, Cohen, Stenger, & Carter, 2000; Carter, Mintun, & Cohen, 1995), the timing of which tends to occur between around 250 and 500 msec poststimulus onset (e.g., Liotti, Woldorff, Perez, & Mayberg, 2000; West & Alain, 1999). Importantly, besides indicating a role of ACC or the nearby pre-supplementary motor area (pre-SMA) in the processing of stimulus conflict, many fMRI studies of conflict have also observed the engagement of lateral frontal regions and/or parietal regions (e.g., Silton et al., 2010; Erickson et al., 2009; Roberts & Hall, 2008; Brass, Derrfuss, & von Cramon, 2005; Egner & Hirsch, 2005b) under instances of incongruency, implicating a role of the frontoparietal attentional control network (Corbetta & Shulman, 2002) during instances of conflict. Moreover, trial-to-trial variations in the modulations of attention during various types of conflict have also been observed (e.g., Weissman, Warner, & Woldorff, 2009; Weissman, Roberts, Visscher, & Woldorff, 2006) by examining variations of neural activity as a function of RT.

Although the aforementioned studies indicate that attentional control processes are being implemented during instances of stimulus conflict, other neuroimaging investigations have taken advantage of the selectivity of visual areas for specific processing modules (faces: Egner & Hirsch, 2005a; colors and words: Erickson et al., 2009; Polk, Drake, Jonides, Smith, & Smith, 2008) to determine if, in response to conflicting stimulus input, selective attention served to enhance the relevant aspects of the stimulus or to suppress that which is irrelevant, although with somewhat mixed results. Egner and Hirsch (2005a) found enhancement in face-specific regions following conflicting multifeatureal stimuli in which faces were the relevant (target) dimension. Their design, however, focused on sequential conflict-adaptation effects (a modulation on a trial as a function of whether the previous stimulus was incongruent versus congruent; see Botvinick, Braver, Barch, Carter, & Cohen, 2001, for a review), rather that modulation of attentional allocation during the current trial. Using a traditional Stroop task in which participants attended to the font color and ignored the word meaning, Erickson and colleagues (2009) and Polk and colleagues (2008) found increased event-related fMRI responses in a color-sensitive region for the incongruent condition relative to neutral and/or congruent trial types (e.g., neutral: Polk et al., 2008; or congruent: Erickson et al., 2009). Polk and colleagues (2008) also found evidence in the Stroop task for suppression of word-processing areas in incongruent versus neutral trials in visual word form areas. In contrast, however, Erickson and colleagues (2009) found that participants with greater fMRI activity in the visual word form area had faster RTs, suggesting an opposite pattern. Moreover, and in addition to these seemingly conflicting results, none of these studies were able to look at the timing of the attentional modulations in the brain because of the sluggish response of the fMRI BOLD signals being measured (Logothetis & Wandell, 2004), leaving the temporal cascade of these within-vision attentional allocation effects during conflict as an open question.

The study of conflict across modalities has been examined to a much lesser degree. However, such study can provide a unique opportunity to selectively assess the relative allocation of attentional resources to both the relevant and irrelevant modalities, in that the major facets of the sensory processing of the stimuli take place in distinct cortical locations with separate temporal profiles (but see Kayser, Petkov, & Logothetis, 2008; Smiley et al., 2007, for evidence of responses to visual [Kayser et al., 2008] and somatosensory [Smiley et al., 2007] stimuli in primary auditory cortex). Weissman, Warner, and Woldorff (2004) found that the occurrence of incongruent audiovisual stimuli led to enhanced fMRI activity in the sensory cortices of the relevant (attended) modality. Furthermore, in a later study with a similar audiovisual conflict paradigm (Weissman et al., 2009), increased activity in the sensory cortices of the irrelevant modality was also observed, regardless of the congruency of the stimuli, on trials to which participants were slow to respond (which were interpreted as trials with a relative waning or lapse of attention). Although these intriguing findings suggest that attention may both serve to enhance processing in the relevant modality in the face of conflict and also enhance processing in the irrelevant modality during momentary lapses, they do not speak to the timing of this variation in attentional allocation during a given trial, nor whether there is differential attentional allocation to the irrelevant modality as a function of congruency. In the current multisensory-conflict study, we sought to bridge this gap with the higher temporal resolution of EEG applied to a novel probe stimulus approach that could directly gauge the amount of neural resources allocated to the irrelevant modality under conditions of high and low conflict.

Previous studies have successfully used a visual probe stimulus to characterize attentional distributions, such as the spatial profile of the attentional spotlight during focused visual attention (e.g., Hopf et al., 2006). Here, we used EEG and visual probes to examine the activity in the sensory cortex of the irrelevant modality (vision) shortly following the presentation of multisensory (audiovisual) stimuli with different conflict characteristics. Participants attended auditorily to discriminate between two spoken letters. Bilateral visual letter stimuli were simultaneously presented with the auditory stimulus that could be fully congruent with it (both sides congruent with the spoken letter), fully incongruent (both sides incongruent), or partially incongruent (incongruent on one side and congruent on the other). On half of the

trials, 500–700 msec after the multisensory stimulus, a bilateral visual probe stimulus was presented in the same location as the visual component of the multisensory stimulus to assess the allocation of attention to the visual modality. We examined time-locked averages to both the multisensory stimuli and the probes, looking for differential modulations in the sensory processing of the probes (reflecting the amount of attention directed toward the visual modality) as a function of the incongruency characteristics of the preceding multisensory stimulus.

There were several key predictions. First, in accordance with previous studies of stimulus conflict, we expected an increased fronto-central negative-polarity ERP peaking at ~350 msec that was greater for fully incongruent versus fully congruent trials (e.g., Atkinson, Drysdale, & Fulham, 2003; Liotti et al., 2000; West & Alain, 1999), with the response to the partially incongruent trials falling between the two. Second, we expected that the sensory-evoked visual sensory activity to the probe stimulus would be modulated by the nature of the previous multisensory stimulus (fully and partially incongruent versus fully congruent) and the attentional processes that it triggered. If, to facilitate task performance, attention were rapidly directed toward the relevant modality and away from the irrelevant one when it conflicted, then the sensory response to the irrelevant-modality visual probe should decrease. Alternatively, it is possible that attention could be drawn to the irrelevant visual modality when its content conflicted, resulting in enhanced sensory processing of the visual probe. This prediction would be in line with findings showing enhanced within-trial processing of incongruent flanker stimuli in a visual-flanker study recently reported by our group (Appelbaum, Smith, Boehler, Chen, & Woldorff, 2011). Additionally, the inclusion of partially incongruent trials (incongruent on only one side) was aimed at taking advantage of the contralaterality of the visual system. In these instances, we hypothesized that the sensory response to the probe might be modulated in a lateralized manner (e.g., modulated mainly on one side or the other), thereby adding further insight into the rapid responses of the brain to multisensory conflict.

METHODS

Participants

Twenty-seven healthy participants from the Duke University community and surrounding areas were included in this study (10 women, 2 left-handed, mean age = 22.3 years, *SD* = 4.4 years). Six additional participants were excluded because of excess physiological noise in their data. All participants gave informed consent and were financially compensated for their time. All procedures were approved by the Duke University Health System Internal Review Board.

Stimuli and Task

Participants were seated 57 cm from a CRT monitor on which a central fixation cross was presented for the duration of the experiment. Auditory stimuli, the spoken letter "X" or the spoken letter "O," with durations of ~335 msec, were centrally presented at 20 dBSL through speakers. These were presented a relatively low volume in order make them somewhat more vulnerable to distraction from other input. Bilateral visual stimuli were presented 5 degrees below and 5 degrees to the right and left of the central fixation cross (see Figure 1), each for a duration of 50 msec.

Each trial began with a spoken letter ("X" or "O") and a simultaneously presented, bilateral, visual stimulus that could be either fully congruent with the spoken letter (spoken "X" with visual X X or spoken "O" with visual O O), fully incongruent with the spoken letter (spoken "X" with visual O O or spoken "O" with visual X X), or partially incongruent (left incongruent: spoken "X" with visual O X or spoken "O" with visual X O; right incongruent: spoken "X" with visual X O or spoken "O" with visual O X). Participants were instructed to attend to the auditory stimuli and to ignore the visual as well as to press one button as rapidly as possible if they heard the letter "X" and another if they heard the letter "O." Response buttons were counterbalanced across participants.

Additionally, on half of the above trials, a neutral bilateral visual probe stimulus (H H) was presented (duration, 50 msec) that onset between 500 and 700 msec following the onset of the multisensory stimulus. On the other half of the trials, no probe was presented, providing a trial-type contrast for extracting the visual-probe evoked response. The trial-onset asynchrony (TOA) was jittered between 950 and 1550 msec. Fifteen blocks of trials were presented, each consisting of 160 trials and lasting 3.3 min, yielding 300 total trials in each condition (e.g., fully incongruent multisensory followed by a probe).

Behavioral Analysis

For each participant and each condition, trials that had RTs on the letter discrimination task that were greater than or less than two standard deviations from the mean of that condition were excluded from analysis. Each of the proportion-correct values in the four main multisensory incongruency conditions (congruent, incongruent, incongruent left, and incongruent right) were entered into a 4 (incongruency) × 2 (probe vs. no probe) two-factor ANOVA.

EEG Recording and Analysis

Using an ANT (Advanced Neuro Technology, The Netherlands) acquisition system, continuous 128-channel EEG data were acquired during the experimental runs, with an on-line average-channel reference. All impedances were maintained below 5 kΩ, and the data were sampled

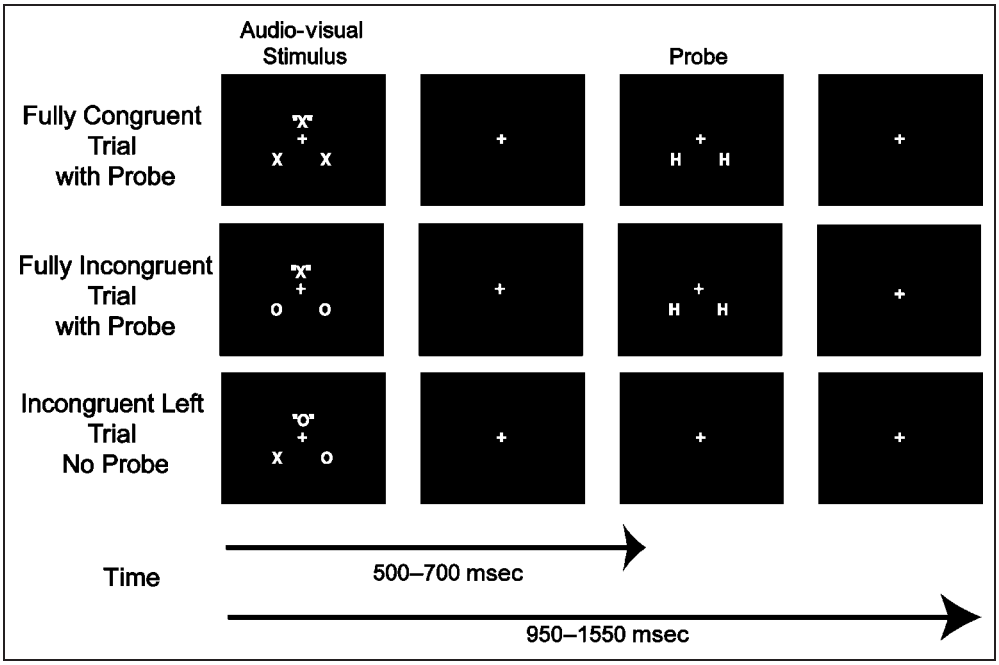

Figure 1. Task depiction. Three example trials are shown. A centrally presented, task-relevant, auditory spoken letter (depicted in quotation marks) was presented while task-irrelevant visual letter stimuli were simultaneously presented bilaterally below fixation. For all trials, the task was to respond to the auditory letter stimulus and to ignore the visual, which was randomly either fully congruent or fully incongruent with the auditory input on both sides or congruent on one side and incongruent on the other ("partially incongruent"). On half of the trials, a visual probe stimulus (see the fully congruent and fully incongruent example trials in the figure) was presented shortly following (jittered onset from 500 to 700 msec post-trial onset) consisting of the letter H, bilaterally presented for 50 msec. For the other half of the trials, no probe stimulus was presented. In all cases, the central fixation cross remained on the screen for the entire duration of the trial.

at 512 Hz per channel. Horizontal and vertical electro-ocular recordings (HEOG and VEOG) were obtained using additional external electrodes placed above and below the right orbit and on the left and right canthi, respectively. Offline, the data were bandpass filtered from 0.01 to 30 Hz and rereferenced to the average of the left and right mastoids. Trials with eye blinks and movements were rejected, along with trials containing excess muscle activity and/or slow drift. Time-locked averages were obtained for the onset of each trail and sorted by the various conditions (multisensory congruency condition, probe vs. no probe, etc.), with the analyses collapsed across X and O stimuli. Additional time-locked averages were obtained for the onset of the visual probe stimuli. Subsequently, difference waves between the conditions were calculated, with specific subtractions described in the results below. For all plots and statistics, the stimuli were baselined from 200 to 0 msec before the relevant stimulus, and all statistical results were Greenhouse–Geisser corrected, where applicable.

RESULTS

Behavior

Participants had a high proportion correct across all conditions (Figure 2A), averaging 96.3% correct. More importantly here, however, there was an effect of multisensory incongruency on accuracy, $F(3, 78) = 4.48$, $p = .006$. Neither a main effect of Probe Presence on Accuracy nor an interaction of Probe Presence and Incongruency were observed (all ps $> .1$). Planned comparisons revealed that the effect of multisensory incongruency was driven by participants being significantly more accurate in the fully congruent condition than in the fully incongruent, incongruent right, and incongruent left conditions (fully congruent vs. fully incongruent, $M = 0.969$ vs. 0.959, $t(26) = 3.17$, $p = .004$; fully congruent vs. incongruent right, $M = 0.969$ vs. 0.961, $t(26) = 2.80$, $p = .01$; fully congruent vs. incongruent left, $M = 0.969$ vs. 0.963, $t(26) = 2.49$, $p = .02$). No other differences in accuracy across the conditions were observed (all ps $> .1$).

The incongruency of the multisensory stimuli also had an effect on RTs (Figure 2B). As was done with the accuracy, the RT data were analyzed using a 4×2 two-way ANOVA, with the factors of Incongruency and Probe Presence. A main effect of Incongruency on RT was found, $F(3, 78) = 10.33$, $p < .001$, but, again, neither a main effect of Probe Presence nor an interaction between Probe Presence and Incongruency were observed (p values $> .1$). Follow-up t tests found that fully incongruent RTs were significantly slower than fully congruent (fully incongruent $M = 430.1$ msec vs. fully congruent $M = 417.6$ msec; $t(26) = 6.40$; $p < .001$), as were incongruent right (incongruent right $M = 423.1$; $t(26) = 2.79$; $p = .01$).

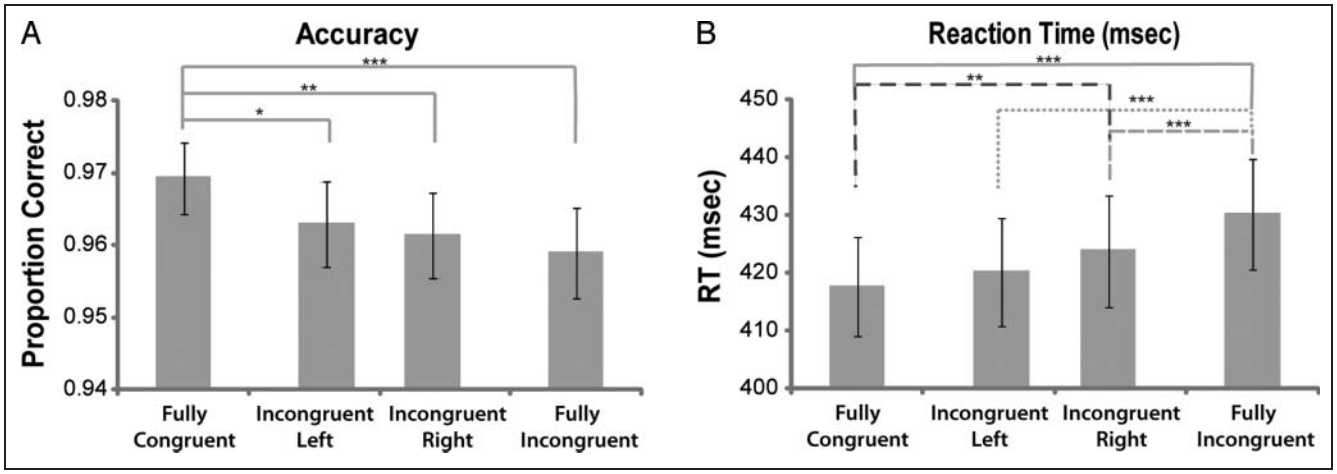

Figure 2. Behavioral data. (A) Accuracy. The proportion correct for the auditory spoken letter discrimination is plotted as a function of trial type. For each condition, proportion correct responses are collapsed across probe and no probe trials. Participants were highly accurate across all conditions, with the congruent condition being the most accurate. (B) RTs. Participants were slowest to respond to the fully incongruent trials and fastest to respond to fully congruent trials. For both figures, asterisks denote the level of significance: *$p < .05$, **$p < .01$, ***$p < .005$, ****$p < .001$.

Additionally, fully incongruent RTs were significantly slower than both incongruent right trials ($t(26) = 3.30$, $p = .003$) and incongruent left trials ($M = 420.1$; $t(26) = 3.86$, $p = .001$). No other significant differences between conditions were observed. Together, these data suggest that the manipulation of incongruency had a clear effect on the participants' accuracy and RTs, with the fully incongruent trials causing the greatest behavioral decrement.

ERPs

Main Effects of Incongruency

Because there were no significant behavioral differences between the incongruent left and incongruent right trials, and because we had no a prior hypotheses that these would differ for the main effect of incongruency, we collapsed across these two conditions for this analysis. Time-locked averages to the multisensory stimulus onset were obtained for fully congruent, fully incongruent, and partially incongruent trial types (collapsed across both probe and no-probe conditions). A fronto-central negative-polarity ERP effect of incongruency was observed from 250 to 450 msec post-trial onset (Figure 3). Specifically, there was a greater negativity for the fully incongruent condition compared with the fully congruent condition, with the partially incongruent condition eliciting a level of activity that fell between these two. A repeated-measures ANOVA to examine this effect was conducted using Incongruency (fully congruent, fully incongruent, partially incongruent) as a factor. This analysis was applied to the ERP amplitudes averaged across the midline site Cz[1] and the four immediately adjacent sites in all directions (see Figure 3) in 100 msec windows (250–350 msec, 350–450 msec). For 250–350 msec, there was a robust main effect of Incongruency, $F(2, 52) = 10.44$, $p < .001$. Within this time period, all three conditions were found to be significantly different from each other, with the fully incongruent being the most negative, followed by the partially incongruent, followed by the fully congruent (fully incongruent vs. partially incongruent: $t(26) = 2.76$,

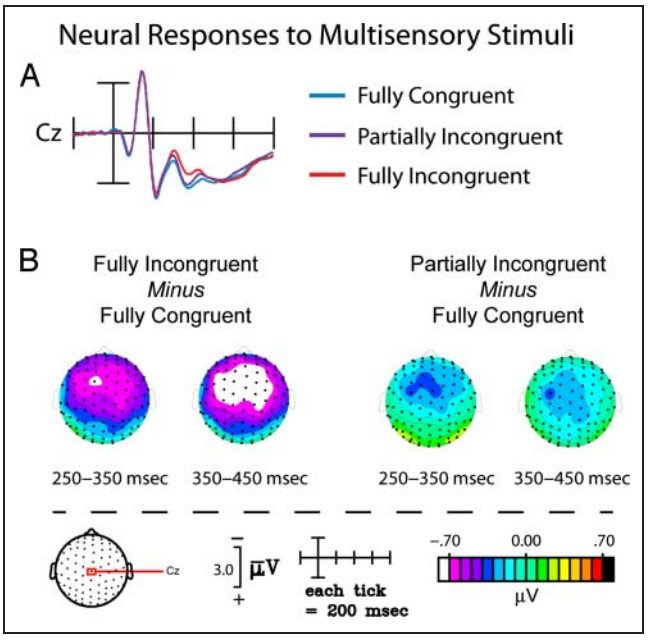

Figure 3. Incongruency neural effects. (A) ERP traces at site Cz showing the fully incongruent, partially incongruent, and fully congruent waveforms. An increased negative-polarity ERP waves is present for the incongruent compared with the congruent condition starting at around 250 msec, with the partially incongruent trials falling in the middle. (B) Topographic distributions plotted for the fully incongruent minus fully congruent trial types (left) and for the partially incongruent minus fully congruent trial types (right). The greatest ERP negativity is observed for the fully incongruent minus fully congruent, whereas the partially incongruent conditions had both a decrease in the duration and amount of this negative incongruency-related effect.

$p = .01$; fully congruent vs. partially incongruent: $t(26) = 2.19$, $p = .04$; fully congruent vs. fully incongruent: $t(26) = 3.93$, $p = .001$). From 350 to 450 msec, there was also a significant main effect of Congruency, $F(2, 52) = 8.31$, $p = .002$. Subsequent analyses revealed that the fully incongruent and fully congruent significantly differed ($t(26) = 3.25$, $p = .003$), as did the fully incongruent and the partially incongruent ($t(26) = 3.56$, $p = .002$), with the partially incongruent and fully congruent trial types not differing during this time window ($t(26) = 1.40$, $p = .18$). Thus, the fully incongruent trial types showed the greatest negativity, which differed from the fully congruent trial types over a 200-msec period, whereas the partially incongruent trial types fell between the two from 250–350 msec, with the differential effect relative to the fully congruent trial type ending sooner (i.e., by around 350 msec).

Effects on the Sensory Response to the Probe Stimulus

To assess the responses in visual cortex to simple stimuli as a function of whether or not the visual component of the preceding multisensory stimulus had been fully congruent, fully incongruent, or partially incongruent with the task-relevant auditory component, we collected time-locked averages to the onset of the visual probe. On trials for which no probe occurred, time-locked averages were also obtained to the times at which a probe could have occurred (i.e., at similarly random points jittered between 500 and 700 msec post-trial onset). By then subtracting the ERP responses to the trials with no probe from those with a probe, separately for the different preceding multisensory congruency trial types, the event-related responses to the visual probes could be extracted for each of the different post-multisensory-conflict context conditions. This subtraction removes the overlapping activity of the processing of the preceding multisensory stimulus in the trials, along with any motor or categorization-related activity associated with the letter-discrimination task, leaving the sensory-evoked ERP responses to the probe stimuli under the different multisensory context conditions. It should additionally be noted that any of such late activity is likely to have different neural generators than the sensory-evoked responses to the probe of interest and were thus unlikely to lead to any neural interactions. The magnitude of these sensory responses (Figure 4) could be evaluated to assess the amount of attention allocated to the task-irrelevant visual modality under these different conditions. Specifically, because the visual probe stimulus was the same physical stimulus across all conditions, any differences in the sensory response to it would presumably be the result of the amount of attentional allocation to the visual modality as a function of the previous multisensory trial type (e.g., fully incongruent vs. fully congruent). These analyses revealed two differential effects on early sensory components to the probe (Figure 4). The first was a small effect on the frontal N1 component, and the second was a larger, posterior (parietal) effect on the P2. No effects were observed on the P1 or the posterior N1 components, and these were therefore not analyzed further.

The mean amplitude from the frontal N1 component (Figure 4A) was first analyzed using a repeated-measures ANOVA with the factor of Incongruency (fully congruent, fully incongruent, incongruent left, and incongruent right), conducted on five midline sites between ~Cz and ~FpFz (and the five channels immediately to the left and right of the midline, averaged together) from 110 to 190 msec following probe onset. This analysis revealed a main effect of Preceding Multisensory Incongruency, $F(3, 78) = 3.04$, $p = .03$, on the N1 component. Subsequent specific comparisons showed that a visual probe following a fully incongruent multisensory stimulus had a marginally greater ERP negativity in the N1 latency range than following a fully congruent one ($t(26) = 1.86$, $p = .07$). The differences between incongruent right and incongruent left followed a similar pattern to that of fully incongruent stimuli, with both showing a greater N1 than the fully congruent condition (incongruent right vs. fully congruent: $t(26) = 2.89$, $p = .008$; incongruent left vs. fully congruent: $t(26) = 2.29$, $p = .03$). No other significant differences on the frontal N1 to the probe as a function of the incongruency characteristics of the preceding multisensory stimulus were observed.

The effect on the P2 sensory-evoked response (Figure 4B) was also analyzed (from 175 to 325 msec post-probe onset) to determine if this component was also modulated as a function of the incongruency in the multisensory target stimulus within the trial. A repeated-measures ANOVA with the four levels of Congruency was run on POz, Oz, and the two immediately adjacent lateral sites averaged together (see Figure 4). This analysis also revealed a particularly large and robust main effect of Incongruency, $F(3, 78) = 10.29$, $p < .001$. Subsequent analyses showed that the probe P2 for all conditions with any incongruency (fully incongruent, incongruent right, and incongruent left) in the preceding target stimulus were significantly larger than for the fully congruent conditions (fully incongruent vs. fully congruent: $t(26) = 4.95$, $p < .001$; incongruent right vs. fully congruent: $t(26) = 3.13$, $p = .004$; incongruent left vs. fully congruent: $t(26) = 3.50$, $p = .002$). There was also a marginal difference between fully incongruent and incongruent right ($t(26) = 1.98$, $p = .06$), with the P2 activity in response to the probe being more positive when it followed a fully incongruent compared with an incongruent right multisensory stimulus. Taken together, the sensory response to the visual probe stimuli clearly varied as a function of the incongruency character of the preceding multisensory stimulus, with the greatest probe response following a fully incongruent multisensory stimulus.

Brain Activity as a Function of RT

In an additional analysis, we sought to determine if the enhanced probe activity that was observed in the N1

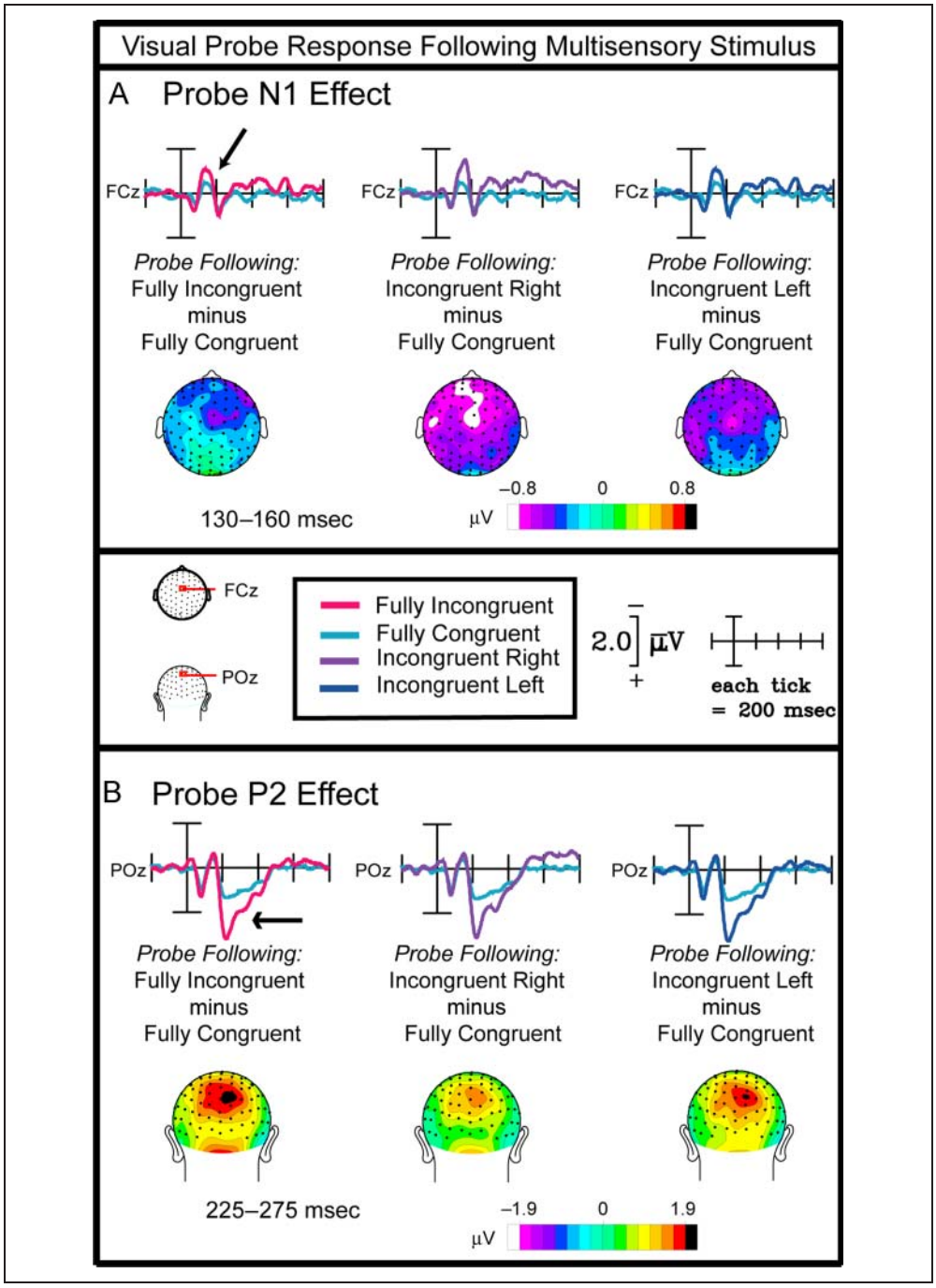

Figure 4. Sensory responses to the visual probes. (A) The three traces from left to right show the anterior responses to the visual probes following the multisensory stimuli that were fully incongruent versus fully congruent, incongruent right versus fully congruent, and incongruent left versus fully congruent at site FCz. An effect on the early anterior N1 component was found to be enhanced for probes following a fully incongruent, incongruent left, and incongruent right versus a fully congruent multisensory stimulus. Topographic distributions depict the difference waves of this effect (e.g., left topo: Probe response following fully incongruent minus probe response following fully congruent) from 130 to 160 msec and to have an anterior focus. (B) The three traces from left to right show the posterior responses to the visual probes following multisensory stimulus that were fully incongruent versus fully congruent, incongruent right versus fully congruent, and incongruent left versus fully congruent at site POz. The topographic distributions reflect the respective difference waves (e.g., left topo: Probe response following fully incongruent minus probe response following fully congruent). Here, a large P2 response can be found for the probe in the incongruent, incongruent right, and incongruent left conditions. This response is located over occipito-parietal regions, as can be seen in the topographic distributions, plotted from 225 to 275 msec.

and P2 responses following an incongruent versus a congruent multisensory stimulus also varied within the incongruency condition as a function of RT. Accordingly, for each condition and for each participant, trials were divided into thirds according to the RT. As such, the top third (i.e., the fastest trials in a given condition) and the bottom third (i.e., the slowest trials in a given condition) were obtained. Within the sets of fastest and slowest trials, respectively, the ERP response on trials without a probe (time-locked to when a probe would have occurred) were subtracted from the probe-locked response on trials with a probe, leaving the sensory responses to the probe separated as a function of response speed with that incongruency trial type. Figure 5 shows the fastest versus slowest trials for the incongruent condition, highlighting the N1 and P2 responses to the probe stimuli following the fully incongruent multisensory stimuli. As can be seen, the trials with the slower RTs (i.e., presumably those for which the incongruency of the preceding multisensory stimuli had the greatest effect) showed a greater negativity for the N1 and a greater positivity for the P2 than the faster trials. Specifically, the N1 effect was tested on the same channels as above and was found to be significantly larger for the slow

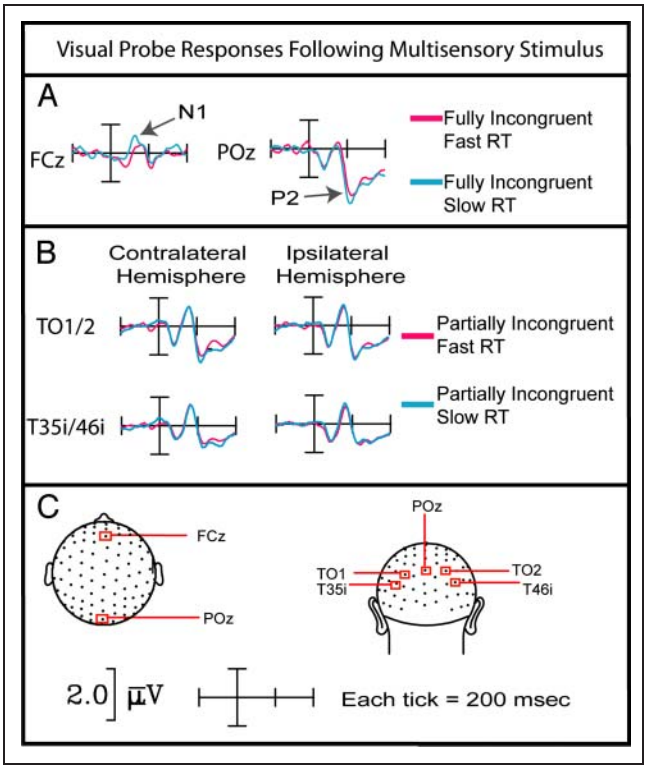

Figure 5. Brain activity as a function of RT. (A) The trials of fully incongruent condition were binned into thirds for each participant as a function of the RTs. The responses to the visual probes are plotted separately for the third slowest and third fastest trials in the fully incongruent condition. The trials in which participants were slow to make their response (i.e., those for which they had the biggest slowing effect of incongruency) also had the biggest visual responses to the probe as indicated by the differential N1 and P2 responses to the same physical stimulus. (B) The plots show the probe responses for the incongruent left and incongruent right trials for which the RTs fell in either the slowest third or the fastest third of trials in those conditions. Here, the data are collapsed across conditions such that the contralateral activity was averaged as was the ipsilateral activity for these parietal–occipital sites. A differential response pattern of enhanced activity to the P2 response is observed for the slow trials compared with the fast trials over the hemisphere contralateral to the preceding incongruency, whereas no difference was observed over the hemisphere ipsilateral to the incongruency. Specifically, an enhanced P2 response to the probes was observed for the slow trials compared with the fast trials on the side contralateral to the side of the incongruent part of the multisensory target stimulus. (C) Schematic of the location of the electrodes shown above in A and B.

versus the fast trials ($t(26) = 2.05$, $p = .05$) from 100 to 130 msec. Moreover, applying the same slow versus fast RT analysis to the fully congruent trials, the N1 effect was found to be not modulated by speed ($p = .83$). The P2 effect, on channels POz, Oz, and the two immediately adjacent lateral electrodes from 175 to 325 msec, also showed a significant difference for fast versus slow fully incongruent trials ($t(26) = 2.28$, $p = .03$; see Figure 5 for channel locations), with a larger probe P2 following the fully incongruent stimuli with the slowest RTs.

Another interesting pattern of findings emerged when looking at the fastest versus slowest trials for the incongruent right and incongruent left conditions. Here, the slowest trials showed a greater P2 response, with this enhancement in both cases being contralateral to the side of the incongruent visual stimulus. That is, for partially incongruent slow RT trials in which the multisensory target stimulus was incongruent on the right, the greater P2 positivity emerged over the left occipital–parietal sites, and for the incongruent left trials, the greater positivity emerged over the right occipital–parietal sites (Figure 5B). To increase the signal-to-noise ratio for this analysis, the data for the incongruent left condition were flipped so that the activity that normally appeared on the left was on the right. These flipped data were then averaged with the incongruent right data so that the contralateral data now appeared on the left for both conditions and the ipsilateral data appeared on the right for both conditions (Figure 5C). These data were the entered into a two-factor (RT: fast vs. slow × Hemisphere: contralateral vs. ipsilateral) repeated-measures ANOVA at the adjacent sites ~P4i, TO2, T46i (right hemisphere) and P3i, TO1, T35i (left hemisphere) for 175–325 msec. This analysis revealed a significant interaction of RT × Hemisphere ($F(1, 26) = 4.76$, $p = .04$), reflecting this contralaterally larger P2 enhancement for slow partially incongruent trials.

DISCUSSION

The current study used an auditory-attention multisensory-conflict probe paradigm to test for modulations of processing in the task-irrelevant modality (vision) as a function of between-modality stimulus incongruency. In terms of overall incongruency effects, behavioral decrements in stimulus discrimination performance were found when the task-relevant auditory stimulus occurred with a simultaneous fully incongruent visual stimulus, reflected by slower RTs and decreased accuracy, as compared with when the multisensory stimulation was fully congruent. The RTs for the partially incongruent trials (incongruent on one side, congruent on the other) fell between those of the fully incongruent and fully congruent trials. The electrophysiological activity elicited by the multisensory stimuli paralleled the behavior, showing an increased fronto-central negativity starting around ~250 msec and lasting until ~450 msec post-multisensory stimulus onset for the fully incongruent relative to the fully congruent trials, with the response to the partially incongruent trials again falling in between those levels and lasting for a shorter time period (until ~350 msec). Thus, in line with previous multisensory studies (e.g., Molholm, Ritter, Javitt, & Foxe, 2004), we show that the current paradigm elicited a negative-polarity neural response to cross-modal conflict, similar to the evoked negative-polarity ERP that has been modeled as arising from ACC in cases of unimodal conflict (e.g., Liotti et al., 2000).

Our main goal here, however, was to use a sensory probe technique to assess for differential allocation of attention to the task-irrelevant visual modality following the multisensory target stimulus as a function of its intermodal incongruency characteristics. To do this, we extracted the ERP responses to visual probe stimuli occurring 500–700 msec after the different multisenosory target stimulus types. The pattern of neural responses to the visual probes, as indexed by larger N1 and P2 sensory-evoked visual ERP responses, strongly suggests that more attention was drawn toward the irrelevant modality when the visual component of the preceding multisensory stimulus was incongruent versus congruent with the task-relevant auditory component. Further, a comparison of the responses to probes following incongruent trials with slower versus faster RTs showed particularly increased N1 and P2 responses for the slower trials. Finally, on slower, relative to faster, partially incongruent trials, the probes elicited a greater P2 sensory component contralateral to the side of the preceding incongruent stimulus. Thus, multiple facets of the results converge to support the view that when input from a task-irrelevant stimulus modality conflicts with task-relevant stimulus input within another modality, the initial response of the brain is for attention to be rapidly drawn toward the irrelevant modality, which then appears to serve as a key mechanism underlying the associated decrement in behavioral task performance. Moreover, the current results strongly suggest that this rapid attentional capture by the source of the incongruent stimulus in the task-irrelevant modality results in a rapid and distracting attention enhancement of the processing of that irrelevant-modality input, rather than a rapid performance-enhancing attentional suppression of the incongruent stimulus input.

Although participants were instructed to attend selectively to the auditory modality and to ignore the visual one, we still found behavioral and neural interference effects induced by the task-irrelevant visual input on incongruent trials. As such, the presence of such conflict might be a bit surprising given that selective attention could potentially act to heavily filter out the irrelevant information in one modality (something which might harder to do in a unimodal Stroop task, because of the typical spatial overlap of the stimuli). However, our evidence is not the only evidence that selective attention is often unable to completely operate unimodally by filtering out distracting information from a different modality, in that other studies, including from our group, have found incongruency effects (behavioral and neural) for conflicting multisensory stimuli (e.g., Bendixen et al., 2010; Zimmer, Itthipanyanan, Grent-'T-Jong, & Woldorff, 2010; Zimmer, Roberts, Harshbarger, & Woldorff, 2010; Weissman et al., 2004, 2009; Molholm et al., 2004). One of the more novel aspects of this study is that the degree for such interference effects was specifically modulated by the amount of interference present, with the effects for the partially incongruent stimuli falling—behaviorally and neurally—between the fully congruent and fully incongruent stimuli. This evidence suggests that the conflict-related decrement in behavioral performance (along with the corresponding neural hallmarks of conflict) are directly related to the proportion of the irrelevant stimulus input the interfering information occupies, with more interference present leading to greater conflict (see also Appelbaum et al., 2011).

In terms of overall incongruency effects, a separate notable aspect of the current results is the timing of the first component of detection of conflict in the multisensory stimuli (the enhanced central negativity at ~250 msec for incongruent relative to congruent stimuli). Most studies of this sort of incongruency effect have been carried out unimodally within vision (e.g., the classic Stroop task) and have typically found this incongurency-related ERP component to start at a substantially longer latency (e.g., ~350 msec; Markela-Lerenc et al., 2004; Liotti et al., 2000; West & Alain, 1999). Our observation of this effect starting at ~250 msec post-stimulus onset suggests that the incongruent stimulus input is detected more quickly when the incongruency is between modalities, at least when the relevant modality is auditory and the irrelevant one visual. The timing of this incongruency-related effect is more similar to what we have previously observed with a within-auditory Stroop task (Donohue, Liotti, Perez, & Woldorff, 2012), suggesting that the rapidity of conflict detection may be related to the relevant attended stimuli being auditory. It is also possible that the simple nature of the visual stimuli used here could have aided in the rapidity of the multisensory conflict detection. The only letters used here were the letter "X" and the letter "O," and given that these have very unique physical forms, it is possible that information gathered in early visual areas would be very fast, facilitating the detection of a match versus nonmatch to the rapidly processed auditory stimuli (in contrast to an entire color word in the classic color-naming Stroop task). Another possibility for the early onset of these effects could be because there were only two response options here (as compared with the four choices typically used in a Stroop task). Appelbaum et al. (2011) also found an earlier onset (~200 msec) of this component in a two-response choice flanker task. Although the simplicity of the multisensory stimuli and/or there being only two response options may be related to the early onset of this component, further work needs to be done using EEG to examine the temporal characteristics of multisensory conflict processing more generally, including determining whether the timing of this conflict detection is consistently earlier than with unimodal visual conflict.

The novel use of a visual probe stimulus in the current study provided a gauge of the processing of stimuli in the task-irrelevant modality 500–700 msec post-multisensory stimulus onset. The finding of greater sensory responses to the probe when it followed an incongruent multisensory stimulus as compared with a congruent one suggests that, at least at that point in the

trial, attention resources had been pulled more toward the irrelevant modality when it had contained an incongruent, and therefore presumably more distracting, stimulus component with respect to the target stimulus in the relevant modality. Moreover, this increased visual-probe response following an incongruent stimulus was particularly large for the most slowed incongruent trials, further supporting that this effect is reflecting performance-impairing distraction. This result is thus in line with a recent within-vision finding from our group (Appelbaum et al., 2011) of the presence of increased activity contralateral to the side of a distracting flanker stimulus which was greatest for participants with the greatest incongruency-related task performance impairment. Importantly, however, this study was able to directly gauge the sensory processing with the task-irrelevant modality through the sensory-probe technique, thus going beyond the somewhat more indirect approach of correlations across participants used in our previous study.

Although the high temporal resolution of the ERPs allowed us to determine that this relative enhancement of visual processing occurred at 500–700 msec post-stimulus onset, we did not probe at various other time points. Thus, we were unable to determine whether attention was pulled more toward the irrelevant modality throughout the duration of the trial, or whether there was a shift toward suppression of distracting input later on in the cascade of temporal processes. Because probing around and before the RT can interfere with the RTs (as indicated in some initial piloting for this study), we decided against probing earlier to look at the attentional modulation in the visual modality and using a probe for this may not be the most ideal technique for studies along these lines.

Although probing earlier in the trial may not be a fruitful endeavor, follow-up studies probing later time periods in the trial to more fully determine the time course of the attraction of attention to the irrelevant modality could be very worthwhile. When Weissman and colleagues (2004) examined fMRI activity in the sensory corticies of the irrelevant modality in a cross-modal conflict paradigm, they observed no difference in activation as a function of congruency. In contrast, in the present electrophysiological study, we did observe an enhancement within the irrelevant modality from 500 to 700 msec. To reconcile these findings, therefore, probing at later points in the trial could reveal suppression of the irrelevant modality, which, when averaged together with our observed early enhancement, would yield a net effect of no enhancement or suppression in an fMRI measure. Furthermore, it is possible that suppression of activity in the irrelevant modality could indeed happen for a longer duration of time during the latter part of the trial, thereby giving rising to the effects seen with the slower temporal resolution and more temporally integrating measures of fMRI. Indeed, two fMRI studies have found enhanced processing in the STS (van Atteveldt, Blau, Blomert, & Goebel, 2010; van Atteveldt, Formisano, Goebel, & Blomert, 2004) for congruent relative to incongruent trials. Accordingly, at some later point in the trial, we hypothesize that there may be some relative suppression of processing of the incongruent stimuli relative to congruent ones, following the rapid enhanced processing observed here.

The modulations of the visual-probe response were evidenced in both the N1 and P2 sensory components. Perhaps a bit surprisingly, there was no difference in the P1 for the probe following the congruent or incongruent multisensory stimuli. The P1 has often been a hallmark of visual attentional modulations, with classic studies of attention finding enhancements at this time period (e.g., Woldorff et al., 1997; Heinze et al., 1994; Vanvoorhis & Hillyard, 1977). However, a study of visual attention observed that components such as the P1 and N1 are affected by the spatial gradients of attention with decreasing modulations of the P1 and N1 as a function of spatial distance from the focus of attention (Mangun & Hillyard, 1988). It is possible that because the visual stimuli in the current paradigm were not directly in the center of the screen (i.e., the spatial location of the auditory stimuli), we may have been less likely to find effects on these components. Furthermore, given that attention was directed toward the auditory system on every trial, these visual modulations come in the form of intermodal attention and may not show the same type of significant modulation of the P1 that attentionally demanding unimodal visual tasks can elicit (O'Donnell, Swearer, Smith, Hokama, & McCarley, 1997).

The N1 modulation of the probe observed here was only found in the anterior N1 with no differential effects for the posterior N1. The lack of a posterior N1 effect may have been because of our bilateral stimulus presentation, which has been shown to reduce posterior N1 attentional effects (Luck, Heinze, Mangun, & Hillyard, 1990). The anterior N1 response was greater to probe stimuli when they followed incongruent multisensory stimuli (marginally significantly for the fully incongruent condition), and within the fully incongruent condition, the anterior N1 response was greatest on trials for which participants were slowest to respond. Although the anterior N1, because of its early timing (<200 msec), likely reflects aspects of sensory processing (also see: Clark & Hillyard, 1996; Mangun & Hillyard, 1988; reviewed in Luck, 1995), it is also possible that this component may have additional contributions from higher-levels of cognitive processing. Furthermore, other studies have demonstrated intermodal attention effects on the anterior visual N1, suggesting that this component may be sensitive to attentional allocation over modalities (De Ruiter, Kok, & van der Schoot, 1998; Eimer & Schroger, 1998).

The P2 component, which showed the largest modulation because of preceding incongruency, also showed a telling interaction with the speed of behavioral task response to the multisensory stimulus. In particular, for the incongruent trials that were slower and presumably

involved more distraction for the participants (see Weissman et al., 2006, 2009), the P2 to the visual probe was particularly large, further suggesting that attention was being pulled particularly strongly toward the visual modality. The partially incongruent trials shed yet more light on the attentional modulations in that on the slowest trials the increase in the probe P2 component was particularly strong on the side contralateral to the side of the preceding incongruent visual stimulus that preceded it within the trial. That is, in those cases where participants were slowest to respond, more visual attention would appear to have been captured to the visual field location of the incongruent stimulus, thereby presumably creating a yet even more salient representation of the conflicting stimulus to compete with the relevant information and thus thereby inducing the largest behavioral performance impairment. Interestingly, using a visual target discrimination task where participants saw a display with one target and one distractor, Hickey, Di Lollo, and McDonald (2009) found a contralateral ERP component, which they named the P_D, that they believed reflected the suppression of distracters. Further work still needs to be done to determine if the P2 effects observed here might reflect related cognitive mechanisms.

Finally, it may be worthwhile to contrast the current multisensory-conflict attentional capture effects to other forms of attentional capture, such as the reflexive attentional orienting observed in exogenous cueing paradigms (Klein, 2000). Notably, the probes in the current study were never relevant or to be responded to and thus were never targets in any functional sense. However, it is the case that our probe response data pattern shows that attention is automatically drawn toward the visual modality when the probe is preceded by a multisensory stimulus in which the auditory component was relevant and the visual component was incongruent versus congruent with that auditory input. Thus, the attentional "capture" observed in this study is not exogenous in nature (such as when attentional orienting can be triggered by a low-level exogenous flash) but is rather a function of high-level processing of the congruency of a preceding stimulus, wherein attention is drawn to the source of conflicting informational input. Accordingly, although both these multisensory-conflict attention-capture effects and exogenous-cueing attentional-capture effects have an automatic or reflexive nature to them, the underlying cognitive and neural mechanisms would seem to fundamentally differ.

In summary, the present results suggests that, in the face of conflicting multisensory stimulus input, attentional resources appear to be drawn toward the irrelevant modality when it contains an incongruent multisensory stimulus. Specifically, when task-irrelevant stimulus input conflicts with input that is relevant, it triggers a rapid increase of attentional resources to be drawn to the irrelevant input, resulting in an enhancement of its processing. This result is thus in sharp contrast to the alternative possibility, which would have been that of rapid task performance-enhancing suppression of the processing of the distracting incongruent input, as this would have been evidenced by reduced sensory responses to the visual probes. Although further work needs to be done to lay out the entire temporal cascade of attentional modulation in response to conflict processing, including later processes in the cascade, the current findings provide strong evidence concerning the initial part of the brain's response to intermodal incongruent stimulus input. In particular, this early response is dominated by the rapid distraction or capture of attentional resources toward the source of the incongruent stimulus input, with this attentional distraction then likely serving as a key underlying mechanism for the behavioral performance decrements that are observed under these circumstances.

Acknowledgments

This work was supported by an NSF graduate student fellowship award to S. E. D. and an NIH RO1 grant (RO1-NS051048) to M. G. W.

Reprint requests should be sent to Sarah E. Donohue, Center for Cognitive Neuroscience, Duke University, Box 90999, Durham, NC 27708-0999, or via e-mail: sarah.donohue@duke.edu.

Note

1. The electrodes over which these and subsequent analyses were conducted are reported with the 10–20 system naming convention. Although the 10–20 sites and our 128-channel sites do not overlap perfectly, the 10–20 sites most proximal to our electrodes are reported.

REFERENCES

Appelbaum, L. G., Smith, D. V., Boehler, C. N., Chen, W. D., & Woldorff, M. G. (2011). Rapid modulation of sensory processing induced by stimulus conflict. *Journal of Cognitive Neuroscience, 23,* 2620–2628.

Atkinson, C. M., Drysdale, K. A., & Fulham, W. R. (2003). Event-related potentials to Stroop and reverse Stroop stimuli. *International Journal of Psychophysiology, 47,* 1–21.

Bendixen, A., Grimm, S., Deouell, L. Y., Wetzel, N., Madebach, A., & Schroger, E. (2010). The time-course of auditory and visual distraction effects in a new crossmodal paradigm. *Neuropsychologia, 48,* 2130–2139.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108,* 624–652.

Brass, M., Derrfuss, J., & von Cramon, D. Y. (2005). The inhibition of imitative and overlearned responses: A functional double dissociation. *Neuropsychologia, 43,* 89–98.

Carter, C. S., Mintun, M., & Cohen, J. D. (1995). Interference and facilitation effects during selective attention: An (H2O)-O-15 PET study of Stroop task performance. *Neuroimage, 2,* 264–272.

Clark, V. P., & Hillyard, S. A. (1996). Spatial selective attention affects early extrastriate but not striate components of the visual evoked potential. *Journal of Cognitive Neuroscience, 8,* 387–402.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3,* 201–215.
De Ruiter, M. B., Kok, A., & van der Schoot, M. (1998). Effects of inter- and intramodal selective attention to non-spatial visual stimuli: An event-related potential analysis. *Biological Psychology, 49,* 269–294.
Donohue, S. E., Liotti, M., Perez, R., & Woldorff, M. W. (2012). Is conflict monitoring supramodal? Spatiotemporal dynamics of cognitive control processes in an auditory Stroop task. *Cognitive, Affective, & Behavioral Neuroscience, 12,* 1–15.
Egner, T., & Hirsch, J. (2005a). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nature Neuroscience, 8,* 1784–1790.
Egner, T., & Hirsch, J. (2005b). The neural correlates and functional integration of cognitive control in a Stroop task. *Neuroimage, 24,* 539–547.
Eimer, M., & Schroger, E. (1998). ERP effects of intermodal attention and cross-modal links in spatial attention. *Psychophysiology, 35,* 313–327.
Erickson, K. I., Prakash, R. S., Kim, J. S., Sutton, B. P., Colcombe, S. J., & Kramer, A. F. (2009). Top–down attentional control in spatially coincident stimuli enhances activity in both task-relevant and task-irrelevant regions of cortex. *Behavioural Brain Research, 197,* 186–197.
Erikson, B., & Erikson, C. (1974). Effects of noise letters upon the identification of a target letter in a non-search task. *Perception & Psychophysics, 16,* 143–149.
Fan, J., Hof, P. R., Guise, K. G., Fossella, J. A., & Posner, M. I. (2008). The functional integration of the anterior cingulate cortex during conflict processing. *Cerebral Cortex, 18,* 796–805.
Heinze, H. J., Mangun, G. R., Burchert, W., Hinrichs, H., Scholz, M., Munte, T. F., et al. (1994). Combined spatial and temporal imaging of brain activity during visual selective attention in humans. *Nature, 372,* 543–546.
Hickey, C., Di Lollo, V., & McDonald, J. J. (2009). Electrophysiological indices of target and distractor processing in visual search. *Journal of Cognitive Neuroscience, 21,* 760–775.
Hopf, J. M., Boehler, C. N., Luck, S. J., Tsotsos, J. K., Heinze, H., & Schoenfeld, M. A. (2006). Direct neurophysiological evidence for spatial suppression surrounding the focus of attention in vision. *Proceedings of the National Academy of Sciences, U.S.A., 103,* 1053–1058.
Kayser, C., Petkov, C. I., & Logothetis, N. K. (2008). Visual modulation of neurons in auditory cortex. *Cerebral Cortex, 18,* 1560–1574.
Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Sciences, 4,* 138–147.
Liotti, M., Woldorff, M. G., Perez, R., & Mayberg, H. S. (2000). An ERP study of the temporal course of the Stroop color-word interference effect. *Neuropsychologia, 38,* 701–711.
Logothetis, N. K., & Wandell, B. A. (2004). Interpreting the BOLD signal. *Annual Review of Physiology, 66,* 735–769.
Luck, S. J. (1995). *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press.
Luck, S. J., Heinze, H. J., Mangun, G. R., & Hillyard, S. A. (1990). Visual event-related potentials index focused attention within bilateral stimulus arrays. 2. Functional dissociation of P1 and N1 components. *Electroencephalography and Clinical Neurophysiology, 75,* 528–542.
MacDonald, A. W., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science, 288,* 1835–1838.
Mangun, G. R., & Hillyard, S. A. (1988). Spatial gradients of visual-attention—Behavioral and electrophysiological evidence. *Electroencephalography and Clinical Neurophysiology, 70,* 417–428.
Mangun, G. R., Hopfinger, J. B., Kussmaul, C. L., Fletcher, E. M., & Heinze, H. J. (1997). Covariations in ERP and PET measures of spatial selective attention in human extrastriate visual cortex. *Human Brain Mapping, 5,* 273–279.
Markela-Lerenc, J., Ille, N., Kaiser, S., Fiedler, P., Mundt, C., & Weisbrod, M. (2004). Prefrontal-cingulate activation during executive control: Which comes first? *Cognitive Brain Research, 18,* 278–287.
Molholm, S., Ritter, W., Javitt, D. C., & Foxe, J. J. (2004). Multisensory visual-auditory object recognition in humans: A high-density electrical mapping study. *Cerebral Cortex, 14,* 452–465.
Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science, 229,* 782–784.
O'Donnell, B. F., Swearer, J. M., Smith, L. T., Hokama, H., & McCarley, R. W. (1997). A topographic study of ERPs elicited by visual feature discrimination. *Brain Topography, 10,* 133–143.
Polk, T. A., Drake, R. M., Jonides, J. J., Smith, M. R., & Smith, E. E. (2008). Attention enhances the neural processing of relevant features and suppresses the processing of irrelevant features in humans: A functional magnetic resonance imaging study of the Stroop task. *Journal of Neuroscience, 28,* 13786–13792.
Roberts, K. L., & Hall, D. A. (2008). Examining a supramodal network for conflict processing: A systematic review and novel functional magnetic resonance imaging data for related visual and auditory Stroop tasks. *Journal of Cognitive Neuroscience, 20,* 1063–1078.
Rugg, M. D., Milner, A. D., Lines, C. R., & Phalp, R. (1987). Modulation of visual event-related potentials by spatial and nonspatial visual selective attention. *Neuropsychologia, 25,* 85–96.
Silton, R. L., Heller, W., Towers, D. N., Engels, A. S., Spielberg, J. M., Edgar, J. C., et al. (2010). The time course of activity in dorsolateral prefrontal cortex and anterior cingulate cortex during top–down attentional control. *Neuroimage, 50,* 1292–1302.
Smiley, J. F., Hackett, T. A., Ulbert, I., Karmas, G., Lakatos, P., Javitt, D. C., et al. (2007). Multisensory convergence in auditory cortex, I. Cortical connections of the caudal superior temporal plane in macaque monkeys. *Journal of Comparative Neurology, 502,* 894–923.
Spitzer, H., Desimone, R., & Moran, J. (1988). Increased attention enhances both behavioral and neuronal performance. *Science, 240,* 338–340.
Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: Current issues from the perspective of the single neuron. *Nature Reviews Neuroscience, 9,* 255–266.
Strayer, D. L., & Johnston, W. A. (2001). Driven to distraction: Dual-task studies of simulated driving and conversing on a cellular telephone. *Psychological Science, 12,* 462–466.
Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology: General, 18,* 643–662.
van Atteveldt, N., Formisano, E., Goebel, R., & Blomert, L. (2004). Integration of letters and speech sounds in the human brain. *Neuron, 43,* 271–282.
van Atteveldt, N. M., Blau, V. C., Blomert, L., & Goebel, R. (2010). fMR-adaptation indicates selectivity to audiovisual content congruency in distributed clusters in human superior temporal cortex. *BMC Neuroscience, 11,* 11.
Vanvoorhis, S., & Hillyard, S. A. (1977). Visual evoked-potentials and selective attention to points in space. *Perception & Psychophysics, 22,* 54–62.

Weissman, D. H., Roberts, K. C., Visscher, K. M., & Woldorff, M. G. (2006). The neural bases of momentary lapses in attention. *Nature Neuroscience, 9,* 971–978.
Weissman, D. H., Warner, L. M., & Woldorff, M. G. (2004). The neural mechanisms for minimizing cross-modal distraction. *Journal of Neuroscience, 24,* 10941–10949.
Weissman, D. H., Warner, L. M., & Woldorff, M. G. (2009). Momentary reductions of attention permit greater processing of irrelevant stimuli. *Neuroimage, 48,* 609–615.
West, R., & Alain, C. (1999). Event-related neural activity associated with the Stroop task. *Cognitive Brain Research, 8,* 157–164.
Woldorff, M. G., Fox, P. T., Matzke, M., Lancaster, J. L., Veeraswamy, S., Zamarripa, F., et al. (1997). Retinotopic organization of early visual spatial attention effects as revealed by PET and ERPs. *Human Brain Mapping, 5,* 280–286.
Zimmer, U., Itthipanyanan, S., Grent-'T-Jong, T., & Woldorff, M. (2010). The electrophysiological time course of the interaction of stimulus conflict and the multisensory spread of attention. *European Journal of Neuroscience, 31,* 1744–1754.
Zimmer, U., Roberts, K. C., Harshbarger, T. B., & Woldorff, M. G. (2010). Multisensory conflict modulates the spread of visual attention across a multisensory object. *Neuroimage, 52,* 606–616.