RESEARCH ARTICLE

Multisensory integration affects ERP components elicited by exogenous cues

**Valerio Santangelo · Rob H. J. Van der Lubbe ·
Marta Olivetti Belardinelli · Albert Postma**

Abstract Previous studies have shown that the amplitude
of event related brain potentials (ERPs) elicited by a com-
bined audiovisual stimulus is larger than the sum of a single
auditory and visual stimulus. This enlargement is thought
to reflect multisensory integration. Based on these data, it
may be hypothesized that the speeding up of responses, due
to exogenous orienting effects induced by bimodal cues,
exceeds the sum of single unimodal cues. Behavioral data,
however, typically revealed no increased orienting effect
following bimodal as compared to unimodal cues, which
could be due to a failure of multisensory integration of the
cues. To examine this possibility, we computed ERPs elic-
ited by both bimodal (audiovisual) and unimodal (either
auditory or visual) cues, and determined their exogenous
orienting effects on responses to a to-be-discriminated
visual target. Interestingly, the posterior P1 component
elicited by bimodal cues was larger than the sum of the P1
components elicited by a single auditory and visual cue
(i.e., a superadditive effect), but no enhanced orienting effect
was found on response speed. The latter result suggests that
multisensory integration elicited by our bimodal cues plays
no special role for spatial orienting, at least in the present
setting.

Keywords Event-related brain potentials · Exogenous
orienting · Spatial attention · Multisensory integration ·
Unimodal cues · Crossmodal cues · Bimodal cues

V. Santangelo (✉) · M. Olivetti Belardinelli
Department of Psychology, University of Rome "La Sapienza",
via dei Marsi 78, 00185 Rome, Italy
e-mail: valerio.santangelo@uniroma1.it

R. H. J. Van der Lubbe · A. Postma
Psychological Laboratory, Helmholtz Institute,
Utrecht University, Utrecht, The Netherlands

M. Olivetti Belardinelli
ECONA, Interuniversity Center for Research on Cognitive
Processing in Natural and Artificial Systems, Rome, Italy

R. H. J. Van der Lubbe
Cognitive Psychology and Ergonomics,
Universiteit Twente, Enschede, The Netherlands

Introduction

In everyday life numerous stimuli reflexively capture our
attention. The ability to focus on those stimuli in order to
extract relevant information is the core of attentional orient-
ing mechanisms. Several studies have shown that behav-
ioral responses to stimuli that combine multimodal
attributes in close spatial and temporal proximity are typi-
cally faster than stimuli that involve unimodal features
(e.g., Molholm et al. 2002; Teder-Sälejärvi et al. 2002).
Usually, multisensory integration is investigated by com-
paring neural responses to each of two separately presented
stimuli of a different modality, with the response to the
bimodal combination of the two (e.g., Stein and Meredith
1993). The rationale is that if the bimodal response is
greater than the sum of the individual unimodal responses,
then it may be concluded that some form of facilitatory
multisensory integration is taking place (though see Stan-
ford and Stein 2007).

Although the neural correlates of this multisensory
enhancement of perception have been investigated exten-
sively in both animals and humans, the underlying mecha-
nisms are still unclear. In non-human mammals, neurons
have been identified in the superior temporal sulcus (STS)
and the superior colliculus (SC) that respond to both
auditory and visual stimulation (Stein et al. 1993). These

neurons show a superadditive effect in case of spatially and temporally coincident multimodal stimulation (Stein and Meredith 1993). In other words, multisensory stimuli significantly enhance the responses of STS and SC neurons above those evoked by either unimodal stimuli (Wallace et al. 1998). In humans, enhancement of neural activity in case of multimodal stimulation has been demonstrated using event-related brain potentials (Giard and Peronnet 1999; Talsma and Woldorff 2005; Teder-Sälejärvi et al. 2002; see also Fort et al. 2002), gamma band responses (Senkowski et al. 2005), and other neuroimaging measures (Calvert et al. 2000; though see Beauchamp, Argall et al. 2004; Dhamala et al. 2007, for null results).

For instance, Teder-Sälejärvi et al. (2002) investigated multisensory integration by presenting random sequences of auditory (brief noise bursts), visual (flashes), and audio-visual (simultaneous noise bursts and flashes) stimuli from a central location at irregular latencies ranging from 600 to 800 ms. Their participants had to press a response button to infrequent and unpredictable target stimuli that could be either a more intense noise burst, a brighter flash, or a combination of the two. Behavioral results showed that bimodal target stimuli were responded to more rapidly and more accurately than the unimodal target stimuli. The neural basis of multisensory interaction was assessed by subtracting the ERPs to the auditory (A) and the visual (V) stimuli alone from the ERP to the combined audiovisual (AV) stimuli [i.e., $AV - (A + V)$]. Using this subtractive method, the authors found clear evidence of ERP interaction effects at both anterior (Fz, F3, F4, C3, C4, Cz) and posterior (Pz, P3, P4, O1, O2, Oz) sites at 100 ms from stimulus onset. However, they acknowledged that slow potentials beginning prior to the stimulus onset might have contributed to the bimodal minus unimodal difference waveform, probably spoiling the early multisensory interactions that they and previous studies (e.g., Giard and Perronet 1999) reported.

Given that these slow potentials likely belong to the CNV family of ERPs that precede perceptual decisions and discriminative responses to task-relevant stimuli (Hillyard 1973), this problem might be avoided using task-irrelevant stimuli. In fact, as Teder-Sälejärvi et al. (2002) argued: "Studies of cross-modal interactions in sensory processing need to take into account task-related neural activity that may be elicited following each stimulus but may not necessarily have anything to do with integration of information across different modalities. [...] On the other hand, designs that do not assign any task-relevance to the unimodal and bimodal stimuli […] would not be likely to encounter this problem." (p. 114).

It may be hypothesized that multisensory integration of task-irrelevant stimuli affects the magnitude of orienting effects as well. In fact, the typical facilitation observed in response to multimodal (as compared to unimodal) events should take place also when bimodal exogenous cues reflexively capture spatial attention. Nevertheless, it cannot be taken for granted that multisensory integration of task-irrelevant events need to play any additional role in generating exogenous orienting effects (see McDonald et al. 2001). Indeed, several studies provided indirect support for the view that involuntary shifts of spatial attention and multisensory integration are independent. For instance, the ventriloquism effect, which is due to multisensory integration processes (the apparent location of a speech or non-speech sound is altered by the synchronized presentation of a transient visual event), has been shown to occur preattentively and independently of both voluntary and involuntary spatial attention shifts (see, for instance, Bertelson et al. 2000; Vroomen et al. 2001). Moreover, it has additionally been argued that reflexive shifts of attention depend on the attentional goals of the observer, while multisensory integration occurs preattentively and without intent (Folk et al. 1992; McDonald et al. 2001; though see Berger et al. 2005; Van der Lubbe and Postma 2005; Van der Lubbe and Van der Helden 2006).

In a previous study (Santangelo et al. 2006), we directly addressed the relevance of multisensory integration for exogenous orienting. We compared a condition in which a to-be-discriminated visual target was preceded by both visual and auditory exogenous cues presented together on the same side (bimodal cue), with conditions in which the target was preceded by either a visual (unimodal cue) or an auditory cue (crossmodal cue) alone. The results revealed comparable magnitudes of cuing effects on RTs among unimodal, crossmodal and bimodal conditions, consistent with a number of previous studies (e.g., Spence and Driver 1999; Ward 1994; see also Ward et al. 1998; Funes et al. 2005). In other words, the bimodal cue did not produce any enlarged cuing effect (in both RT and error rate data) with respect to both the unimodal and the crossmodal cue.

This result could be due to a failure in audio–visual integration mechanisms, (i.e., no multisensory integration), or alternatively, integration might have taken place, but had no effect on exogenous orienting (e.g., Giard and Perronet 1999). A possible reason for the absence of multisensory integration is the use of quite simple cues that implied no task demands, although this opposes the earlier comments regarding the ventriloquism effect. Nevertheless, one could argue that multisensory integration only occurs in case of task-relevance of specific stimuli.

The aim of this study was therefore to examine whether or not multisensory integration took place in case of bimodal stimulation, which may be inferred from elicited neural activity, and to compare this effect with a behavioral

measure indicative for exogenous orienting. To this end, we measured ERPs following nonpredictive unimodal (visual), crossmodal (auditory), and bimodal (audiovisual) cues, and determined their exogenous orienting effects on RT in a visuo-spatial discrimination task. Obviously, if multisensory integration affects attentional orienting then we not only expect to find larger ERPs for bimodal (i.e., audiovisual) than for unimodal (i.e., visual) and crossmodal (i.e., auditory) cues, but additionally an enlarged exogenous cuing effect on RT for bimodal as compared to unimodal and crossmodal cues.

Method

Participants

Twelve participants (mean age 22.1 years, range 19–32, 4 male, 11 right-handed, 1 left-handed), students of Utrecht University, participated. They had normal or corrected-to-normal vision and normal hearing, and were paid 25 € for their participation. Before the start of the experiment, they had to perform a test in which they had to indicate the side (left vs. right) of auditory stimuli at the cued locations. Participants had to score at least 95% correct in order to participate in the subsequent experiment. The study was approved by a local ethics board of the University of Utrecht, and informed consent was obtained from all participants.

Stimuli and materials

Stimuli were presented on three crossmodal units (21 × 12 cm), consisting of a sound passing 8 × 8 green LED display (10 × 10 cm) in front of a loudspeaker. The units were hung at 160 cm from the participant, either 29.4° to the left or the right, or in front of the participant. The fixation point was a dot (0.7 × 0.7°). A line (0.2 × 3.1°) presented for 100 ms served as visual cue. The auditory cue was a burst of white noise presented for 100 ms. Pointing up- or downwards triangles (2.6 × 1.4°) presented for 100 ms served as visual targets. The visual cues were presented at the bottom of crossmodal units while the targets were presented at the top of them (see Fig. 1), in order to avoid presentation from overlapping spatial locations (cf. Spence and Driver 1994).

Task and procedure

Three conditions were included in which visual and/or auditory cues preceded visual targets. Cue modality varied between conditions in order to avoid interference from subsequent trials with different cues, and their order was counterbalanced. In the unimodal condition, visual targets were preceded by a visual cue; in the crossmodal condition, visual targets were preceded by an auditory cue; in the bimodal condition, visual targets were preceded by both a visual and an auditory cue simultaneously delivered at the same unit.
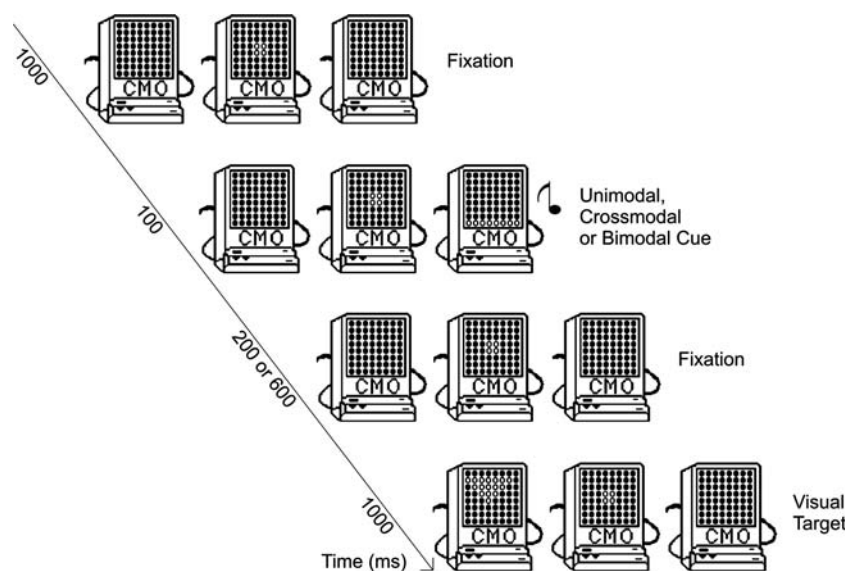

Fig. 1 Schematic representation of the sequence of events and their duration (starting from the top) in the experimental trials. Each trial began with a fixation point, which remained till the end of the trial. After 1,000 ms, either a visual (i.e., unimodal), auditory (i.e., crossmodal), or audiovisual (i.e., bimodal) cue was presented equiprobably on the left or right side for 100 ms. After a variable interval (either 200 or 600 ms), the visual target (i.e., the *triangle*) was presented equiprobably on the left or right side for 100 ms, requiring a left or right button press when the triangle pointed up- or downwards, respectively

Figure 1 illustrates the sequence of events on each experimental trial. The participants were comfortably seated in an armchair in a silent and darkened room. Left and right response buttons were placed on a hand rest in front of the participant. Each trial started with the presentation of a fixation point on the middle unit. After a delay of 1,000 ms, the cue (either unimodal, crossmodal or bimodal) was presented on the left or right unit for 100 ms. After a further delay (SOA of 200 or 600 ms), the visual target was presented on either the left or the right unit. SOA, cue location and target location were all equally probable, and varied randomly from trial to trial. Hence, cue location was non-predictive with regard to the forthcoming target position; the target occurred either at the same (cued trials) or the opposite (uncued trials) side of the cue.

Participants performed a visuo-spatial discrimination task. They were required to press a left or right button as rapidly and accurately as possible for a triangle (i.e., the visual target) pointing up- or downwards, respectively. The next trial started after either a response or 1,500 ms from target onset. On 20% of trials, the cue was not followed by the target, in order to reduce anticipatory responses (i.e. catch trials; cf. Low and Miller 1999).

Each type of trial [48 combinations—cue type (3) × cue side (2) × target side (2) × target type (2) × SOA (2)], was repeated 50 times (2,400 trials). Six hundred catch trials were additionally presented (200 for cue type), totally amounting to 3,000 trials. Consequently, there were 100 trials per experimental condition when collapsed across the side of target presentation and the required response. The trials were delivered in six different blocks, two for each cue type. Each block comprised 500 trials (including 100 catch trials) and lasted approximately 25 min. Participants were allowed to rest 5 min between the blocks.

Electrophysiological recording

The presentation of stimuli and the emission of triggers signaling the moment and the type of stimulus was controlled by a CMO-module (version 3.7f; see also Van der Lubbe and Postma 2005; Van der Lubbe et al. 2006; Santangelo et al. 2006). The triggers were received by Vision Recorder (version 1.0b, BrainProducts, GmbH), which measured the participants' electroencephalographic (EEG) signals, eye movements [the horizontal and vertical electro-oculogram (EOG)], and button presses.

EEG signals were recorded from 64 Ag/AgCl ring electrodes, including 60 standard sites from the 10 to 10 system (Fpz, Fp1, Fp2, AFz, AF3, AF4, AF7, AF8, Fz, F1, F2, F3, F4, F5, F6, F7, F8, FCz, FC1, FC2, FC3, FC4, FC5, FC6, FT7, FT8, C1, C2, C3, C4, C5, C6, T7, T8, CPz, CP1, CP2, CP3, CP4, CP5, CP6, Pz, P1, P2, P3, P4, P5, P6, P7, P8, POz, PO3, PO4, PO7, PO8, Oz, O1, O2) and were online

referenced to Cz. Vertical and horizontal EOG was recorded using two Ag/AgCl ring electrodes placed above and below the left eye and at the outer canthi of both eyes, respectively. The EEG was digitized at 250 Hz (time constant of 4 s) with an amplifier bandpass of 0.04–100 Hz (half amplitude low- and high-frequency cut-offs, respectively) together with a 50 Hz notch filter and was stored for off-line analyses. Impedances were kept below 5 kΩ. The level of impedances was checked every two blocks during participants' time rest.

Data analysis

Offline, the EEG and EOG were segmented into 2,000-ms epochs, with a 100-ms pre-cue baseline. Trials with amplitudes on the EOG channels exceeding 60 μV from cue onset until target offset (determined by Brain Vision Analyzer 1.03) were excluded, which implies that no trials remained with eye movements larger than approximately 3° (cf. Van der Lubbe and Postma 2005; Santangelo et al. 2006).

Computerized artefact rejection was performed before signal averaging to discard epochs related to cue onset in which deviations in eye position, blinks, or muscular activity occurred. On average, about 9% of the trials were rejected. Trials in which participants incorrectly responded to target stimuli were also removed before averaging. Artefact-free data were then used to create averaged waveforms. All trials (including catch trials) were included in the grand averages. Separate cue ERPs were created for left and right cues. All final ERPs were reduced to a length of 200 ms (just prior to a possible target onset). Average activity was measured from 130 to 150 ms after cue onset on parieto-occipital (i.e., PO7–PO8) and central (i.e., C1–C2) sites with respect to the 100 ms pre-stimulus baseline. Additionally, peak latencies and amplitudes[1] of the P1 and N1 components on these sites were determined within time windows of 80–180 ms and 80–200 ms, respectively.

Results

Behavioral data

Mean RT cuing effects (i.e., uncued–cued trials) are displayed in Fig. 2. Trials with incorrect responses were eliminated from the RT analysis, as were trials in which the RT

---

[1] Given the fact that the signal to noise ratio (SNR) of amplitude measures depends on the number of trials used to create the averages, we examined whether conditions differed in these amounts. A two-way ANOVA with the within-participants factors of Cue type (unimodal, crossmodal, or bimodal) and Cue Side (left or right) revealed no significant effects (all $Fs < 2.7$; $Ps > 0.128$), indicating that there were no differences in the number of trials between conditions.
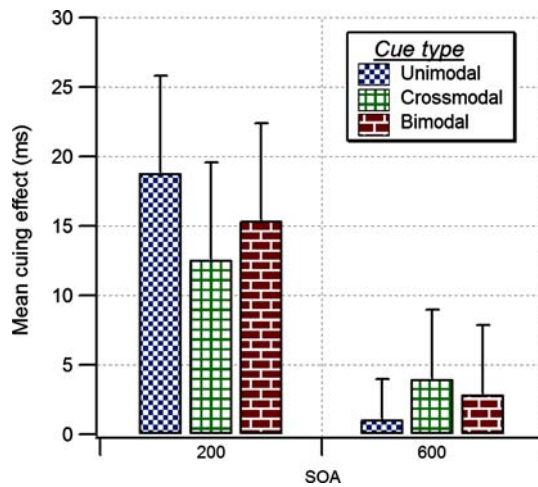
Fig. 2 Magnitude of cuing effect (i.e., uncued–cued trials) and standard error of the mean (i.e., *error bars*) for unimodal, crossmodal, and bimodal cues at 200 and 600 ms of SOA

was shorter than 100 (i.e., premature) or longer than 1,500 ms (i.e., misses). Both premature responses and misses occurred seldomly (an average of 0.1% of the trials). Trials with eye movements (8.7% of the trials) were excluded from the analysis as well.

A three-way within-participants ANOVA with the factors of Cue type (unimodal, crossmodal, or bimodal), SOA (200 or 600 ms), and Cuing (cued vs. uncued) was performed on the RT data. This analysis revealed that the participants responded significantly faster for cued (565 ms) than for uncued targets (575 ms) [$F(1, 11) = 6.0$, $P = 0.033$]. The participants' responses were also significantly slower in case of an SOA of 200 ms SOA (576 ms) than in case of an SOA of 600 ms (564 ms) [$F(1, 11) = 13.5$, $P = 0.004$]. Moreover, responses were faster when the cue was crossmodal (545 ms) than when it was bimodal (575 ms), which, in turn, were faster than when it was unimodal (590 ms) [$F(2, 22) = 5.7$, $P = 0.015$].
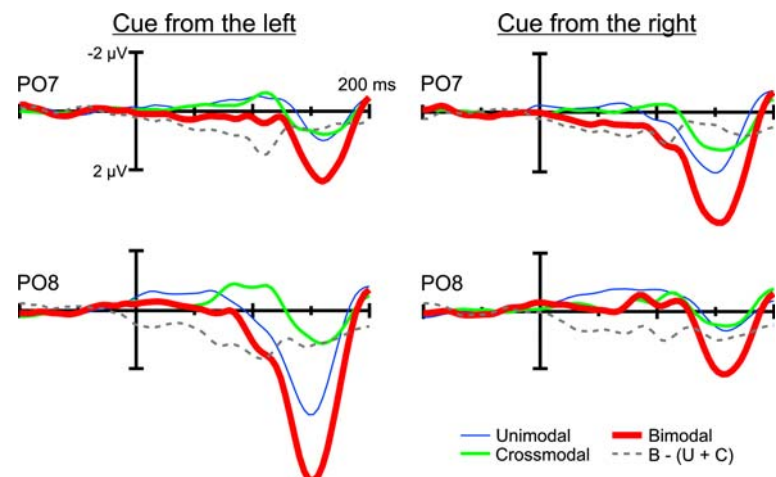
Crucially, neither the SOA × Cue type × Cuing interaction [$F(2, 22) = 0.3$, $P = 0.727$], nor the Cue type × Cuing interaction were significant [$F(2, 22) = 0.3$, $P = 0.922$]. These results, which are consistent with a number of previous findings (e.g., Santangelo et al. 2006, Experiment 1; Spence and Driver 1999; Ward 1994; Ward et al. 1998), clearly indicate that cuing effects elicited by unimodal, crossmodal, and bimodal cues are of comparable magnitude. The analysis did not reveal any other significant effect, as well as a similar analysis performed on the error data (all $Fs < 2.4$; $Ps > 0.111$).

Electrophysiological data

Figures 3 and 4 show the ERP components elicited by the different types of cue for channels at which the amplitude of components elicited by the bimodal cues were maximal (isopotential—spline—maps of these components are illustrated in Fig. 5 at three different time windows, i.e., 100–120, 120–140, and 140–160 ms). The waveforms show a visual P1 and an auditory N1 component. In particular, both an auditory N1 component and a visual P1 component appeared between 130 and 150 ms from cue onset. The visual P1 was maximal at parieto-occipital sites (i.e., PO7/PO8; see Fig. 3), while the auditory N1 was maximal at central sites (i.e., C1/C2; see Fig. 4). The average activity within these time windows was collapsed across the side of cue presentation (left vs. right) in order to determine ipsilateral and contralateral P1 and N1 components, and then used for the statistical analyses. Not surprisingly, the visual P1 component was larger on contralateral (relative to the side on which the stimulus was presented) than on ipsilateral sites ($t = 6.2$, $P < 0.001$), while no significant difference was found between ipsilateral and contralateral sites for the auditory N1 component ($t = -0.815$, $P = 0.432$). However, the analysis of peak voltage deflections additionally revealed a lateralization effect for the auditory N1 component,

Fig. 3 Event-related brain potentials [and difference waves B − (U + C)] at parieto-occipital sites (i.e., PO7/8) elicited by unimodal, crossmodal, and bimodal cues presented on either the left or right side
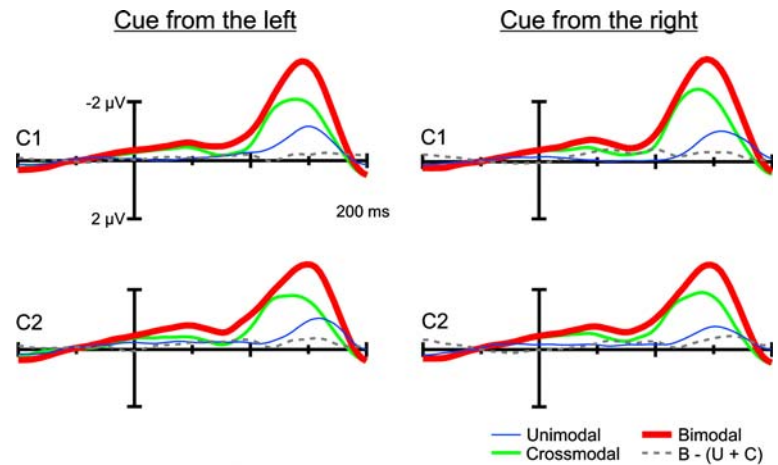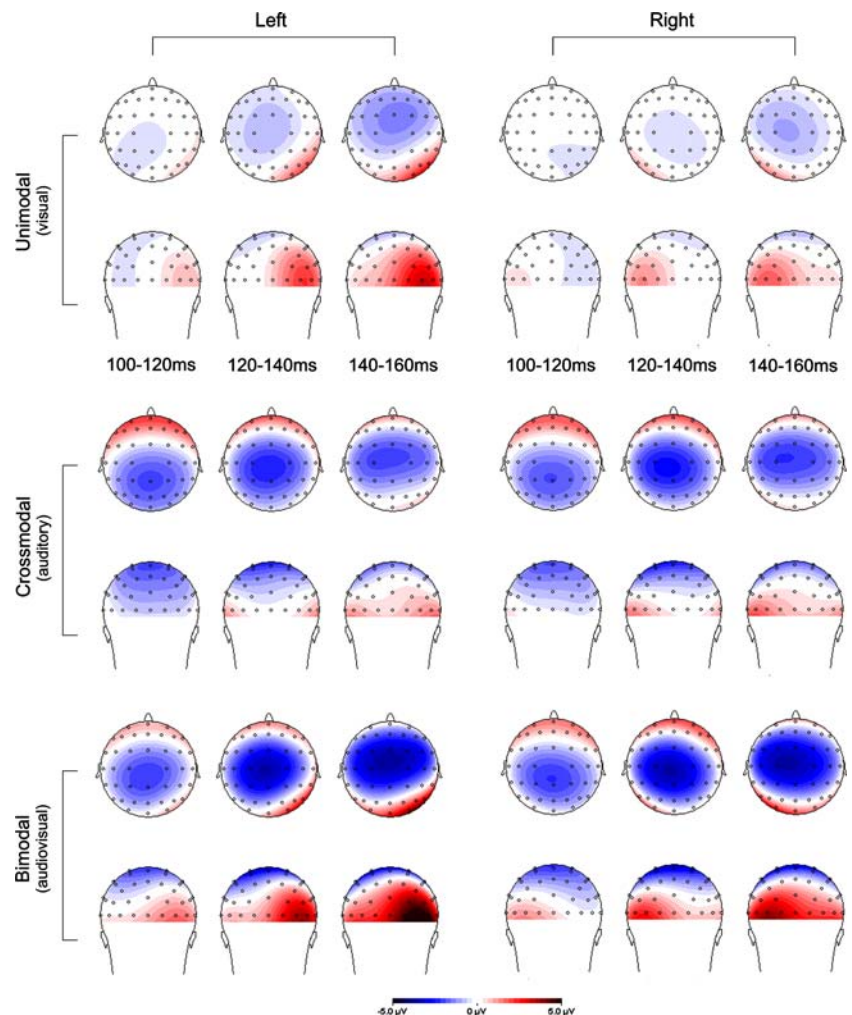
showing a larger peak amplitude at contralateral as com-
pared to ipsilateral central sites ($t = -2.6$, $P = 0.023$). Simi-
larly, there was a larger peak voltage deflection at
contralateral as compared to ipsilateral sites for the visual
P1 component ($t = 4.7$, $P = 0.001$), while no significant
differences were revealed by means of the analysis of peak

latencies for both the visual P1 and the auditory N1 compo-
nents ($t = 0.1$, $P = 1.0$ and $t = .8$, $P = 0.436$, respectively).

In order to investigate multisensory interactions, ERP
components elicited by unimodal (U) and crossmodal (C)
cues alone were subtracted from ERP components elicited
by bimodal (B) cues (cf. Teder-Sälejärvi et al. 2002) at both

parieto-occipital (PO7/8; P1 components) and central (C1/2; N1 components) sites. The resulting waveforms are included in Figs. 3 and 4 (dotted lines). The [B − (U + C)] contralateral and ipsilateral difference waves were additionally collapsed across cue location (left vs. right) and are displayed in Fig. 6.

A two-tailed t test on each difference wave (i.e., ipsilateral and contralateral PO7/8 and ipsilateral and contralateral C1/2) within the 130–150 ms time window was performed to examine whether residual neural activity [i.e., B − (U + C)] deviated from 0. The results showed residual activity for the ipsilateral and contralateral PO7/8 sites (P1: $t = 2.6$, $P = 0.024$ and $t = 2.8$, $P = 0.018$, respectively) but not for the ipsilateral and contralateral C1/2 difference waves (N1: $t = -0.4$, $P = 0.692$ and $t = -0.2$, $P = 0.843$, respectively; this null effect on central sites was also confirmed by the analyses on peak latencies and amplitudes, $t = -0.1$, $P = 0.821$ and $t = -0.3$, $P = 0.744$, respectively). These results indicate that the P1 component elicited by the bimodal cue was larger than the sum of the P1 components elicited by unimodal and crossmodal cues at parieto-occipital sites. In particular, contralateral difference waves were larger than ipsilateral difference waves at parieto-occipital sites ($t = 3.3$, $P = 0.007$), which crucially indicates that bimodal (audiovisual) exogenous cues enhanced neural activity on contralateral parieto-occipital sites as compared to unimodal (visual) or crossmodal (auditory) cues at P1 latencies. This pattern was also confirmed by the analysis of peak voltage deflections, which once again indicated larger activity (i.e., peak amplitude) on contralateral parieto-occipital sites ($t = 4.2$, $P = 0.004$).

Discussion

The major question addressed in this study was to examine whether audiovisual (bimodal) spatially nonpredictive (exogenous) cues elicited larger neural activity than the sum of the activities elicited by unimodal visual and auditory cues. This question was raised in order to determine whether multisensory integration affects the magnitude of exogenous orienting effects. We computed ERPs elicited by spatially nonpredictive visual, auditory, and audiovisual cues (which do not suffer from the neural activity related to perceptual decision and discriminative responses, cf. Teder-Sälejärvi et al. 2002), and assessed their exogenous orienting effects in a visuo-spatial discrimination task.

Behavioral results showed comparable magnitudes of cuing effects between unimodal, crossmodal and bimodal conditions, which is consistent with several previous findings (Spence and Driver 1999; Ward 1994; see also Santangelo et al. 2006, Experiment 1; Ward et al. 1998). By contrast, ERP components at P1 latencies elicited by bimodal (audiovisual) cues showed a superadditive effect as compared to ERP components elicited by unimodal (either visual or auditory) cues. Specifically, the P1 elicited by the audiovisual cue was larger that the sum of P1 elicited by visual and auditory cues on parieto-occipital sites (see Figs. 3, 6).

The present results provide the first empirical evidence (at least to our knowledge) that multisensory integration affects neural activity elicited by peripheral (exogenous) audiovisual cues on contralateral parieto-occipital sites. Concerning the absence of superadditive effects at central sites, or the absence of lateralized effects for the auditory N1 in terms of average activity (even though the contralateral component was larger than the ipsilateral component in terms of peak voltage deflection), we believe that it may be related to the fact that auditory stimuli were entirely task-irrelevant in our paradigm. That is, they were spatially nonpredictive (as were the visual and bimodal cues) but also modality-irrelevant (note that in the present study the targets were never presented in the auditory modality).

It remains open the reason why the enhanced neural activity elicited by bimodal as compared to unimodal spatial cues does not fulfil any special role in terms of

Fig. 6 Parieto-occipital difference waves [B − (U + C)] ipsilateral and contralateral to the cue side, collapsed over cue-left and cue-right conditions and left and right scalp sites. For example, the difference wave from PO7 when the cue was presented on the left side was averaged together with the difference wave from PO8 when the cue was presented on the right side to form the ipsilateral difference wave
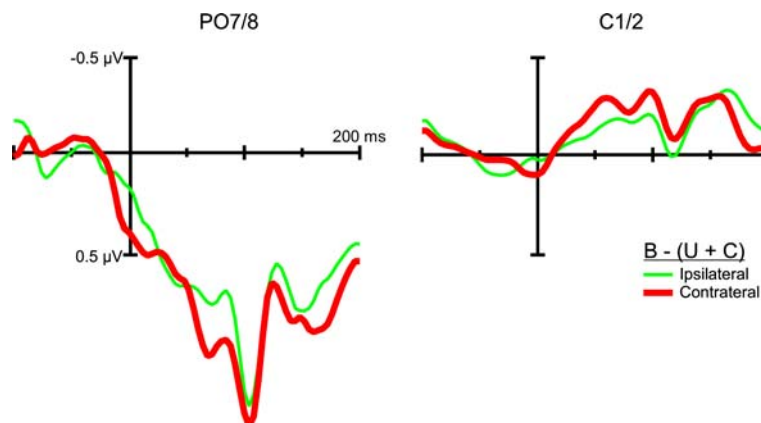
behavioral orienting effects. The easiest explanation is that multisensory integration simply does not affect the exogenous orienting of spatial attention. In terms of neural mechanisms, it may be the case that there is not any specific "link" between cerebral areas involved with the integration of multisensory stimuli and cerebral areas involved with the exogenous orienting of spatial attention. Another possible explanation for this null effect of multisensory integration on exogenous orienting is that unimodal (either visual or auditory) spatially nonpredictive cues already provide maximal attentional capture (i.e., the maximal magnitude of exogenous orienting effect), so that bimodal cues cannot provide any larger orienting effect as compared to unimodal cues. It is worth noting that this does not necessarily imply that multisensory integration cannot fulfil any special role on the exogenous orienting of spatial attention in every situation. In fact, the notion that bimodal cues as compared to unimodal cues resulted in a larger neural activity might be consistent with the view that some qualitative difference exist in the exogenous orienting mechanisms triggered by bimodal or unimodal stimuli, though this is not visible at the level of orienting effects.

This view appears to be consistent with recent evidence. For instance, Santangelo and Spence (2007a) recently compared the ability of unimodal (either visual or auditory) and bimodal (audiovisual) spatially non-predictive peripheral stimuli in capturing visuo-spatial attention under no load and high perceptual load conditions. The participants' exogenous orienting was measured by means of their performance on an orthogonal spatial cuing paradigm, in which the elevation (up vs. down) of a visual target (preceded by either a unimodal or bimodal exogenous cue) had to be discriminated. A high perceptual load condition consisted of participants having to monitor a rapidly presented central stream of visual letters for occasionally presented target digits. In the no load condition, the stream was replaced by a fixation point (i.e., there was no concurrent perceptually demanding task). The results showed that all three cues captured visuo-spatial attention in the no load condition. However, in line with other recent findings (Santangelo et al. 2007b; see also Santangelo and Spence 2007b), unimodal visual and auditory cues failed to capture visuo-spatial attention under conditions of high perceptual load (see also Schwartz et al. 2005). By contrast, audiovisual cuing effects were unaffected by the perceptual load manipulation. Santangelo and Spence's results indicate that multisensory integration (i.e., their audiovisual cue) was immune to a perceptually demanding concurrent task, showing that multisensory integration can play a key role in the behavioral orienting of spatial attention (see also Santangelo et al. 2007a).

The findings reported in the present study may be in line with this view: although the contralateral enhanced neural activity elicited by bimodal as compared unimodal cues did not result in larger cuing effects following bimodal than unimodal cues, it might nevertheless play a role in the exogenous orienting of spatial attention. For instance, it might contribute to the more resistance of cuing effects elicited by bimodal cues to disruption by perceptually demanding tasks (see Santangelo and Spence 2007a). Taken together these findings might suggest that multisensory integration (which has been shown at a physiological level in the present study following the presentation of bimodal—audiovisual—cues) promotes the exogenous orienting of spatial attention only when the task in which the participants are engaged is demanding enough. Future research may investigate, for instance, the possibility that multisensory integration (i.e., bimodal versus unimodal spatially nonpredictive cues) improves behavioral performance under conditions of difficult target discrimination (e.g., degraded stimuli), as well as been shown under conditions of increased perceptual load (Santangelo et al. 2007a; Santangelo and Spence 2007a).

References

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

Berger A, Henik A, Rafal R (2005) Competition between endogenous and exogenous orienting of visual attention. J Exp Psychol Gen 134:207–221

Bertelson P, Vroomen J, De Gelder B, Driver J (2000) The ventriloquist effect does not depend on the direction of deliberate visual attention. Percept Psychophys 62:321–332

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

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

Folk CL, Remington RW, Johnson JC (1992) Involuntary covert orienting is contingent on attentional control settings. J Exp Psychol Hum Percept 18:1030–1044

Fort A, Delpuech C, Pernier J, Giard MH (2002) Early auditory-visual interactions in human cortex during nonredundant target identification. Cogn Brain Res 14:20–30

Funes MJ, Lupianez J, Milliken B (2005) The role of spatial attention and other processes on the magnitude and time course of cueing effects. Cogn Process 6:98–116

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

Hillyard SA (1973) The CNV and human behaviour. In: McCallum WC, Knott JR (eds) Event-related slow potentials of the brain: their relation to behaviour. Elsevier, Amsterdam, pp 161–171

Low KA, Miller J (1999) The usefulness of partial information: effects of go probability in the choice/Nogo task. Psychophysiology 36:288–297

McDonald JJ, Teder-Salejarvi WA, Ward LM (2001) Multisensory integration and crossmodal attention effects in the human brain. Science 292:1791

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

Santangelo V, Ho C, Spence C (2007a) Capturing spatial attention with multisensory cues. Psychon Bull Rev (in press)

Santangelo V, Olivetti Belardinelli M, Spence C (2007b) The suppression of reflexive visual and auditory orienting when attention is otherwise engaged. J Exp Psychol Hum Percept 33:137–148

Santangelo V, Spence C (2007a) Multisensory cues capture spatial attention regardless of perceptual load. J Exp Psychol Hum Percept (in press)

Santangelo V, Spence C (2007b) Assessing the automaticity of the exogenous orienting of tactile attention. Perception (in press)

Santangelo V, Van der Lubbe RHJ, Olivetti Belardinelli M, Postma A (2006) Spatial attention triggered by unimodal, crossmodal, and bimodal exogenous cues: a comparison on reflexive orienting mechanisms. Exp Brain Res 173:40–48

Schwartz S, Vuilleumier P, Hutton C, Maravita A, Dolan RJ, Driver J (2005) Attentional load and sensory competition in human vision: modulation of fMRI responses by load at fixation during task-irrelevant stimulation in the peripheral visual field. Cereb Cortex 15:770–786

Senkowski D, Talsma D, Herrmann CS, Woldorff MG (2005) Multisensory processing and oscillatory gamma responses: effects of spatial selective attention. Exp Brain Res 166:411–426

Spence C, Driver J (1994) Covert spatial orienting in audition: exogenous and endogenous mechanisms. J Exp Psychol Human Percept 20:555–574

Spence C, Driver J (1999) A new approach to the design of multimodal warning signals. In: Harris D (ed) Engineering psychology and cognitive ergonomics. vol IV. Ashgate Publishing, Aldershot, pp 455–461

Stanford TR, Stein BE (2007) Superadditivity in multisensory integration: putting the computation in context. NeuroReport 18:787–792

Stein BE, Meredith MA (1993) The merging of the senses. MIT, Cambridge

Stein BE, Meredith MA, Wallace MT (1993) The visually responsive neuron and beyond: multisensory integration in cat and monkey. Prog Brain Res 95:79–90

Talsma D, Woldorff MG (2005) Selective attention and multisensory integration: multiple phases of effects on the evoked brain activity. J Cogn Neurosci 17:1098–1114

Teder-Sälejärvi WA, McDonald JJ, Di Russo F, Hillyard SA (2002) An analysis of audio–visual crossmodal integration by means of event-related potential (ERP) recordings. Cogn Brain Res 14:106–114

Van der Lubbe RHJ, Havik MM, Bekker EM, Postma A (2006) Task-dependent exogenous cuing effects depend on cue modality. Psychophysiol 43:145–160

Van der Lubbe RHJ, Postma A (2005) Interruption from irrelevant auditory and visual onsets even when attention is in a focused state. Exp Brain Res 164:464–471

Van der Lubbe RHJ, Van der Helden J (2006) Failure of the extended contingent attentional capture account in multimodal settings. Adv Cogn Psychol 2:255–267

Vroomen J, Bertelson P, De Gelder B (2001) The ventriloquist effect does not depend on the direction of automatic visual attention. Percept Psychophys 63:651–659

Wallace MT, Meredith MA, Stein BE (1998) Multisensory integration in the superior colliculus of the alert cat. J Neurophysiol 80:1006–1010

Ward LM (1994) Supramodal and modality-specific mechanisms for stimulus-driven shifts of auditory and visual attention. Can J Exp Psychol 48:242–259

Ward LM, McDonald JJ, Golestani N (1998) Cross-modal control of attention shifts. In: Wright RD (eds) Visual attention. Oxford University Press, New York, pp 232–268