·Original Article·

Spatio-temporal measures of electrophysiological correlates for behavioral multisensory enhancement during visual, auditory and somatosensory stimulation: A behavioral and ERP study

Wuyi Wang^{1,2}, Li Hu³, Hongyan Cui², Xiaobo Xie², Yong Hu^{2,4}

¹School of Precision Instruments and Opto-Electronics Engineering, Tianjin University, Tianjin 300072, China

²Institute of Biomedical Engineering, Chinese Academy of Medical Science and Peking Union Medical College, Tianjin 300192, China

³Key Laboratory of Cognition and Personality (Ministry of Education) and School of Psychology, Southwest University, Chongging 400715, China

⁴Department of Orthopaedics and Traumatology, the University of Hong Kong, Hong Kong SAR Corresponding author: Yong Hu. E-mail: yhud@hku.hk

© Shanghai Institutes for Biological Sciences, CAS and Springer-Verlag Berlin Heidelberg 2013

ABSTRACT

Multisensory enhancement, as a facilitation phenomenon, is responsible for superior behavioral performance when an individual is responding to cross-modal versus modality-specific stimuli. However, the event-related potential (ERP) counterparts of behavioral multisensory enhancement are not well known. We recorded ERPs and behavioral data from 14 healthy volunteers with three types of target stimuli (modality-specific, bimodal, and trimodal) to examine the spatio-temporal electrophysiological characteristics of multisensory enhancement by comparing behavioral data with ERPs. We found a strong correlation between P3 latency and behavioral performance in terms of reaction time (RT) (R = 0.98, P < 0.001), suggesting that P3 latency constitutes a temporal measure of behavioral multisensory enhancement. In addition, a fast RT and short P3 latency were found when comparing the modality-specific visual target with the modality-specific auditory and somatosensory targets. Our results indicate that behavioral multisensory enhancement can be identified by the latency and source distribution of the P3 component. These findings may advance our understanding of the

neuronal mechanisms of multisensory enhancement.

Keywords: multisensory enhancement; temporal measure; P3 latency; visual dominance

INTRODUCTION

Information from different sensory inputs and modalities is often perceived simultaneously in daily life. Accordingly, the brain must continuously combine sensory information to construct a coherent percept of the external world^[1]. Multisensory enhancement refers to a facilitation effect by which responses to cross-modal stimuli are superior to those elicited by modality-specific stimuli. This phenomenon is typically observed during the detection of bimodal and trimodal targets by the visual, auditory, and somatosensory systems. In behavioral studies, the facilitation effect has been documented in terms of decreased reaction time (RT) during target detection^[2]. Concurrently, the underlying mechanisms of multisensory enhancement have been investigated using electrophysiological and functional neuroimaging, including multi-unit neuronal recordings, event-related potentials (ERPs), and functional magnetic resonance imaging (fMRI)^[3, 4]. Compared with modalityspecific targets, the detection of cross-modal targets is associated with a higher spiking frequency in the superior colliculus^[5, 6], increased amplitude and earlier peak latency in some ERP components^[4, 7], and in fMRI studies with a more extensive distribution of activation in both sensoryspecific cortices and regions involved in multisensory integration including the superior temporal sulcus and the posterior parietal cortex^[3, 8].

In terms of behavioral evaluation, several primary models have been adopted to evaluate the processing of multisensory enhancement, including the redundant target effect (RTE)^[9], the race model^[10], and the co-activation model. Todd reported that the RT during target detection is faster when the target is accompanied by other sensoryrelevant stimuli^[11]. Subsequent studies have produced similar findings, and the multisensory enhancement effect has been suggested to be attributable to the RTE^[12]. Furthermore, a faster RT when detecting a cross-modal target in the race model was confirmed to be not simply a result of probability summation (i.e., not solely due to the summed activity of each modality)^[12, 13]. Subsequent studies have proposed a co-activation model^[9], which states that the processing of a target in one modality is influenced by the processing of targets in other modalities. Although probability summation does appear to contribute to behavioral multisensory enhancement, the neuronal response interactions between different modalities are clearly important^[14].

Several previous studies have ascribed the multisensory enhancement effect to two general stages of neuronal signal processing: stimulus-driven, bottom-up, early sensory processes^[13, 15-17] and selection-driven, top-down, late cognitive control^[18, 19]. Although recent studies have primarily focused on the early sensory processing of multisensory integration, late cognitive processing during cross-modal target detection may also play a crucial role in multisensory enhancement^[13, 20, 21]. Busse et al. reported greater brain activity when a taskirrelevant sound stimulus is accompanied by a visual target during a late perceptual selection stage^[20]. In another study, a larger P3 ERP component was elicited during the detection of a bimodal target compared with a modality-specific target, indicating that enhanced awareness processing may be implicated in multisensory enhancement^[22]. Furthermore, the P3 component has been reported to reflect updates in working memory^[23]. Molholm *et al.* also reported that a simultaneous visual-auditory stimulus elicits a more positive P2 than the sum of two modality-specific stimuli^[13]. Given these findings, the P2 and P3 components may represent an electrophysiological index of perceptual selection processing during multisensory enhancement.

Several studies have reported similar behavioral and ERP results from investigations of cross-modal stimulation with combined visual, auditory, and somatosensory target stimuli^[24-27]. Diederich and Colonius reported differences in RTs when participants responded to modality-specific, bimodal, and trimodal targets combining these three types of stimuli^[28]. In that study, the responses to trimodal stimuli were faster than those to bimodal stimuli, while bimodal stimuli elicited faster responses than modality-specific stimuli. In an ERP study of multisensory enhancement, Karns *et al.* provided evidence for intermodal attentional enhancement by comparing early ERP components elicited by attended and unattended visual, auditory, and tactile stimuli^[27].

The multisensory enhancement effect has been investigated at both the behavioral and the electrophysiological levels in terms of vision, audition, and somatosensation. However, the relationship between behavior and ERPs during multisensory enhancement remains unclear, especially in terms of the late ERP components (such as P3), which are elicited in a well-defined manner by target detection^[29].

In a previous study, we explored the dynamic characteristics of multisensory integration using a target detection task, to identify the sensory-specific cognitive processes underlying responses to visual, auditory, and somatosensory targets paired with simultaneous modality-specific and bimodal background stimuli^[30]. In this study, by comparing behavioral data with changes in ERP components (P2 and P3), we examined the temporal and spatial dynamics of neural activation underlying behavioral multisensory enhancement. In addition, we compared the dominance of the visual, auditory, and somatosensory modalities in multisensory processing.

METHODS

Participants

Fourteen right-handed, healthy undergraduate and graduate student volunteers from Tianjin University (eight females and six males) aged 20–24 years (22.3 \pm 0.4, mean \pm SD) participated in this study. All volunteers had normal sensorimotor and hearing abilities, and normal or corrected-to-normal vision. All gave written informed

consent prior to the study. The experimental procedure was compliant with the Helsinki Declaration, and approved by the local Ethics Committee (Institute of Biomedical Engineering, Chinese Academy of Medical Science).

Stimuli

The visual stimuli were adapted from previous studies^[13,31]. The target visual stimulus (V) was a green, 3-cm-diameter, solid circle displayed for 60 ms at a visual angle of 2.29°. The non-target visual stimulus (v) was a red solid circle of the same size, presented in the same way as the target stimulus. The visual stimulus was presented in the center of a 19-inch CRT monitor at a viewing distance of 75 cm. The target auditory stimulus (A) was a high-frequency, pure tone of 2 kHz with a duration of 60 ms and a 5 ms rise/ fall time. The non-target auditory stimulus (a) was a lowfrequency, pure tone of 1 kHz for 60 ms. The target and non-target auditory stimuli had the same intensity of 80 dB, with different tones according to previous studies^[13, 31]. The auditory stimuli were presented binaurally via earphones (Stim Audio System; NeuroScan Lab, Charlotte, NC). The target somatosensory stimulus (S) was a high-intensity electrical stimulus with a constant current of 1.5 ± 0.1 mA, presented for 1 ms. The non-target somatosensory stimulus (s) was a low-intensity electrical stimulus with a constant current of 1.0 \pm 0.1 mA for 1 ms^[32]. The somatosensory stimuli were applied as square-wave pulses on the left index finger via two metal rings. In sum, the difference between target and non-target stimuli was color for visual, frequency for auditory, and intensity for somatosensory. There were seven different types of non-target stimuli, resulting from a combination of the modality-specific, bimodal, and trimodal stimulus conditions (termed a, s, v, as, av, sv, and asv), and there were seven equivalent types of target stimulus (termed A, S, V, AS, AV, SV, and ASV) (Fig. 1). Presentation of the stimuli was controlled by software (STIM system; NeuroScan Lab), which aligned the presentation time for each stimulus.

Each participant was seated in a comfortable chair in a dimly-lit room. They were required to pay full attention to the sensory events, which were presented in random order (Fig. 1). Participants were asked to respond immediately to any target stimulus by pressing a response button using their right index finger (STIM System Switch Response Pad; NeuroScan Lab). Approximately 50% of the trials in each stimulus condition contained no target (i.e., no response was required). Each type of target was presented 72 times over the course of the study. There was a random interval of 1 500–2 500 ms between stimuli. Before formal data collection, participants were shown 200 stimuli, including those of all stimulus types, for familiarization with the experimental task. Altogether, an entire session was composed of six runs, each lasting ~13 min, at ~5-min intervals. One session took ~110 min.

Electroencephalographic Recording

EEG signals were recorded on a 64-channel NeuroScan system (sampling rate, 1 000 Hz; band-pass, 0.05–100 Hz) with a standard 10-20 electrode positioning system EEG cap (NeuroScan Company). The Cz electrode was used as a reference, and the impedances in all channels were <5 k Ω during the experiment. The electro-oculographic (EOG) signal was synchronously monitored using four electrodes to evaluate eye-blinks and ocular movements.

The EEG signals were processed using Scan4.3 software from NeuroScan Lab and EEGLAB^[34], and were band-pass filtered between 0.1 and 30 Hz. The EEG epochs were divided into time-windows of 1 200 ms (200 ms pre-stimulus and 1 000 ms post-stimulus). Baseline correction was performed during the pre-stimulus interval. All epochs contaminated by eye-blinks and eye movements were rejected using automatic artifact rejection processing with Scan4.3. After baseline correction and artifact rejection, all EEG epochs were re-referenced to the double mastoid electrodes. For each participant in each stimulus type, average waveforms were calculated and time-locked to the onset of each event. Average waveforms from single participants were subsequently averaged to acquire group-level average waveforms for each stimulus type.

On the average ERP waveforms from each participant, we measured the peak latencies and baseline-to-peak amplitudes of P2 (150–280 ms, at Cz) and P3 (280–600 ms, at Pz) components for all target conditions^[35]. Because the P3 component of the ERP was previously defined as an index of neuronal processes related to updating working memory, the P3 peak amplitude and latency from each subject were taken to reflect a temporal process of neuronal activation during target detection^[36]. The P2 wave in the central cortex is considered to reflect the general neuronal processing that occurs when a simple sensory target is



Fig. 1. Target stimuli and experimental procedure. Seven target stimuli (A, S, V, AS, AV, SV, ASV) are shown (upper panel). Each column represents one stimulus type. For visual stimuli, the target event (V) was a green flash (solid circle; duration, 60 ms), whereas the non-target event (v) was a red flash (solid circle; duration, 60 ms). For auditory stimuli, the target event (A) was a high-frequency monotonic sound burst (frequency, 2 kHz; duration, 60 ms), whereas the non-target event (a) was a low-frequency monotonic sound burst (frequency, 1 kHz; duration, 60 ms). For somatosensory stimuli, the target event (S) was a higher intensity electrical stimulus (intensity, 1.5 ± 0.1 mA; duration, 1 ms), whereas the non-target event (s) was a lower intensity stimulus (intensity, 1.0 ± 0.1 mA; duration, 1 ms). Both target stimuli and non-target stimuli were presented in random order with the inter-stimulus interval (ISI) varying randomly between 1 500 ms and 2 500 ms (lower panel).

matched with a stored memory representation^[37]. For each stimulus type, the scalp topography at the group-level of P2 and P3 components was computed by spline interpolation. One-way ANOVA was used to compare the peak amplitudes and latencies of P2 and P3 among the seven target conditions. For significant effects, we ran a *post hoc* analysis using Tukey's correction method to compare the amplitudes and latencies in different target conditions.

Source analyses were performed for all seven target conditions using BESA 5.3 (Brain Electrical Source Analysis software; BESA GmbH, Graefelfing, Germany). We used Classical LORETA Analysis Recursively Applied (CLARA) distributed source analysis to locate the sources of the P3 components, with LORETA mean slow resolution brain electromagnetic tomography^[38, 39].

The singular value decomposition was regularized using a cutoff of 0.01%, and a three-iteration scheme was adopted to execute the CLARA source analysis^[38]. The

generators of P3 were reconstructed with grand-averaged ERP data using the electrodes common to all subjects, but removing all deviant electrodes. At the same time, all significant regions were identified based on the maximum intensity in the obtained source volumes. Finally, source locations were transformed into normalized Talairach space using a realistic approximation (adult, cr80) model.

Behavioral Performance

RTs between 200 ms and 1 000 ms were considered acceptable for all participants. The difference in RTs in response to different target stimuli (A, S, V, AS, AV, SV, and ASV) was assessed by one-way repeated-measures analysis of variance (ANOVA). If the difference was statistically significant, a *post-hoc* analysis using Tukey's correction further compared the RTs in the modality-specific, bimodal, and trimodal target conditions.

RT distribution analysis was performed to verify the existence of the RTE. The RT distributions from the trimodal,

bimodal, and modality-specific target conditions were divided into 20 quantiles (5% bin for each subject)^[12, 25]. The race model inequality was used to verify violation of the race model^[33]. Three bimodal inequalities and one trimodal inequality were built according to the race model inequality, while bimodal visual-auditory stimuli presented $p(AV) \le p(A) + p(V)$, bimodal visual-somatosensory stimuli presented $p(SV) \le p(S) + p(V)$, bimodal auditory-somatosensory stimuli presented $p(ASV) \le p(A) + p(V)$, bimodal auditory-somatosensory stimuli presented $p(ASV) \le p(A) + p(V) + p(S)$. By plotting the curve of the cumulative density function (CDF) of the RT for all conditions, these inequalities were tested. If these inequalities were violated, the race model was substituted with a co-activation model.

The correlations of the measured P2 and P3 amplitudes and latencies with the corresponding mean RTs were examined. For each stimulus type, the amplitudes and latencies of the P2 and P3 components were averaged across subjects and compared with the corresponding RTs, which were also averaged across subjects. The correlation coefficients and significance were computed for all comparisons.

RESULTS

Behavioral Results

The RTs in the different target conditions were significantly different (Fig. 2, Table 1). *Post hoc* comparisons revealed



Fig. 2. Comparison of reaction time (left panel), P3 latency (middle), and P3 amplitude (right) among the seven target stimuli. Mean ± SE. **P* <0.05, one-way ANOVA.

Table 1. Behavioral data (reaction time and miss rate), ERP data (peak latencies and amplitudes of P2 and P3), and their comparison (one-way ANOVA) among the seven target conditions

	А	S	V	AS	AV	SV	ASV	F	Ρ	Significantly different pairs
RT (ms)	523 ± 16.2	571 ± 22.6	416 ± 16.8	451 ± 21.7	388 ± 17.5	389 ± 15.5	368 ± 18.8	17	<0.001	I (A-V, AV, SV, ASV)
										(S-V, AS, AV, SV, ASV)
										(AS-ASV)
P2 latency	213 ± 4.0	206 ± 6.4	199 ± 7.4	209 ± 4.8	207 ± 7.0	201 ± 9.0	183 ± 5.5	2.3	0.044	(A-ASV)
(ms)										
P2 amplitude	4.9 ± 1.4	6.3 ± 1.1	4.0 ± 1.3	10.9 ± 2.2	4.3 ± 1.7	6.6 ± 1.3	8.7 ± 2.1	2.3	0.039	
(µV)										
P3 latency	485 ± 21.9	503 ± 21.6	388 ± 10.2	388 ± 16.1	344 ± 8.5	369 ± 8.0	319 ± 6.4	23.3	<0.001	(A-V, AS, AV, SV, ASV)
(ms)										(S-V, AS, AV, SV, ASV)
										(V-ASV) (AS-ASV)
P3 amplitude	9.8 ± 1.2	13.1 ± 1.4	13.0 ± 1.5	13.1 ± 1.3	13.8 ± 1.5	15.0 ± 1.6	15.7 ± 1.6	1.7	0.13	
(µV)										

that the RT in response to S was longer than those to other target stimuli except A (all P < 0.001), and the RT in response to A was longer than those to V, AV, SV, and ASV (V vs A: P = 0.002; others all P < 0.001). Besides, the RT to AS was longer than that to ASV (P = 0.037) (Table 1). These data indicate faster target detection in crossmodal stimulation than in modality-specific stimulation, with the trimodal stimuli being more effective than the bimodal stimuli. These findings are consistent with previous studies^[25].

For the race model test, paired *t*-tests showed that: (1) the bimodal auditory-somatosensory (AS) CDF was greater than the sum of the A and S CDFs in three bins (bins 3–5); (2) the AV CDF was greater than the sum of the A and V CDFs in three bins (2–4); (3) the SV CDF was greater than the sum of the S and V CDFs in two bins (3 and 4); and (4) the trimodal ASV CDF was greater than the sum of the A+S+V CDFs in four bins (1–4) (Fig. 3). Also, the multisensory RT gains in the bimodal and trimodal redundant targets partially exceeded the prediction of the race model (Fig. 3), resulting in a Miller inequality violation^[9]. All those results showed the RT pattern in a co-activation model, demonstrating the multisensory enhancement effect for the bimodal and trimodal redundant targets.

EEG Results

The P2 and P3 latencies were significantly different among the seven target conditions (P2 latency: F = 2.3, P = 0.044; P3 latency: F = 23.3, P < 0.001; one-way ANOVA; Table



Fig. 3. Race model testing for bimodal and trimodal stimuli. The cumulative density functions (CDFs) represent probability scores for modality-specific, bimodal, and trimodal targets. *P <0.05, **P <0.01, paired *t*-tests.



Fig. 4. Group-level average waveforms and scalp topographies (left panel) of P3 in the seven target conditions (A, S, V, AS, AV, SV, ASV) measured at Pz. The gray region shows the time interval in which the ERP peaks were measured; each target condition is presented as one color. Right panel: source analysis results for the P3 component among the seven target conditions. The red, blue, green, and black circles represent the sources from conditions V and A, S and SV, AV and AS, and ASV, respectively, in the precuneus (parietal lobe).

1). Post hoc analyses further revealed a longer P2 latency in response to A than to ASV (P = 0.025). In addition, the P3 latencies in response to S and A were both longer than those to V, AS, AV, SV, and ASV (all P < 0.001). Finally, the P3 latency in response to V was longer than that to ASV (P= 0.017), and the P3 latency in response to AS was longer than that to ASV (P = 0.017; Fig. 2).

Figure 4 shows the group-level average waveforms and scalp topography of P3 across the seven target conditions (A, S, V, AS, AV, SV, and ASV). The scalp topography of P3 (positive activity with a maximum over the parietal region) was similar across the seven target conditions. Overall, our ERP analyses suggested that the P3 latencies among the modality-specific, bimodal, and trimodal conditions were significantly different. The P3 amplitude and scalp topography results did not reveal significant differences among the seven target conditions.

Source analysis

The imaging results from the source analysis of the P3 component in the seven target conditions are shown in Figure 4. The activity was determined to arise from four sources in response to the V and A, S and SV, AV and AS, and ASV targets. All these sources were located in the precuneus (parietal lobe), which has been implicated

in visuospatial processing, episodic memory, and self-reflection^[40]. Furthermore, the sources of the responses to



Fig. 5. Correlation between behavioral data (reaction time) and ERP data (P3 latency). Vertical and horizontal error bars represent, for each condition, the variance across subjects (expressed as SEM). Gray solid line represents the best linear fit. The correlation between RT and P3 latency was significant (R = 0.98, P < 0.001).

visual stimuli were in all four locations, while the sources of the responses to auditory and somatosensory stimuli were only in three of the locations.

Correlation Analysis

Correlation analysis between the behavioral (RT) and ERP results (peak amplitudes and peak latencies of P2 and P3) in all the stimulus conditions revealed a significant correlation only between RT and the P3 latency (R = 0.98, P < 0.001; Fig. 5). Taken together, these results indicate that among the ERP components, the P3 latency exhibited the strongest correlation with behavioral responses across the seven target conditions.

DISCUSSION

The main goal of the present study was to investigate the relationship between behavioral multisensory enhancement and late ERP components, and to describe the spatio-temporal characteristics of behavioral multisensory enhancement. Our results revealed the multisensory enhancement effect at the behavioral level.

Measurement of Multisensory Enhancement

Although the P2 and P3 latencies elicited during multisensory enhancement were affected by different types of sensory input, only the P3 latency was significantly correlated with the behavioral data. The temporal aspects of behavioral multisensory enhancement in response to different stimuli may be related to the P3 latency. The distributions of P3 latencies in response to the seven target conditions were significantly different among the modalityspecific, bimodal, and trimodal conditions, and between the bimodal and trimodal conditions. Since the P3 latency is considered to be a measure of classification speed for a stimulus^[41, 42], it is likely that it is also correlated with the behavioral data. These behavioral data varied among the seven target conditions only as a result of stimulus-related differences, as a single response button was pressed by each subject in all target conditions when they detected a target stimulus. The significant correlation between the behavioral results and the P3 latency indicated that P3 latency may serve as a temporal measure of the neuronal processing underlying behavioral multisensory enhancement.

Cognitive Resource Allocation and Visual Dominance

In addition to the temporal aspects of behavioral multisensory enhancement, we examined the spatial distribution of P3 sources in the brain among the different target conditions using BESA 5.3, and revealed four source locations (Fig. 4). The locations of the modality-specific, bimodal, and trimodal P3 sources during multisensory enhancement in this study are consistent with the findings of previous fMRI studies^[43]. For instance, responses to the modality-specific stimuli (V, A, and S) had sources close to the posterior precuneus, responses to the bimodal stimuli (SV, AV, and AS) had sources in the middle of the precuneus, and responses to the trimodal stimulus (ASV) had sources in the anterior precuneus^[3, 44]. Previous studies have reported that the precuneus is involved in multiple brain functions, including multisensory attention^[45] and multisensory-spatial processing^[46]. Here, we found the sources of responses to visual stimuli over four locations, while the sources of responses to auditory and somatosensory stimuli were in only three locations. These data suggested that neuronal resources are allocated more extensively when the brain detects visual targets than auditory and somatosensory targets^[26, 47, 48]. Our results demonstrated a general visual dominance over touch and audition in terms of the spatial distribution of P3 sources.

ACKNOWLEDGEMENTS

This work was partly supported by the National Natural Science Foundation of China (81271685), Key Project of Science and Technology Fund of Tianjin Municipality, China (10JCZDJC16100), the China Postdoctoral Science Foundation (2012M510754), and the Independent Innovative Fund of Tianjin University, Tianjin Municipality, China (1102). We thank Edanz Editing China for language editing.

Received date: 2012-11-25; Accepted date: 2013-04-19

REFERENCES

- Stein BE, Stanford TR. Multisensory integration: current issues from the perspective of the single neuron. Nat Rev Neurosci 2008, 9: 255–266.
- [2] Lovelace CT, Stein BE, Wallace MT. An irrelevant light enhances auditory detection in humans: a psychophysical analysis of multisensory integration in stimulus detection. Brain Res Cogn Brain Res 2003, 17: 447–453.

- [3] Macaluso E, Driver J. Multisensory spatial interactions: a window onto functional integration in the human brain. Trends Neurosci 2005, 28: 264–271.
- [4] Murray MM, Molholm S, Michel CM, Heslenfeld DJ, Ritter W, Javitt DC, et al. Grabbing your ear: Rapid auditorysomatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. Cereb Cortex 2005, 15: 963–974.
- [5] Stein BE, Meredith MA. The merging of the senses. Cambridge, MA: MIT Press, 1993.
- [6] Stein BE, London N, Wilkinson LK, Price DD. Enhancement of perceived visual intensity by auditory stimuli: A psychophysical analysis. J Cogn Neurosci 1996, 8: 497–506.
- [7] McDonald JJ, Teder-Salejarvi WA, Di Russo F, Hillyard SA. Neural basis of auditory-induced shifts in visual time-order perception. Nat Neurosci 2005, 8: 1197–1202.
- [8] Sedda A, Scarpina F. Dorsal and ventral streams across sensory modalities. Neurosci Bull 2012, 28: 291–300.
- [9] Miller J. Divided attention: evidence for coactivation with redundant signals. Cogn Psychol 1982, 14: 247–279.
- [10] Raab DH. Statistical facilitation of simple reaction times. Trans N Y Acad Sci 1962, 24: 574–590.
- [11] Todd JW. Reaction to Multiple Stimuli. pp. iii. 65. 1912.
- [12] Sinnett S, Soto-Faraco S, Spence C. The co-occurrence of multisensory competition and facilitation. Acta Psychol (Amst) 2008, 128: 153–161.
- [13] Molholm S, Ritter W, Murray MM, Javitt DC, Schroeder CE, Foxe JJ. Multisensory auditory-visual interactions during early sensory processing in humans: a high-density electrical mapping study. Brain Res Cogn Brain Res 2002, 14: 115– 128.
- [14] Sperdin HF, Cappe C, Foxe JJ, Murray MM. Early, low-level auditory-somatosensory multisensory interactions impact reaction time speed. Front Integr Neurosci 2009, 3: 2.
- [15] Senkowski D, Saint-Amour D, Hofle M, Foxe JJ. Multisensory interactions in early evoked brain activity follow the principle of inverse effectiveness. Neuroimage 2011, 56: 2200–2208.
- [16] Rowland BA, Stein BE. Multisensory integration produces an initial response enhancement. Front Integr Neurosci 2007, 1: 4.
- [17] Giard MH, Peronnet F. Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. J Cogn Neurosci 1999, 11: 473– 490.
- [18] Diaconescu AO, Alain C, McIntosh AR. The Co-occurrence of multisensory facilitation and crossmodal conflict in the human brain. J Neurophysiol 2011, 106(6): 2896–2909.
- [19] McDonald JJ, Teder-Salejarvi WA, Di Russo F, Hillyard SA. Neural substrates of perceptual enhancement by crossmodal spatial attention. J Cogn Neurosci 2003, 15: 10–19.

- [20] Busse L, Roberts KC, Crist RE, Weissman DH, Woldorff MG. The spread of attention across modalities and space in a multisensory object. Proc Natl Acad Sci U S A 2005, 102: 18751–18756.
- [21] Talsma D, Senkowski D, Soto-Faraco S, Woldorff MG. The multifaceted interplay between attention and multisensory integration. Trends Cogn Sci 2010, 14: 400–410.
- [22] Van der Burg E, Talsma D, Olivers CN, Hickey C, Theeuwes J. Early multisensory interactions affect the competition among multiple visual objects. Neuroimage 2011, 55: 1208–1218.
- [23] Sutton S, Braren M, Zubin J, John ER. Evoked-potential correlates of stimulus uncertainty. Science 1965, 150: 1187– 1188.
- [24] Bresciani JP, Dammeier F, Ernst MO. Tri-modal integration of visual, tactile and auditory signals for the perception of sequences of events. Brain Res Bull 2008, 75: 753–760.
- [25] Gondan M, Roder B. A new method for detecting interactions between the senses in event-related potentials. Brain Res 2006, 1073-1074: 389–397.
- [26] Hecht D, Reiner M. Sensory dominance in combinations of audio, visual and haptic stimuli. Exp Brain Res 2009, 193: 307–314.
- [27] Karns CM, Knight RT. Intermodal auditory, visual, and tactile attention modulates early stages of neural processing. J Cogn Neurosci 2009, 21: 669–683.
- [28] Diederich A, Colonius H. Bimodal and trimodal multisensory enhancement: effects of stimulus onset and intensity on reaction time. Percept Psychophys 2004, 66: 1388–1404.
- [29] Zhu WN, Zhang JJ, Liu HW, Ding XJ, Ma YY, Zhou CL. Differential cognitive responses to guqin music and piano music in Chinese subjects: an event-related potential study. Neurosci Bull 2008, 24: 21–28.
- [30] Wang W, Hu L, Valentini E, Xie X, Cui H, Hu Y. Dynamic characteristics of multisensory facilitation and inhibition. Cogn Neurodyn 2012, 6(5): 409–419.
- [31] Feng T, Qiu Y, Zhu Y, Tong S. Attention rivalry under irrelevant audiovisual stimulation. Neurosci Lett 2008, 438: 6–9.
- [32] Tanaka E, Inui K, Kida T, Miyazaki T, Takeshima Y, Kakigi R. A transition from unimodal to multimodal activations in four sensory modalities in humans: an electrophysiological study. BMC Neurosci 2008, 9: 116.
- [33] Ulrich R, Miller J, Schroter H. Testing the race model inequality: an algorithm and computer programs. Behav Res Methods 2007, 39: 291–302.
- [34] Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 2004, 134: 9–21.
- [35] Luck SJ. An introduction to the event-related potential technique. Cambridge, MA: MIT Press, 2005.
- [36] Polich J, Alexander JE, Bauer LO, Kuperman S, Morzorati

S, O'Connor SJ, *et al.* P300 topography of amplitude/latency correlations. Brain Topogr 1997, 9: 275–282.

- [37] Freunberger R, Klimesch W, Doppelmayr M, Holler Y. Visual P2 component is related to theta phase-locking. Neurosci Lett 2007, 426: 181–186.
- [38] Hoechstetter K, Berg P, Scherg M. BESA Research Tutorial 4: Distributed Source Imaging. 2010.
- [39] Hamalainen JA, Ortiz-Mantilla S, Benasich AA. Source localization of event-related potentials to pitch change mapped onto age-appropriate MRIs at 6months of age. Neuroimage 2011, 54(3): 1910–1918.
- [40] Dorfel D, Werner A, Schaefer M, von Kummer R, Karl A. Distinct brain networks in recognition memory share a defined region in the precuneus. Eur J Neurosci 2009, 30: 1947–1959.
- [41] Polich J. Attention, probability, and task demands as determinants of P300 latency from auditory stimuli. Electroencephalogr Clin Neurophysiol 1986, 63: 251–259.
- [42] Kutas M, McCarthy G, Donchin E. Augmenting mental chronometry: the P300 as a measure of stimulus evaluation time. Science 1977, 197: 792–795.

- [43] Driver J, Noesselt T. Multisensory interplay reveals crossmodal influences on 'sensory-specific' brain regions, neural responses, and judgments. Neuron 2008, 57: 11–23.
- [44] Bremmer F, Schlack A, Shah NJ, Zafiris O, Kubischik M, Hoffmann K, et al. Polymodal motion processing in posterior parietal and premotor cortex: a human fMRI study strongly implies equivalencies between humans and monkeys. Neuron 2001, 29: 287–296.
- [45] Anderson JS, Ferguson MA, Lopez-Larson M, Yurgelun-Todd D. Topographic maps of multisensory attention. Proc Natl Acad Sci U S A 2010, 107: 20110–20114.
- [46] Renier LA, Anurova I, De Volder AG, Carlson S, VanMeter J, Rauschecker JP. Multisensory integration of sounds and vibrotactile stimuli in processing streams for "what" and "where". J Neurosci 2009, 29: 10950–10960.
- [47] Koppen C, Levitan CA, Spence C. A signal detection study of the Colavita visual dominance effect. Exp Brain Res 2009, 196: 353–360.
- [48] Van Damme S, Crombez G, Spence C. Is visual dominance modulated by the threat value of visual and auditory stimuli? Exp Brain Res 2009, 193: 197–204.