

SCIENTIFIC REPORTS



OPEN

Semantic congruent audiovisual integration during the encoding stage of working memory: an ERP and sLORETA study

Yuanjun Xie, Yuanyuan Xu, Chen Bian & Min Li

Although multisensory integration is an inherent component of functional brain organization, multisensory integration during working memory (WM) has attracted little attention. The present study investigated the neural properties underlying the multisensory integration of WM by comparing semantically related bimodal stimulus presentations with unimodal stimulus presentations and analysing the results using the standardized low-resolution brain electromagnetic tomography (sLORETA) source location approach. The results showed that the memory retrieval reaction times during congruent audiovisual conditions were faster than those during unisensory conditions. Moreover, our findings indicated that the event-related potential (ERP) for simultaneous audiovisual stimuli differed from the ERP for the sum of unisensory constituents during the encoding stage and occurred within a 236–530 ms timeframe over the frontal and parietal-occipital electrodes. The sLORETA images revealed a distributed network of brain areas that participate in the multisensory integration of WM. These results suggested that information inputs from different WM subsystems yielded nonlinear multisensory interactions and became integrated during the encoding stage. The multicomponent model of WM indicates that the central executive could play a critical role in the integration of information from different slave systems.

The integrated information derived from different sensory modalities is indispensable for providing unified perceptions¹. Multisensory integration is the process by which information from different sensory systems is combined by nervous system units², and it considerably improves accuracy³, reduces reaction times^{4,5}, and enhances precision^{6–8}. The enhancement in behavioural performance in response to multisensory inputs lies in the convergence of information rather than in the simple redundancy of targets⁹ because multisensory stimuli increase the firing rates of cells to a level that exceeds the rate predicted by the summation of responses to unisensory stimuli^{10,11}.

Certain basic principles underlie the integration effect^{12,13}. Temporally and spatially aligned sensory inputs in different modalities contribute to further processing and result in behavioural facilitations^{14–16}, and higher level factors, such as semantics or semantic congruence, are capable of influencing the integration of information across sensory modalities. In particular, semantic congruence refers to combinations of auditory and visual stimuli that are presented in terms of their meaning¹⁷. By assessing the consequences of presenting matching (the image and sound belong to the same object) or mismatching (the image and sound belong to different objects) information, judgements on the congruence or incongruence of stimuli pairings can be made. The impact of semantics on multisensory integration has been investigated in several studies^{18,19}, and significantly faster reaction times were found for semantically congruent audiovisual pairings compared with unisensory stimulation and significantly longer response times were found for semantically mismatched conditions. Thus, semantic congruence is a critical factor for determining multisensory behavioural performance¹⁸.

Additionally, the neocortex has been found to be necessary for multisensory integration to increase the efficiency of the underlying processes. Several higher-order association cortical areas were considered multisensory, including the intraparietal complex, the superior temporal sulcus, and the frontal cortex^{20–23}. Intracranial

The department of military psychology, College of Psychology, the Third Military Medical University, Chongqing, +86023730030, China. Correspondence and requests for materials should be addressed to M.L. (email: limin52267@tmmu.edu.cn)

recoding and neuroimaging studies in humans demonstrated that multisensory inputs were indeed co-localized to regions of the parietal lobe^{22, 24–26}, including the inferior parietal sulcus, the inferior parietal lobule, and the superior parietal lobule.

Previous studies have demonstrated the benefits of multisensory perception and characterized the properties of multisensory integration by manipulating spatial, temporal and semantic signals from different modalities. However, the impact of multisensory integration on learning and memory have received limited attention²⁷. Recently, the contribution of cross-modal information to perceptual learning has been studied, and subjects trained with congruent audiovisual stimuli showed an advantage when learning visual tasks compared with subjects trained only with visual stimuli^{28, 29}. In addition, researchers have investigated the effect of multisensory experiences on subsequent unisensory memory performance with a continuous recognition task^{3, 30–32}, and they found that recognition memory for repeated pictures was improved when the picture was initially presented with semantically congruent sounds compared to when it was presented with semantically inconsistent sounds. Similarly, subsequent studies on the relationship between sounds and recognition tasks showed that improved recognition occurred after an initially congruent picture-sound pairings^{33–35}.

Electrophysiological studies from a group of researchers demonstrated an incidental effect of multisensory memory on subsequent visual stimuli processing, and they showed that the rapid differentiation of repeated visual stimuli occurred according to the subjects' multisensory (auditory-visual) or unisensory (visual only) experiences approximately 60–136 ms post-stimulus onset³⁰. The effect was observed even when initial multisensory experiences consisted of images and meaningless sounds³². Regions within the lateral-occipital complex were more active in response to visual stimuli associated with multisensory rather than unisensory experiences³¹. These results suggested that multisensory interactions continued to affect subsequent memory recognition³⁶.

Furthermore, several studies in humans and primates indicated the presence of multisensory facilitation in working memory (WM). Heikkilä and colleagues³⁷ examined the effect of audiovisual encoding in WM on later unisensory memory recognition. The participants were instructed to memorize auditory or visual stimuli that co-occurred with either a congruent stimulus, an incongruent stimulus or a neutral stimulus combinations during the encoding stage. Subsequent memory performance was improved overall when the sound stimulus was initially paired with a semantically congruent picture than a neutral stimulus. In addition, a primate behaviour study³⁸ trained monkeys to learn auditory stimuli and similar visual stimuli in a delay matching-to-sample task and then tested the animals with a concurrent audiovisual matching task with equal proportions of auditory, visual, and audiovisual stimuli. Consistent with outcomes in human studies, the study found that the accuracy was higher and response times were faster for audiovisual test stimuli than those for either of the unisensory test stimuli.

Although the above findings indicate that multisensory objects processing impacts later memory performance and brain activity, the mechanisms underlying multisensory processes are not unambiguous. For multisensory experiences in WM, the influence of encoding information from different subsystems on subsequent memory performance is poorly understood and whether the information can be integrated during the encoding stage remains unclear.

The purpose of the present study was to investigate the integration effect of multisensory WM via an event-related potential (ERP) approach, and the addition criterion was applied to examine the integration effect. ERPs in response to a singular auditory stimulus (A) and visual stimulus (V) were summed and compared with the ERPs in response to simultaneously presented audiovisual stimuli (AV). If the neural response to AV differed with the sum of A and V ($ERP_{AV} \neq ERP_A + ERP_V$), then multisensory integration was thought to have occurred. This approach has been repeatedly applied in multisensory research on ERPs in humans^{39–41}. We predicted that multisensory integration would be formed in the encoding stage of WM and influence later memory recognition. Finally, cortical sources of the integration effect in multisensory WM were modelled by the standardized low-resolution brain electromagnetic tomography software (sLORETA)⁴².

Results

Behavioural results. A significant main effect of the modalities was observed on the mean reaction times (RTs) [$F_{2, 35} = 24.31$, $p < 0.001$, $\eta^2 = 0.42$]. Post hoc comparisons revealed that the RTs for memory retrieval under the AV condition were faster than those under the unimodal conditions (A and V). However, the accurate response rate (ACRs) for memory recognition reached a ceiling in all trial types, and significant effects were not observed during these conditions [$F_{2, 35} = 2.13$, $p > 0.05$, $\eta^2 = 0.07$]. The details of the behavioural analysis are shown in Fig. 1B and C.

ERP results. The difference in potentials between ERP_{AV} and ERP_{A+V} via Mass Univariate ERP Toolbox (MUET) was examined to determine the emergence of the integration effect during the multisensory encoding stage of WM. The smallest significant t-score was 2.03, which corresponded to a test-wise alpha level of 0.05. The results showed that a super-additivity effect ($ERP_{AV} > ERP_{A+V}$) of multisensory information occurred during the encoding stage of WM within 236–530 ms after stimulus onset over the parieto-occipital electrodes and the frontal electrodes (beginning at 392 ms) (Fig. 2). The overall average ERP waveforms at six primary electrodes under AV and A + V conditions are displayed in Fig. 3. The mean amplitudes of specific electrodes significantly differed between the AV and A + V conditions over a latency period of 236–530 ms (Fz, 392–530 ms; the smallest $F_{1, 35} = 10.82$, $p < 0.01$, $\eta^2 = 2.24$). The scalp topography distributions of the three conditions (AV, A + V, and Addition) showed that audiovisual interactions primarily activated the posterior brain regions, with the frontal areas recruited in the processes at a relatively later stage.

sLORETA results. Figure 4 showed the sLORETA brain maps that represent neural generators of scalp voltages within the temporal window (236–530 ms) of the significant integration effect. Several brain areas were revealed, including the frontal (inferior frontal gyrus, BA47; middle frontal gyrus, BA9), parietal (inferior parietal lobule,

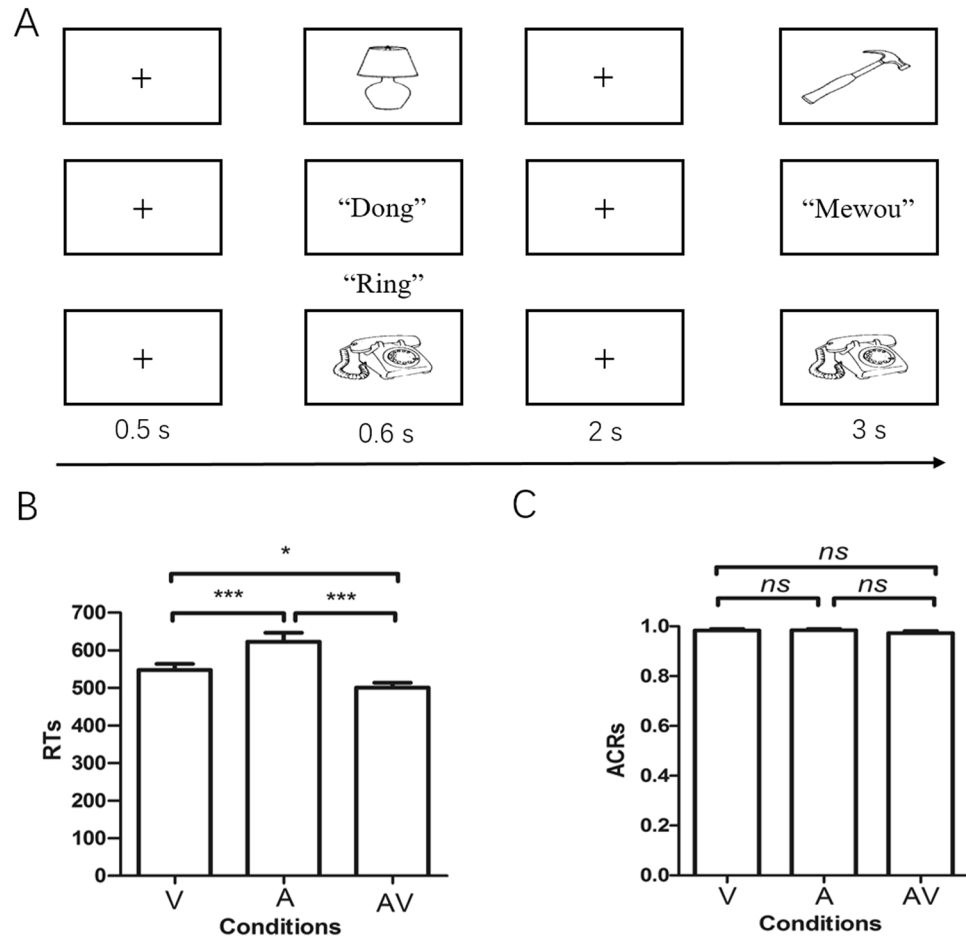


Figure 1. Behavioural data results. (A) Three-block design of the behavioural experiment. In each trial with three blocks, a fixation cross was shown for 0.5 s, and then a stimulus (visual, auditory, and congruent audiovisual) was presented, a blank screen was shown after a 2 s delay, and finally the test stimulus appeared within a 3 s time limit. (B) Corrected mean RTs for memory retrieval during the three conditions. (C) Mean ACRs for memory recognition in the three blocks. Error bars represent the standard error of the mean. Note: * $p < 0.05$, *** $p < 0.001$. *ns*, no significant difference.

BA39/BA40; precuneus, BA7/BA31), temporal (middle temporal gyrus, BA21; superior temporal gyrus, BA41), and occipital (middle occipital gyrus, BA19) areas. The maximum sources were found in the inferior parietal lobule.

Discussion

The present study investigated the integration effect of multisensory WM via an ERP analysis and sLORETA source location method. Behaviourally, the mean RTs for memory recognition during the semantically congruent audiovisual trial type were faster than the mean RTs of both unisensory trial types. Moreover, the ERP for AV stimuli observed in the encoding stage of WM differed from the sum of the ERPs for the A and V stimuli and appeared within 236–530 ms after stimuli onset. The typical outcome of multisensory integration is behavioural facilitation and differences in brain activity between multisensory stimuli and the summed unisensory signals⁴³. The results of the present study were consistent with this proposition and suggested that the multisensory representation of WM was established during the encoding stage and later to facilitate memory retrieval. The neural origins of the integration effect in the encoding stage of WM involve distributed brain networks. This finding is consistent with the results of previous multisensory research which multiple distributed networks were implicated in multisensory integration⁴⁴ and support the opinion that WM processing involves distributed neuronal networks⁴⁵.

The organization of information within WM must be discussed in order to completely understand WM processing from a multisensory perspective⁴⁶. The feature binding of WM is an important issue that indicates whether each feature of an object is arranged separately. Multiple studies have shown that the processing of colour and shape features or an object's identity and location is automatically connected^{47–50}. Other studies have demonstrated that the binding of features can even occur across sensory modalities^{51,52}. The hippocampus also appears to play an important role in the process of binding⁵³ (but see ref. 54). To determine how information from different subsystems can be integrated based on mutual basic codes, the episodic buffer has been proposed as a new component of WM⁵⁵, and it combines information presented in the different storage systems into a complex representation that depends on executive resources. However, further investigation showed that the episodic buffer is a passive domain-general buffer for the storage of information from a visuospatial sketchpad and a phonological

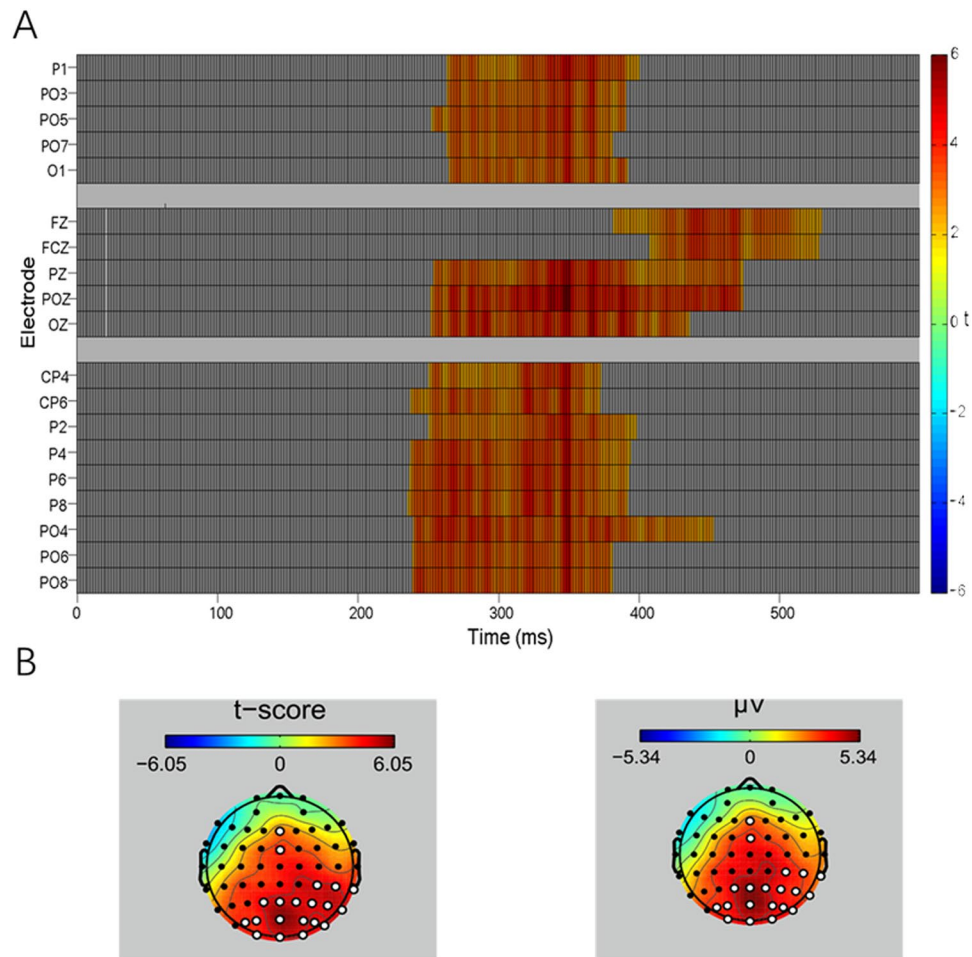


Figure 2. Point-wise cluster test of the integration effect with MUET. (A) Significant clusters in time points and electrodes (only significantly different electrodes are displayed) that occurred in the time window of approximately 236–530 ms. (B) Distributions of significant electrodes in the scalp topography (white dash).

loop according to evidence from the cross-modal binding of WM, because binding the features of colour and shape to objects across auditory and visual modalities appears to be automatic and does not require explicit attentional control⁵⁶. Thus, although the central executive (CE) was initially considered to be a supervisory attention system, we propose that it may play a crucial role in the integration of information from different code systems. In other words, the CE involved in attention control and monitors various slave systems in a coordinated manner⁵⁷, and it may also provide an integration function that merges the information derived from these subsystems into a temporary unitary representation. Because integration or the capability of performing two tasks simultaneously is regarded as a necessary function of the CE⁵⁸. For instance, Alzheimer's patients with a defective executive function show an apparent deterioration in the coordination of the operation of two slave systems: the phonological loop and the visuospatial sketchpad^{59,60}.

The neural substrates of CE were originally assumed to be located in the frontal cortex because patients or primates with lesions in those brain regions frequently demonstrate impaired performance on a range of tasks that assess executive functions^{61–63}. However, evidence also indicates that the CE component of WM is distributed between anterior and posterior brain areas⁶⁴. Therefore, the CE not only recruits frontal areas but also relies on posterior parietal regions^{65–68}.

By all accounts, the CE could be separated into distinct sub-functions that are performed by discrete cortical regions. Hence, the integration function was assumed to occur in the posterior parietal cortex and attention control was assumed to occur in the frontal cortex. These assumptions were supported by the notion that attention and information representations of WM were assigned to the frontal and parietal regions⁶⁹. The two functions of the CE jointly control multisensory processing in WM. Specifically, the posterior parietal cortex is known to play a role in cognition functions^{70,71}, and this cortex can be considered a hub of multisensory information^{44,71–74} that combines information on different sensory modalities to form a unified sensory space⁷⁵. In particular, the inferior parietal lobule is a key location in the brain that is extensively connected with the auditory, visual, and somatosensory cortices. The neurons in this lobule are multimodal because they can process different types of stimuli simultaneously. This multisensory association of traits makes this lobule an ideal candidate for the integrating-information from different modalities. Therefore, according to the classic theory of hierarchical processing, it is possible that auditory and visual information of WM is initially processed within a specialized phonological loop

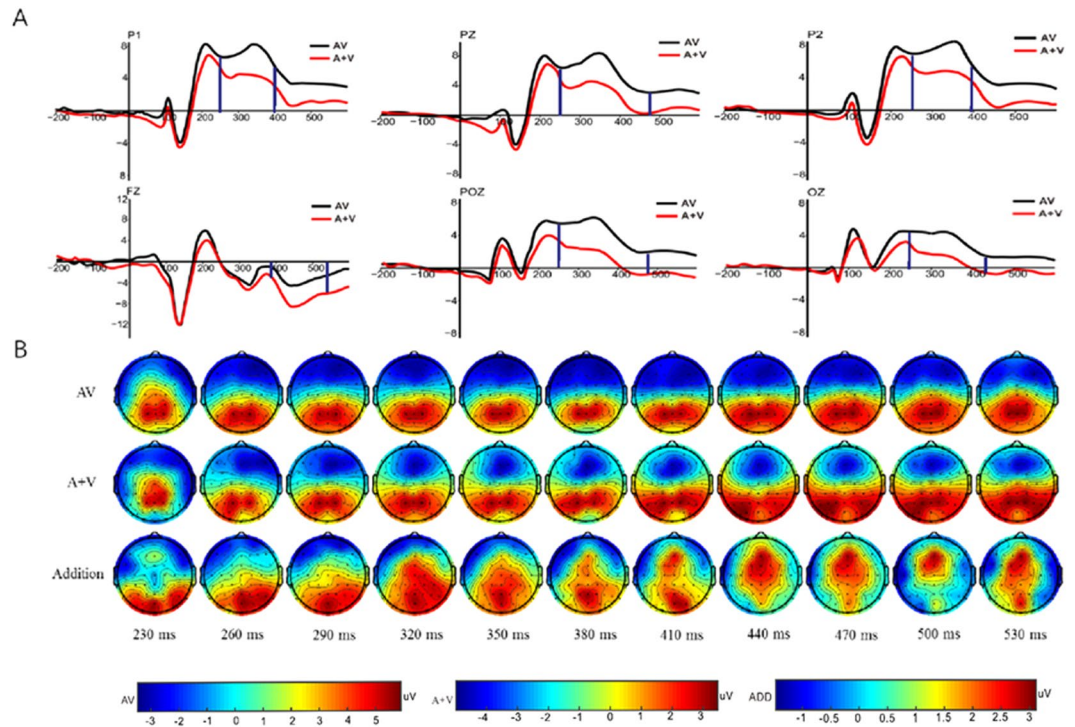


Figure 3. (A) Grand average ERP waveforms under the AV and A + V conditions at the six primary electrodes (P1, PZ, P2, FZ, POZ, and OZ); substantial differences were observed between the AV and A + V conditions for these electrodes over a latent period of 236–530 ms (between the two vertical blue lines). (B) Scalp topography distributions for the steps of instantaneous amplitude values during the AV, A + V, and addition[ADD, AV – (A + V)] conditions.

and a visuospatial sketchpad⁷⁶, and then these subcomponents of WM from different code systems are integrated in the posterior parietal cortex such as interior parietal lobule through the coordination of CE, thereby forming a unified multisensory presentation.

In addition, the results from the point-wise t-test and the scalp topography distributions of ERPs showed that frontal electrodes or areas participated in the integration effect at approximately 400 ms after stimulus onset. This finding may suggest that the frontal cortex exerts top-down attention control at a relatively later stage. This action of the frontal cortex could contribute to the monitoring of information integration because both the inferior frontal gyrus and the middle frontal gyrus involve in attention control^{77,78}. The modulation of inferior frontal gyrus excitability via rTMS induces a significant improvement in attentional task performance⁷⁹, and the middle frontal gyrus is also engaged in a sustained mnemonic response in high-demand WM tasks⁸⁰. The information integration appearing in the posterior parietal cortex is supervised via top-down attention control of the frontal cortices to facilitate retrieval in the future.

Early ERP studies in humans demonstrated that brain responses differ between paired multisensory stimuli and unisensory stimuli presented alone starting within the first 100 ms after stimulus onset^{81–83}, and they showed that primary sensory cortices were potential source of these differences^{84–86}. The early multisensory interactions are strongly related to the stimulus input characteristics and could reflect bottom-up processing because the presence of such interactions can be retained despite attentional suppression⁸⁷. Thus, early multisensory interactions are a hallmark of bottom-up multisensory processes⁴³ and may be automatic^{88,89}. However, these findings differed from the result of the present study that semantic integration of multisensory WM occurred at a relatively later stage, because multisensory integration involving semantic stimuli reflects the influence of top-down processing^{90,91} and may happen in the relatively later. For instance, the first convergence of audiovisual speech (letters and sounds) has been found to occur at latencies of approximately 225 ms and present delays between 380 and 450 ms after stimulus onset⁹². This result is consistent with the classical view that the integration of senses occurs at a later stage of processing after the initial unisensory information has been processed⁹¹.

In summary, the information from WM subsystems could be integrated into unified memory traces during the encoding stage, thereby facilitating memory retrieval. The multisensory integration of WM involves the coordination of distributed brain networks. The integration effect of WM occurs in the posterior parietal cortex, and the frontal cortices subsequently exert attention control over the information integration.

Methods

Participants. Thirty-eight neurologically normal paid volunteers (all right handed; aged 20–24 years old, mean = 23.16 years old; 22 males and 16 females) who reported having normal hearing and normal or corrected to normal vision participated in this study. Two participants (males) were excluded from the analysis because of

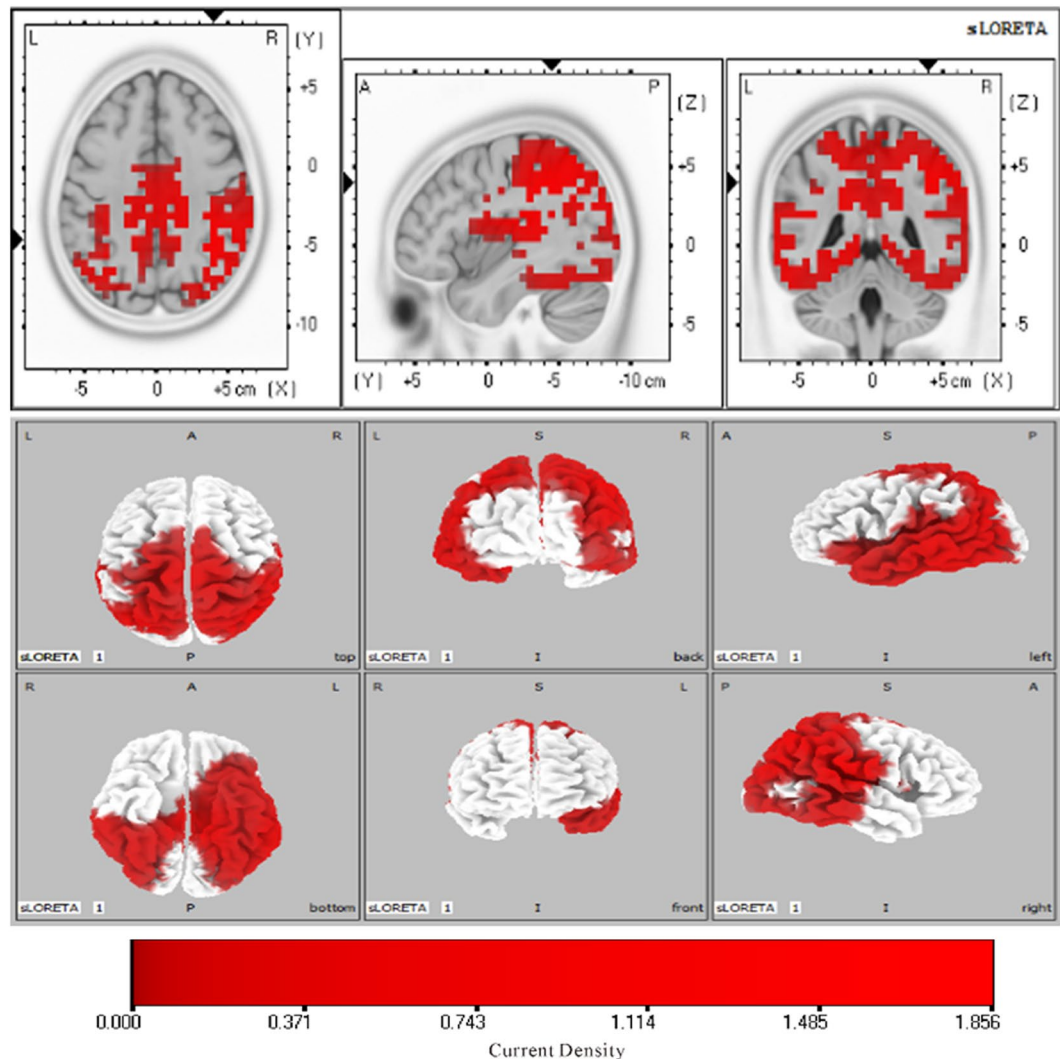


Figure 4. sLORETA images for the integration effect of multisensory WM over a timeframe of 236–530 ms (slice view and 3D view). The highest level of activation was localized in the inferior parietal lobule (BA 39/40).

extensive oculomotor artefacts. This study was approved by the Ethics Committee of the Third Military Medical University, China. All participants provided written informed consent prior to the start of the experiment. All experimental methods were conducted in accordance with the ethical guidelines determined by the National Ministry of Health, Labour and Welfare and the Declaration of Helsinki (BMJ 1991; 302:1194).

Stimuli. Visual stimuli were obtained from the standard set of outline drawing figures^{93,94}. The selected pictures contained an equivalent number of objects from different semantic categories (e.g., animals, tools, instruments, vehicles, etc.) and were divided equally among experimental conditions. The auditory stimuli consisted of verbalizations that corresponded to the visual stimuli. For example, the sound of a cat meowing was paired with the picture of a cat. All of the sound files were downloaded from a website (<http://www.findsounds.com>) and modified with audio editing software (Adobe Audition version 5.0) according to the following parameters: 16 bit; 44,100 Hz digitization; and 0.6 s sound duration with a 10 ms linear amplitude enveloping at sound onset and offset to avoid clicks.

In the present study, we only considered the semantic matching of AV stimuli relative to unisensory stimuli because prior studies (as mentioned in the Introduction) demonstrated that in many cases, a positive effect on subsequent memory recognition was only observed for semantically congruent and not incongruent AV presentations. Thus, three stimulus types were employed: auditory unimodal, visual unimodal, and congruent audiovisual (see appendix A). During the congruent AV condition, the recognition stimuli were only visual objects, which was similar to the condition in previous studies.

Experimental procedures. Participants sat in a quiet and comfortable room and performed a delay-matched WM task during the three experimental blocks (A, V, and AV), which each contained 30 trials. The details of each trial in the three blocks are displayed in Fig. 1A. All three stimuli types were presented at the

encoding stage for a duration of 0.6 s, which was followed by 2 s of delay, and then the probe stimulus appeared until the participant responded within the 3 s time limit. The visual stimuli were presented on a 17-inch computer monitor that subtended the visual angle by approximately 6.5 degrees and had a black background, and semantically related sounds were delivered binaurally at an intensity level of 70 dB via earphones. Congruent AV stimuli were simultaneously presented. The inter-trial interval ranged from 1500 to 3000 ms, and blocks were randomly varied between participants. The participants were asked to judge whether the probe stimulus was same or not with the stimuli presenting during the encoding stage with a key response (Yes and No responses corresponding to the "F" and "K" keys on the keyboard), and presented and un-presented probe stimuli were split in half. Breaks were encouraged between blocks to maintain high concentration and prevent fatigue. At the end of each block, the percentage of correct responses was displayed on the monitor. The stimulus delivery and behavioural responses recordings were controlled using E-prime 2.0 (Psychology Software Tools, Inc., Pittsburgh, PA; <http://www.pstnet.com/eprime>).

EEG acquisition and preprocessing. A high-density electroencephalography (EEG) recording was acquired with a QuickAmp amplifier (Compumedics USA, El Paso, TX, USA) using 64 Ag/AgCl electrodes (extended 10–20 system). Horizontal and vertical electrooculograms were recorded with four bipolar electrodes placed on the outer canthus of each eye and above and below the right eye. Reference electrodes were placed on the bilateral mastoids. The electrode impedances were kept below 5 K Ω on average. The EEG activity was amplified using 0.01–100 Hz band-passed filters and sampled at 1000 Hz. EEG data were pre-processed with MATLAB 2013b (MathWorks, USA) and the EEGLAB13.4.4b toolbox⁹⁵. EEG data at each electrode were re-referenced to the average of the left and right mastoids before further analysis. Then, the signal was digitally filtered off-line with a band-pass filter of 0.1–100 Hz. Time windows of 200 ms before and 600 ms after the onset of stimulus were segmented from the EEG (only focused on the encoding stage of WM). Epochs with amplitude values exceeding ± 100 μ V at any electrode and those containing blinks, eye movements, or incorrect responses were rejected and excluded from the analysis. In addition, the datasets analyzed during the current study are available from the corresponding author on reasonable request.

Data analysis

Behavioural analysis. RTs were recorded on-line for the three types of trials. Only RT values associated with correct responses were considered for the data analysis. Accuracy rates were calculated as the percentage of correct responses (correct hit and correct rejection). Further statistical analyses were conducted using IBM SPSS software V22 (IBM Corp., Armonk, NY, USA).

ERP analysis. Because of the paucity of prior knowledge on defined latencies and locations for the integration effect in multisensory WM, point-wise comparisons between the ERP_{AV} and ERP_{A+V} from each electrode and time point were performed to detect when and where the integration effect occurring by the MUE⁹⁶. A cluster-based permutation test was used to control the family-wise error rate based on the statistical significance and the proximity of nearby time points and electrodes into clusters. This approach is probably the most powerful mass univariate procedure for the detection of broadly distributed effects⁹⁷. The threshold of the test-wise alpha value for cluster inclusion was set at the 0.05. The number of permutations in the present test was 5000. Pairs of time values specified the beginning and end of a time window at approximately 0–600 ms.

sLORETA analysis. Source localization was conducted to explore the originator in the integration effect of WM by sLORETA. sLORETA resolves the inverse problem by assuming synchronous and simultaneous activation of neighbouring neurons⁹⁸. The localization accuracy of sLORETA has been validated by combining LORETA with tomographic techniques, such as EEG/fMRI⁹⁹. In sLORETA, the intracerebral volume is partitioned in 6239 voxels at a 5 mm spatial resolution, and the standardized current density at each voxel is then calculated in a realistic head model. In the present study, the source location was investigated using voxel-wise randomization tests with 5000 permutations based on statistical non-parametric mapping¹⁰⁰. Voxels with significant differences (corrected p-value < 0.01) were located in the MNI brain and Brodmann areas provided by the software.

Statistics. A repeated measures analysis of variance was used to analyse the behavioural and ERP data. The condition (A, V, and AV) and electrode (only for ERP data) factors were used as within-subject factors. A Greenhouse-Geisser correction was applied to adjust the degrees of freedom. When necessary, post hoc comparisons were performed with the Bonferroni correction. Kolmogorov-Smirnov tests indicated that all variables used for the analysis were normally distributed (all $z < 0.8$, $p > 0.3$).

References

- King, A. J. & Calvert, G. A. Multisensory integration: perceptual grouping by eye and ear. *Curr Biol.* **11**, 322–325 (2001).
- Abbas Shangari, T., Falahi, M., Bakouie, F. & Gharibzadeh, S. Multisensory integration using dynamical Bayesian networks. *Front Comput Neurosci* **9**, 58–60 (2015).
- Lehmann, S. & Murray, M. M. The role of multisensory memories in unisensory object discrimination. *Cogn Brain Res.* **24**, 326–334 (2005).
- Forster, B., Cavina-Pratesi, C., Aglioti, S. M. & Berlucchi, G. Redundant target effect and intersensory facilitation from visual-tactile interactions in simple reaction time. *Exp Brain Res.* **143**, 480–487 (2002).
- Schröger, E. & Widmann, A. Speeded responses to audiovisual signal changes result from bimodal integration. *Psychophysiology.* **35**, 755–759 (1998).
- Jordanescu, L., Guzman-Martinez, E., Grabowecky, M. & Suzuki, S. Characteristic sounds facilitate visual search. *Psychon Bull Rev.* **15**, 548–554 (2008).

7. Frassinetti, F., Bolognini, N. & Làdavas, E. Enhancement of visual perception by crossmodal visuo-auditory interaction. *Exp Brain Res.* **147**, 332–343 (2002).
8. Molholm, S., Ritter, W., Javitt, D. C. & Foxe, J. J. Multisensory visual–auditory object recognition in humans: a high-density electrical mapping study. *Cereb Cortex.* **14**, 452–465 (2004).
9. Gingras, G., Rowland, B. A. & Stein, B. E. The differing impact of multisensory and unisensory integration on behavior. *J Neurosci.* **29**, 4897–4902 (2009).
10. Alvarado, J. C., Vaughan, J. W., Stanford, T. R. & Stein, B. E. Multisensory versus unisensory integration: contrasting modes in the superior colliculus. *J Neurophysiol.* **97**, 3193–3205 (2007).
11. Wallace, M. T., Meredith, M. A. & Stein, B. E. Multisensory integration in the superior colliculus of the alert cat. *J Neurophysiol.* **80**, 1006–1010 (1998).
12. Bausenhardt, K. M. Mechanisms of multisensory integration in the time domain. *Procedia Soc Behav Sci.* **126**, 4–5 (2014).
13. Gondan, M., Niederhaus, B., Rösler, F. & Röder, B. Multisensory processing in the redundant-target effect: a behavioral and event-related potential study. *Percept Psychophys.* **67**, 713–726 (2005).
14. Goldring, J. E., Dorris, M. C., Corneil, B. D., Ballantyne, P. A. & Munoz, D. R. Combined eye-head gaze shifts to visual and auditory targets in humans. *Exp Brain Res.* **111**, 68–78 (1996).
15. Gielen, S. C. A. M., Schmidt, R. A. & Heuvel, P. J. M. On the nature of intersensory facilitation of reaction time. *Percept Psychophys.* **34**, 161–168 (1983).
16. Harrington, K. L. & Peck, K. C. Spatial disparity affects visual–auditory interactions in human sensorimotor processing. *Exp Brain Res.* **122**, 247–252 (1998).
17. Deroy, O. *et al.* The complex interplay between multisensory integration and perceptual awareness. *Multisens Res.* **29**, 585–606 (2016).
18. Laurienti, P., Kraft, R., Maldjian, J., Burdette, J. & Wallace, M. Semantic congruence is a critical factor in multisensory behavioral performance. *Exp Brain Res.* **158**, 405–414 (2004).
19. Laurienti, P. J. *et al.* Cross-modal sensory processing in the anterior cingulate and medial prefrontal cortices. *Hum Brain Mapp.* **19**, 213–223 (2003).
20. Fuster, J. M., Bodner, M. & Kroger, J. K. Cross-modal and cross-temporal association in neurons of frontal cortex. *Nature.* **405**, 347–351 (2000).
21. Amedi, A., von Kriegstein, K., van Atteveldt, N. M., Beauchamp, M. S. & Naumer, M. J. Functional imaging of human crossmodal identification and object recognition. *Exp Brain Res.* **166**, 559–571 (2005).
22. Calvert, G. A., Hansen, P. C., Iversen, S. D. & Brammer, M. J. Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the BOLD effect. *Neuroimage.* **14**, 427–438 (2001).
23. Ghazanfar, A. A. & Schroeder, C. E. Is neocortex essentially multisensory? *Trends Cogn Sci.* **10**, 278–285 (2006).
24. Molholm, S. *et al.* Audio-visual multisensory integration in superior parietal lobule revealed by human intracranial recordings. *J Neurophysiol.* **96**, 721–729 (2006).
25. Grefkes, C., Weiss, P. H., Zilles, K. & Fink, G. R. Crossmodal processing of object features in human anterior intraparietal cortex: an fMRI study implies equivalencies between humans and monkeys. *Neuron.* **35**, 173–184 (2002).
26. Moran, R. J., Molholm, S., Reilly, R. B. & Foxe, J. J. Changes in effective connectivity of human superior parietal lobule under multisensory and unisensory stimulation. *Eur J Neurosci.* **27**, 2303–2312 (2008).
27. Shams, L. & Seitz, A. R. Benefits of multisensory learning. *Trends Cogn Sci.* **12**, 411–417 (2008).
28. Kim, R. S., Seitz, A. R. & Shams, L. Benefits of stimulus congruency for multisensory facilitation of visual learning. *PLoS One.* **3**, e1532 (2008).
29. Seitz, A. R., Kim, R. & Shams, L. Sound facilitates visual learning. *Curr Biol.* **16**, 1422–1427 (2006).
30. Murray, M. M. *et al.* Rapid discrimination of visual and multisensory memories revealed by electrical neuroimaging. *Neuroimage.* **21**, 125–135 (2004).
31. Murray, M. M., Foxe, J. J. & Wylie, G. R. The brain uses single-trial multisensory memories to discriminate without awareness. *Neuroimage.* **27**, 473–478 (2005).
32. Thelen, A., Cappe, C. & Murray, M. M. Electrical neuroimaging of memory discrimination based on single-trial multisensory learning. *Neuroimage.* **62**, 1478–1488 (2012).
33. Moran, Z. D. *et al.* Multisensory encoding improves auditory recognition. *Multisens Res.* **26**, 581–592 (2013).
34. Matusz, P. J. *et al.* The role of auditory cortices in the retrieval of single-trial auditory-visual object memories. *Eur J Neurosci.* **41**, 699–708 (2015).
35. Thelen, A., Talsma, D. & Murray, M. M. Single-trial multisensory memories affect later auditory and visual object discrimination. *Cognition.* **138**, 148–165 (2015).
36. Thelen, A. & Murray, M. M. The efficacy of single-trial multisensory memories. *Multisens Res.* **26**, 483–502 (2013).
37. Heikkilä, J., Alho, K., Hyvönen, H. & Tiippana, K. Audiovisual semantic congruency during encoding enhances memory performance. *Exp Psychol.* **62**, 123–130 (2015).
38. Bigelow, J. & Poremba, A. Audiovisual integration facilitates monkeys’ short-term memory. *Anim Cogn.* **19**, 799–811 (2016).
39. Miniussi, C., Girelli, M. & Marzi, C. A. Neural site of the redundant target effect: electrophysiological evidence. *J Cogn Neurosci.* **10**, 216–230 (1998).
40. Brefczynski-Lewis, J., Lowitzsch, S., Parsons, M., Lemieux, S. & Puce, A. Audiovisual non-verbal dynamic faces elicit converging fMRI and ERP responses. *Brain Topogr.* **21**, 193–206 (2009).
41. Sperdin, H. F., Cappe, C., Foxe, J. J. & Murray, M. M. Early, Low-Level Auditory-Somatosensory Multisensory Interactions Impact Reaction Time Speed. *Front Integr Neurosci.* **3**, 1–10 (2009).
42. Pascual-Marqui, R. D. Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. *Methods Find Exp Clin Pharmacol.* **24**, 5–12 (2002).
43. Baart, M., Stekelenburg, J. J. & Vroomen, J. Electrophysiological evidence for speech-specific audiovisual integration. *Neuropsychologia.* **53**, 115–121 (2014).
44. Lewis, J. W. & Van Essen, D. C. Corticocortical connections of visual, sensorimotor, and multimodal processing areas in the parietal lobe of the macaque monkey. *J Comp Neurol.* **428**, 112–137 (2000).
45. Constantinidis, C. & Procyk, E. The primate working memory networks. *Cogn Affect Behav Neurosci.* **4**, 444–465 (2004).
46. Quak, M., London, R. E. & Talsma, D. A multisensory perspective of working memory. *Front Hum Neurosci.* **9**, 197–108 (2015).
47. Allen, R. J., Hitch, G. J., Mate, J. & Baddeley, A. D. Feature binding and attention in working memory: a resolution of previous contradictory findings. *Q J Exp Psychol.* **65**, 2369–2382 (2012).
48. Karlsen, P. J., Allen, R. J., Baddeley, A. D. & Hitch, G. J. Binding across space and time in visual working memory. *Mem Cognit.* **38**, 292–303 (2010).
49. Jiang, Y., Olson, I. R. & Chun, M. M. Organization of visual short-term memory. *J Exp Psychol Learn Mem Cogn.* **26**, 683–702 (2000).
50. Olsson, H. & Poom, L. Visual memory needs categories. *Proc Natl Acad Sci.* **102**, 8776–8780 (2005).
51. Allen, R. J., Hitch, G. J. & Baddeley, A. D. Cross-modal binding and working memory. *Vis Cogn.* **17**, 83–102 (2009).
52. Baddeley, A. D., Hitch, G. J. & Allen, R. J. Working memory and binding in sentence recall. *J Mem Lang.* **61**, 438–456 (2009).

53. Olson, I. R., Moore, K. S., Stark, M. & Chatterjee, A. Visual working memory is impaired when the medial temporal lobe is damaged. *J Cogn Neurosci*. **18**, 1087–1097 (2006).
54. Baddeley, A., Allen, R. & Vargha-Khadem, F. Is the hippocampus necessary for visual and verbal binding in working memory? *Neuropsychologia*. **48**, 1089–1095 (2010).
55. Baddeley, A. The episodic buffer: a new component of working memory? *Trends Cogn Sci*. **4**, 417–423 (2000).
56. Baddeley, A., Allen, R. & Hitch, G. Investigating the episodic buffer. *Psychol Belg*. **50**, 223–243 (2010).
57. Baddeley, A. The central executive: a concept and some misconceptions. *J Int Neuropsych Soc*. **4**, 523–526 (1998).
58. Baddeley, A. Exploring the central executive. *Q J Exp Psychol A*. **49**, 5–28 (1996).
59. Baddeley, A., Logie, R., Bressi, S., Sala, S. D. & Spinnler, H. Dementia and working memory. *Q J Exp Psychol A*. **38**, 603–618 (1986).
60. Baddeley, A. D., Bressi, S., Della Sala, S., Logie, R. & Spinnler, H. The decline of working memory in Alzheimer's disease. *Brain*. **114**, 2521–2542 (1991).
61. Owen, A. M., Sahakian, B. J., Semple, J., Polkey, C. E. & Robbins, T. W. Visuo-spatial short-term recognition memory and learning after temporal lobe excisions, frontal lobe excisions or amygdalo-hippocampectomy in man. *Neuropsychologia*. **33**, 1–24 (1995).
62. Baddeley, A., Della Sala, S., Robbins, T. W. & Baddeley, A. Working memory and executive control. *Philos Trans R Soc Lond B Biol Sci*. **351**, 1397–1404 (1996).
63. Rushworth, M. F., Nixon, P. D., Eacott, M. J. & Passingham, R. E. Ventral prefrontal cortex is not essential for working memory. *J Neurosci*. **17**, 4829–4838 (1997).
64. Collette, F. *et al.* Regional brain activity during tasks devoted to the central executive of working memory. *Cogn Brain Res*. **7**, 411–417 (1999).
65. Quintana, J. *et al.* Prefrontal-posterior parietal networks in schizophrenia: primary dysfunctions and secondary compensations. *Biol Psychiat*. **53**, 12–24 (2003).
66. Collette, F., Hogge, M. & Salmon, E. & Van der Linden, M., Exploration of the neural substrates of executive functioning by functional neuroimaging. *Neuroscience*. **139**, 209–221 (2006).
67. Collette, F. *et al.* Involvement of both prefrontal and inferior parietal cortex in dual-task performance. *Cogn Brain Res*. **24**, 237–251 (2005).
68. Wager, T. D. & Smith, E. E. Neuroimaging studies of working memory: a meta-analysis. *Cogn Affect Behav Neurosci*. **3**, 255–274 (2003).
69. Luckmann, H. C., Jacobs, H. I. L. & Sack, A. T. The cross-functional role of frontoparietal regions in cognition: internal attention as the overarching mechanism. *Prog Neurobiol*. **116**, 66–86 (2014).
70. Olson, I. R. & Berryhill, M. Some surprising findings on the involvement of the parietal lobe in human memory. *Neurobiol Learn Mem*. **91**, 155–165 (2009).
71. Duhamel, J., Colby, C. L. & Goldberg, M. E. Ventral intraparietal area of the macaque: congruent visual and somatic response properties. *J Neurophysiol*. **79**, 126–136 (1998).
72. Colby, C. L., Duhamel, J. R. & Goldberg, M. E. Ventral intraparietal area of the macaque: anatomic location and visual response properties. *J Neurophysiol*. **69**, 902–914 (1993).
73. Schlack, A., Hoffmann, K. P. & Bremmer, F. Interaction of linear vestibular and visual stimulation in the macaque ventral intraparietal area (VIP). *Eur J Neurosci*. **16**, 1877–1886 (2002).
74. Schlack, A., Sterbing-D'Angelo, S. J., Hartung, K., Hoffmann, K. & Bremmer, F. Multisensory space representations in the macaque ventral intraparietal area. *J Neurosci*. **25**, 4616–4625 (2005).
75. Andersen, R. A. Multimodal integration for the representation of space in the posterior parietal cortex. *Philos Trans R Soc Lond B Biol Sci*. **352**, 1421–1428 (1997).
76. Funahashi, S. & Andreau, J. M. Prefrontal cortex and neural mechanisms of executive function. *J Physiol-Paris*. **107**, 471–481 (2013).
77. Hampshire, A., Chamberlain, S. R., Monti, M. M., Duncan, J. & Owen, A. M. The role of the right inferior frontal gyrus: inhibition and attentional control. *Neuroimage*. **50**, 1313–1319 (2010).
78. Japee, S., Holiday, K., Satyshur, M. D., Mukai, I. & Ungerleider, L. G. A role of right middle frontal gyrus in reorienting of attention: a case study. *Front Syst Neurosci*. **9**, 1–16 (2015).
79. Eliasova, I., Anderkova, L., Marecek, R. & Rektorova, I. Non-invasive brain stimulation of the right inferior frontal gyrus may improve attention in early Alzheimer's disease: a pilot study. *J Neurosci*. **346**, 318–322 (2014).
80. Leung, H. C., Gore, J. C. & Goldman-Rakic, P. S. Sustained mnemonic response in the human middle frontal gyrus during on-line storage of spatial memoranda. *J Cogn Neurosci*. **14**, 659–671 (2002).
81. Molholm, S. *et al.* Multisensory auditory-visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Cogn Brain Res*. **14**, 115–128 (2002).
82. Giard, M. H. & Peronnet, F. Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *J Cogn Neurosci*. **11**, 473–490 (1999).
83. Schröger, E. & Widmann, A. Speeded responses to audiovisual signal changes result from bimodal integration. *Psychophysiology*. **35**, 755–759 (1998).
84. Murray, M. M. *et al.* The multisensory function of the human primary visual cortex. *Neuropsychologia*. **83**, 161–169 (2016).
85. Cappe, C., Thut, G., Romei, V. & Murray, M. M. Auditory-visual multisensory interactions in humans: timing, topography, directionality, and sources. *J Neurosci*. **30**, 12572–12580 (2010).
86. Raji, T. *et al.* Onset timing of cross-sensory activations and multisensory interactions in auditory and visual sensory cortices. *Eur J Neurosci*. **31**, 1772–1782 (2010).
87. Ten Oever, S. *et al.* The COGs (context, object, and goals) in multisensory processing. *Exp Brain Res*. **234**, 1307–1323 (2016).
88. Van der Burg, E., Talsma, D., Olivers, C. N. L., Hickey, C. & Theeuwes, J. Early multisensory interactions affect the competition among multiple visual objects. *Neuroimage*. **55**, 1208–1281 (2011).
89. Van der Burg, E., Olivers, C. N., Bronkhorst, A. W. & Theeuwes, J. Pip and pop: nonspatial auditory signals improve spatial visual search. *J Exp Psychol Hum Percept Perform*. **34**, 1053–1065 (2008).
90. Murray, M. M., Lewkowicz, D. J., Amedi, A. & Wallace, M. T. Multisensory processes: a balancing act across the lifespan. *Trends Neurosci*. **39**, 567–579 (2016).
91. Talsma, D. Predictive coding and multisensory integration: an attentional account of the multisensory mind. *Neurosci*. **9**, 1–13 (2015).
92. Raji, T., Uutela, K. & Hari, R. Audiovisual integration of letters in the human brain. *Neuron*. **28**, 617–625 (2000).
93. Sanfeliu, M. C. & Fernandez, A. A set of 254 Snodgrass-Vanderwart pictures standardized for Spanish: Norms for name agreement, image agreement, familiarity, and visual complexity. *Behav Res Methods Instrum Comput*. **28**, 537–555 (1996).
94. Snodgrass, J. G. & Vanderwart, M. A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *J Exp Psychol Learn Mem*. **6**, 174–215 (1980).
95. Delorme, A. & Makeig, S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Meth*. **134**, 9–21 (2004).
96. Groppe, D. M., Urbach, T. P. & Kutas, M. Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology*. **48**, 1711–1725 (2011).

97. Groppe, D. M., Urbach, T. P. & Kutas, M. Mass univariate analysis of event-related brain potentials/fields II: Simulation studies. *Psychophysiology*. **48**, 1726–1737 (2011).
98. Pascual-Marqui, R. D., Michel, C. M. & Lehmann, D. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. *Int J Psychophysiol*. **18**, 49–65 (1994).
99. Mulert, C. *et al.* Integration of fMRI and simultaneous EEG: towards a comprehensive understanding of localization and time-course of brain activity in target detection. *Neuroimage*. **22**, 83–94 (2004).
100. Holmes, A. P., Blair, R. C., Watson, J. & Ford, I. Nonparametric analysis of statistic images from functional mapping experiments. *J Cereb Blood Flow Metab*. **16**, 7–22 (1996).

Acknowledgements

This study was financially supported by National Natural Science Foundation of China granted to Min Li (No. 31170994), to Chen Bian (No. 31500896), and Major Program of "Twelfth Five-year Plan" for Medical Development of PLA to Min Li (No. AWS13J002).

Author Contributions

X.Y.J. and L.M. conceived the experiment, X.Y.J., X.Y.Y. and B.C. conducted the experiment, X.Y.J. analyzed the results and wrote the manuscript. All authors reviewed the manuscript.

Additional Information

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017