Proceedings of ICEGEE 2026 Symposuim: AI-Based Medicine and Biological Data Analysis
DOI: 10.54254/2753-8818/2026.AU33914

Spatial Attention under Visual-Auditory Crossmodal Paradigm

Yumeng Su

First Clinical Medical College, Southern Medical University, Guangzhou, China
3422237449@qq.com

Abstract. To investigate the differences in brain region activation and functional connectivity between spatial attention and neutral attention under a visual cue-auditory target crossmodal paradigm, and to examine whether there is a trade-off characteristic between the scope of brain network collaboration and connection strength in these two attention modes. Forty-nine healthy young adults participated as subjects. Task-state data were collected using 3.0T functional magnetic resonance imaging (fMRI). After using SPM12 to preprocess, core regions of interest (ROIs) were selected according to the AAL3 template and prior knowledge in the field of visual attention. Whole-brain activation was analyzed using the general linear model. Functional connectivity analysis was estimated using Pearson correlation and Fisher-z transformation. Paired t-tests were applied for statistical comparison between the two attention conditions. Relative to neutral attention, spatial attention activated a broader range of brain regions, evoked significantly broader activation across the occipital lobe, inferior parietal lobule, and motor-related areas of the frontal lobe. Neutral attention activation was highly concentrated in the core regions of the occipital visual cortex. Spatial attention supported a widely collaborative network spanning visual, parietal, and frontal areas was observed, whereas neutral attention showed concentrated collaborative activation within the occipital lobe with higher core node connection strength. Exploratory analysis revealed spatial attention strengthened functional connectivity between the precuneus and motor-related regions, while neutral attention mainly enhanced internal synchronization within the occipital lobe. The occipital visual cortex serves as the core hub for visual cue processing. Spatial attention relies on a large-scale distributed collaborative network pattern, whereas neutral attention adopts a core-region focused mode, reflecting a dynamic trade-off mechanism between brain network coverage and connection strength. These results provide experimental references at the functional connectivity level for understanding crossmodal regulation of visual spatial attention.

Keywords: Visual spatial attention, Neutral attention, Functional magnetic resonance imaging, Functional connectivity

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-8818/2026.AU33914

1. Introduction

Visual attention is a core mechanism that enables the human cognitive system to selectively processes and prioritizes external sensory information. The neural differences between spatially oriented attention and neutrally maintained attention have long been a central topic in cognitive neuroscience [1]. In natural settings, attentional processing typically operates in a crossmodal manner, in which individuals receive visual cues and prepare motor responses to auditory targets. Such processing involves visual recognition, auditory detection, and multisensory integration, all of which depend on dynamic activation and coordinated functional connectivity across multiple brain regions [2]. With high spatial resolution, functional magnetic resonance imaging (fMRI) has become a key technical tool for characterizing the topological structure of attentional processing brain networks, while functional connectivity analysis based on blood-oxygen-level-dependent signals has emerged as a core method for quantifying the synchronicity and collaborative patterns of neural activity between brain regions [3, 4].

According to the principle of retinotopic organization, the primary visual cortex and dorsal visual pathway form specific representations for central and peripheral visual fields, providing an anatomical foundation for distinct attentional control modes [5]. The dorsal attention network, centered on the inferior parietal lobule and precuneus, supports spatial orienting via top-down attentional control [1, 6]. However, in visual-cue auditory-target crossmodal tasks, the cooperative architecture among the dorsal attention network, visual processing streams, and motor preparation networks remains to be systematically verified using standardized paradigms.

Existing studies have shown that crossmodal attentional processing is accompanied by the collaborative activation of extensive brain regions, with brain region activation collaboration being the primary mode [2]. Spatial attention can significantly expand the coverage of brain networks [7]. However, most studies have not finely divided the functional subregions of the visual cortex, have paid less attention to the differential regulation of central and peripheral visual field representations, and have not systematically examined whether spatial attention and neutral attention follow an organizational rule of network coverage and connection strength trade-offs [8]. Additionally, most studies have focused on unimodal conditions, and there is still a lack of stable and consistent functional connectivity evidence regarding the brain network differences between the two attentional modes under crossmodal paradigms.

This study employed a standardized visual cue-auditory target crossmodal paradigm, setting spatial attention and neutral attention conditions, and used 3.0T fMRI to collect task-state data. Through standardized preprocessing and a priori ROI analysis, combined with general linear model and functional connectivity analysis, we investigated brain region activation differences and functional connectivity patterns under the two attentional conditions. This study aimed to clarify the significantly activated brain regions and functional subregion representation differences between spatial attention and neutral attention, characterize the number, strength, and topological features of core brain network connections under the two conditions, and provide experimental evidence for the neural mechanisms underlying crossmodal visual spatial attention.

2. Methods

2.1. Subjects and experimental paradigm

Fifty-two healthy young adults aged 18-27 were recruited for this study. Inclusion criteria included being right-handed, having normal or corrected-to-normal vision, normal hearing, no history of

DOI: 10.54254/2753-8818/2026.AU33914

brain injury, no history or diagnosis of mental illness, no use of psychotropic drugs, and no MRI contraindications (e.g., claustrophobia or metal implants). Three subjects were excluded due to excessive head motion or incomplete data collection, leaving 49 subjects (mean age 22.7 years, 20 males, 29 females) with complete task-state fMRI data for analysis. Data were collected from the Department of Radiology, Beijing Youan Hospital, Capital Medical University. This study was approved by the Ethics Committee of Capital Medical University, and all subjects were informed of the experimental details and signed written informed consent before participation, receiving transportation compensation.

The experiment used a visual cue-auditory target crossmodal task paradigm, with auditory targets being unilateral single-channel stimuli to which subjects responded by pressing a key indicating the side of target presentation. The full task consisted of two independent runs, each containing four blocks (S-block for spatial attention condition, N-block for neutral attention condition, N-block, S-block) to balance block order effects. Random static black screen intervals of 14s, 16s, 18s, or 20s were set between adjacent blocks. Each block consisted of 10 trials, with all trials in a block having the same type of cue stimulus. Each trial lasted 4s, with the presentation sequence being: fixation point background (1500 ms) → cue stimulus (100 ms, spatial cue/neutral cue) → background → target stimulus (50 ms, unilateral auditory stimulus) → background, with an interval of 300 ms or 1500 ms between cue and target, and the final background duration adjusted to complete 4s based on the interval time. The experimental paradigm and timing structure are shown in Figure 1.

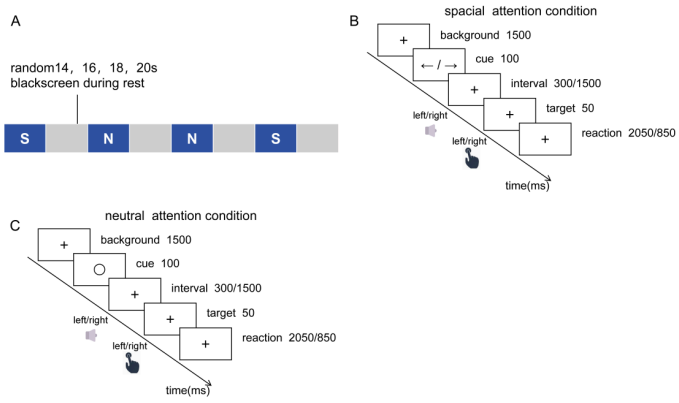


Figure 1. Illustrates the experimental paradigm

Note: A. Each task run consisted of four blocks, with the block order being S, N, N, S, where S represents the spatial attention condition and N represents the neutral attention condition. Random baseline gaps of 14, 16, 18, or 20s were present between every two blocks. B. The presentation sequence over time for each trial under spatial cue conditions in the visual-auditory task. C. The presentation sequence over time for each trial under neutral cue conditions in the visual-auditory task. A static black screen was set after each trial.

To avoid subjects forming positional predictions and ensure the experimental validity of spatial attention orientation, 20% invalid cues (where the cue stimulus and target stimulus directions were inconsistent) were set, randomly distributed in S-blocks, and the two interval times were balanced throughout the experiment.

Before participating in the experiment, subjects were informed of the relevant details and familiarized with the task paradigm. During the experiment, subjects lay in an MRI scanner, wearing MRI-compatible audiovisual equipment. Auditory signals were transmitted via headphones connected to a computer outside the MRI room, and visual signals were displayed on a video screen connected via a video cable. Subjects held a self-made key with their right hand, pressing the left

DOI: 10.54254/2753-8818/2026.AU33914

key with their index finger and the right key with their ring finger to respond. A trigger was used to achieve synchronous triggering. Before each task, subjects were prompted about the task and asked about their comfort via the MRI-compatible audiovisual equipment.

2.2. fMRI data acquisition

All fMRI data were acquired using a 3.0T MRI scanner (GE SIGNA Pioneer) with a 32-channel head coil at the Department of Radiology, Beijing Youan Hospital, Capital Medical University. Subjects were asked to complete Task A while remaining still and breathing evenly. The scanning time for subjects to complete the task was 8.5 minutes. Specific scanning parameters were as follows: fMRI scanning parameters were repetition time (TR) = 2000ms, echo time (TE) = 30ms, flip angle (FA) = 90°, slice thickness = 3.5mm, no slice gap, field of view (FOV) = 24cm×24cm, matrix size = 64×64, 36 slices. Simultaneously, whole-brain axial T1-weighted three-dimensional structural images were acquired with TR = 7.032ms, TE = 3.016ms, FA = 12°, number of excitations (NEX) = 1.0, FOV = 24cm×24cm, matrix size = 256×256, slice thickness = 1.0mm, no slice gap, 192 slices. For subjects with larger head circumferences, the head fixation device was appropriately adjusted to ensure comfort and head stability.

2.3. fMRI data preprocessing

The collected task-state fMRI data were preprocessed using SPM12 software on the MATLAB R2018b platform, with the specific workflow covering the removal of initial time points, slice timing correction, head motion correction, spatial normalization, and smoothing [9].

First, the first six time points of image data were removed to eliminate artifacts caused by magnetic field instability on early-acquired images, leaving 253 effective time points for processing. Given that the GE MRI scanner uses an interleaved scanning mode, there are differences in the acquisition times of brain images at different layers. Therefore, time layer correction was performed by interpolation, selecting the middle time layer as the reference layer to compensate for time deviations between different layers.

During scanning, even with subjects' heads fixed, fatigue-induced head motion may occur over long acquisition times, and micro-head motions can significantly affect fMRI data. Therefore, head motion parameters were estimated for all subjects, and data from subjects with head translations exceeding 3 mm or rotations exceeding 3° were excluded to ensure the reliability of subsequent analysis [4]. The 49 subjects included in this study all met the quality control criteria, with an average frame displacement (mean FD) < 0.2mm.

Due to individual differences in brain size, shape, and spatial position among subjects, it was necessary to register them to a unified standard space for subsequent analysis. This study used a two-step registration method: first, the T1 structural image was registered to the functional image, and gray matter, white matter, and cerebrospinal fluid segmentation of the registered structural image was performed to obtain a transformation matrix from the original space to the Montreal Neurological Institute (MNI) standard space. Then, based on this matrix, the functional image was transformed into MNI standard space, and after normalization, the image voxel size was resampled to 3×3×3 mm^3 [5].

Finally, smoothing was performed using a 6×6×6 mm^3 full-width at half-maximum smoothing kernel to reduce the impact of spatial normalization and improve image signal-to-noise ratio [7].

DOI: 10.54254/2753-8818/2026.AU33914

2.4. Activation analysis

This study used a general linear model (GLM) in SPM12 software for brain region activation analysis, with the overall analysis workflow divided into two consecutive levels: individual-level first-level analysis and group-level second-level analysis, with the analysis logic consistent with the block design of the experimental paradigm [9].

In the individual-level first-level analysis, regression variables (block-level regressors) were constructed for the spatial attention block (S-block) and neutral attention block (N-block), respectively. The task time series was convolved with the standard hemodynamic response function (HRF) to form the core predictive factor of the model, and head motion parameters were included as interference covariates in the design matrix to control for the impact of head motion on the blood-oxygen-level-dependent (BOLD) signal. The model regression coefficients (β values) were estimated using the ordinary least squares (OLS) method to reflect the response intensity of brain regions under different conditions, and two opposite-direction statistical contrasts were set: S task > N task and N task > S task, generating individual-level contrast maps for subsequent group statistical inference.

In the group-level second-level analysis, the first-level contrast images of all subjects were included in a second-sample t-test to test the statistical significance of brain region activation differences between the two attentional conditions at the group level. Cluster-level familywise error (FWE) correction was used for multiple comparison correction, with the correction threshold set at $p < 0.05$ (uncorrected $p < 0.001$ at the voxel level), ultimately determining statistically significant activated brain region clusters under the two conditions [4, 9].

2.5. Functional connectivity analysis

2.5.1. Selection of regions of interest (ROIs) and brain region segmentation

The selection of regions of interest (ROIs) is directly based on the whole-brain activation results. Core brain regions with significant activation under the conditions of spatial attention and neutral attention, and with cluster sizes greater than 100 voxels, are extracted as the target ROIs. Ultimately, 29 core ROIs are identified, covering the occipital visual cortex, the dorsal attention network in the parietal lobe, the motor-related cortex in the frontal lobe, subcortical nuclei, and cerebellar-related regions. The anatomical and functional localizations of all ROIs are referenced from classic research findings in the field of visual attention.

This study employs the Automated Anatomical Labeling 3 (AAL3) template, based on MRI images, for brain region segmentation. The AAL3 template provides relatively fine spatial resolution and stable anatomical references [5], making it suitable for subsequent functional connectivity analysis.

2.5.2. Functional connectivity

This study utilizes the residual functional connectivity (residual FC) method to analyze the functional coupling characteristics between brain regions, aiming to eliminate task-induced signals and noise interference and more purely reflect the intrinsic neural activity synchrony between brain regions. The analysis process is implemented in MATLAB R2018b and Python environments, with the following specific steps:

DOI: 10.54254/2753-8818/2026.AU33914

Extract the blood-oxygen-level-dependent (BOLD) signal time series from the 29 ROIs in the preprocessed functional images. Perform task-effect removal on the time series by regressing out noise components such as head motion parameters, cerebrospinal fluid signals, and white matter signals. Use a finite impulse response (FIR) model to correct for hemodynamic delays, obtaining optimized residual time series.

Group and organize the time series of the 29 core ROIs according to the two experimental conditions of spatial attention and neutral attention. Each group of data includes 49 subjects, 2 runs, and 253 time points, forming a standard data matrix with dimensions of 29×253.

For individual-level functional connectivity construction, analyze the S task and N task separately. Conduct Pearson correlation analysis on the residual time series of ROIs for each subject and each task condition to construct a 29×29 functional connectivity matrix. To meet the normal distribution assumption for subsequent statistical tests, perform Fisher-z transformation on the correlation coefficients. Take the arithmetic mean of the Fisher-z matrices for the two runs to obtain the final functional connectivity matrix for each subject and each condition (a single-subject single-condition matrix is 29×29, corresponding to the pairwise connections of the 29 ROIs) [4].

Group-level statistical tests are divided into two steps. The first step uses a one-sample t-test, with 406 unique upper-triangular connections as independent statistical units (non-repetitive pairwise connections of the 29 ROIs). Conduct statistical tests on the functional connectivity results of the two conditions against 0 at a 95% confidence level. Use family-wise error (FWE) correction ($p < 0.05$) to screen for statistically significant positive and negative functional connections [4, 10]. The second step employs a paired t-test to compare the Fisher-z values between the two conditions at a 95% confidence level, with 406 unique connections as statistical units. This step tests whether the mean of the difference values between the spatial attention and neutral attention conditions is significantly different from 0, thereby clarifying the differential characteristics of functional connectivity in brain networks under the two attention modes [4, 9].

Through the above analysis, we explore the dynamic changes in functional connectivity between brain regions under different comparison conditions and reveal the regulatory role of tasks in the functional integration of large-scale brain networks.

3. Research results

The average task accuracy rates of all subjects are above 95%, indicating stable and reliable task performance.

3.1. Significantly activated brain regions under spatial attention and neutral attention

Compared to neutral attention, spatial attention can significantly activate a wider range of brain regions. After family-wise error correction ($p < 0.05$), the activated regions are mainly distributed in the bilateral occipital visual cortex (MOG.L/R, IOG.L/R, SOG.L/R, Cal.L/R, Ling.L/R, Cun.R), the left inferior parietal lobule (IPL.L), the left motor-related cortex in the frontal lobe (SMA.L, IFGtri.L, IFGope.L), as well as regions such as the left insula (Ins.L), the left putamen (Put.L), and the right cerebellar region 6 (Cereb6.R). The peak activation coordinates and brain region information are presented in Table 1. The activation pattern under neutral attention is more concentrated compared to spatial attention, with highly focused activation in the core regions of the occipital visual cortex, including the bilateral cuneus (Cun.L/R), the bilateral calcarine fissure cortex (Cal.L/R), and the left superior occipital gyrus (SOG.L). There is no significant activation in brain regions such as the parietal and frontal lobes. The above results are obtained through whole-brain

DOI: 10.54254/2753-8818/2026.AU33914

one-sample t-tests. The peak activation coordinates and brain region localizations are shown in Table 1 and Table 2, and the spatial distribution of activation differences between conditions is illustrated in Figure 2.

Table 1. Significantly activated brain regions with cluster sizes > 100 voxels in the S>N contrast and their peak MNI coordinates

	Number	Cluster Size	Anatomy Region	t Value	x	y	z
S>N	01	1269	MOG.L	6.24	-22	-90	12
	02	685	IOG.L	6.54	-24	-98	-8
	03	493	MOG.R	7.15	36	-94	4
	04	450	IPL.L	4.27	-28	-48	46
	05	372	Cal.R	5.05	14	-96	-2
	06	356	IOG.R	6.01	36	-92	-2
	07	314	Cal.L	4.74	-10	-90	0
	08	282	MTG.L	5.01	-48	-68	0
	09	250	Ling.L	5.43	-20	-84	-16
	10	240	Ling.R	5.47	28	-86	-6
	11	238	SMA.L	4.75	-8	-12	56
	12	227	IFGtri.L	5.33	-50	16	4
	13	225	Fus.L	4.90	-30	-78	-12
	14	221	IFGope.L	5.53	-50	14	4
	15	191	SOG.L	5.01	-16	-94	12
	16	182	SOG.R	5.70	24	-88	10
	17	173	Fus.R	5.42	32	-82	-4
	18	152	Ins.L	4.29	-40	6	6
	19	153	Ang.L	4.25	-44	-56	40
	20	138	Cun.R	4.68	20	-98	10
	21	106	Smg.L	3.78	-64	-46	24
	22	102	Cereb6.R	4.72	24	-76	-20
	23	100	Put.L	4.54	-24	-12	10

Table 2. Significantly activated brain regions with cluster sizes > 100 voxels in the N>S contrast and their peak MNI coordinates.

	Number	Cluster Size	Anatomy Region	t Value	x	y	z
N>S	24	613	Cun.L	-5.47	-14	-68	20
	25	351	Cun.R	-4.88	18	-80	32
	26	271	Cal.R	-5.05	18	-60	16
	27	244	Cal.L	-5.20	-14	-66	18
	28	155	Prec.R	-4.53	28	-50	2
	29	137	SOG.L	-4.42	-14	-82	40

DOI: 10.54254/2753-8818/2026.AU33914

Note: Peak coordinates are located using xjView. MOG = Middle Occipital Gyrus, IOG = Inferior Occipital Gyrus, IPL = Inferior Parietal Lobule, Cal = Calcarine Cortex, MTG = Middle Temporal Gyrus, Ling = Lingual Gyrus, SMA = Supplementary Motor Area, IFGtri = Inferior Frontal Gyrus, Triangular Part, Fus = Fusiform Gyrus, IFGope = Inferior Frontal Gyrus, Opercular Part, SOG = Superior Occipital Gyrus, Ins = Insula, Ang = Angular Gyrus, Cun = Cuneus, Smg = Supramarginal Gyrus, Cereb6 = Cerebellar Lobule VI, Put = Putamen, Prec = Precuneus, L = Left, R = Right.

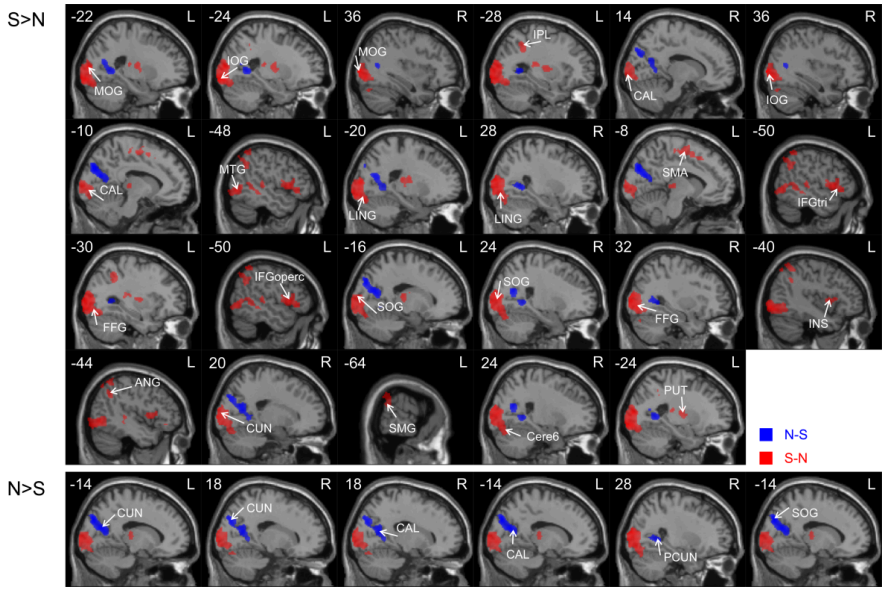


Figure 2. Brain activation difference map between spatial attention and neutral attention contrasts

Note: The colored regions represent the spatial distribution of brain activation differences between conditions: red indicates S>N (enhanced activation under spatial attention), and blue indicates N>S (enhanced activation under neutral attention). The abbreviations in the figure have the same meanings as those in Table 1. The numerical values in the top-left corner represent the corresponding x-axis coordinates. The results reflect the core differences between the two attention modes at the levels of the visual cortex and attention network.

3.2. Functional connectivity matrix

Based on the method introduced in Section 2.5.2, Pearson correlations were applied to the 29 ROIs to construct a 29×29 functional connectivity matrix. Figure 3 shows the functional connectivity network matrix of the first participant. This matrix is only used to visually display the topological structure of the functional connectivity matrix and the typical connection pattern at the individual level, and is not involved in group-level statistical inference. The matrix color follows the visualization standard set in this study. Red represents positive functional connectivity, indicating enhanced synchronization of neural activity between brain regions. Blue represents negative functional connectivity, indicating weakened synchronization of neural activity between brain regions. The basic characteristics of functional interaction among individual brain regions are clearly presented.

DOI: 10.54254/2753-8818/2026.AU33914

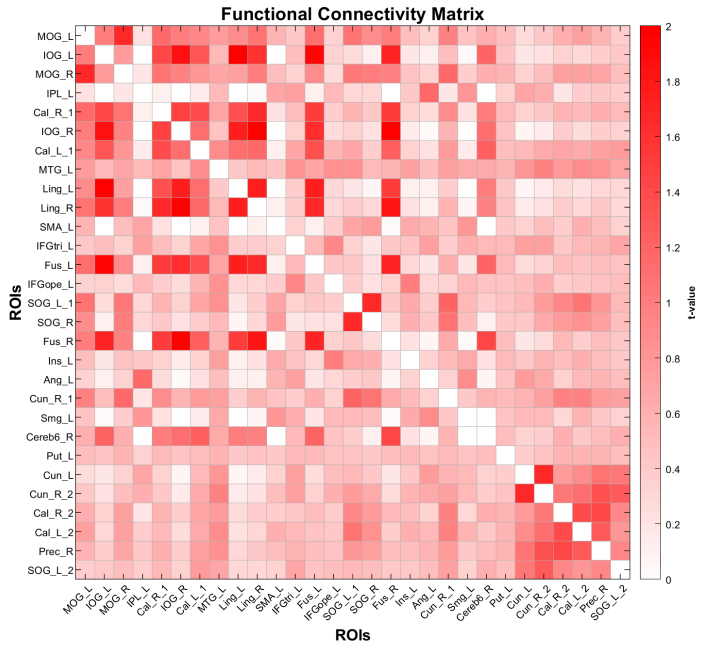


Figure 3. Functional connectivity network matrix. The figure is a 29×29 matrix diagram representing the functional connectivity among the 29 ROIs, i.e., the correlations between the 29 ROIs

3.3. Functional connectivity under spatial attention and neutral attention conditions

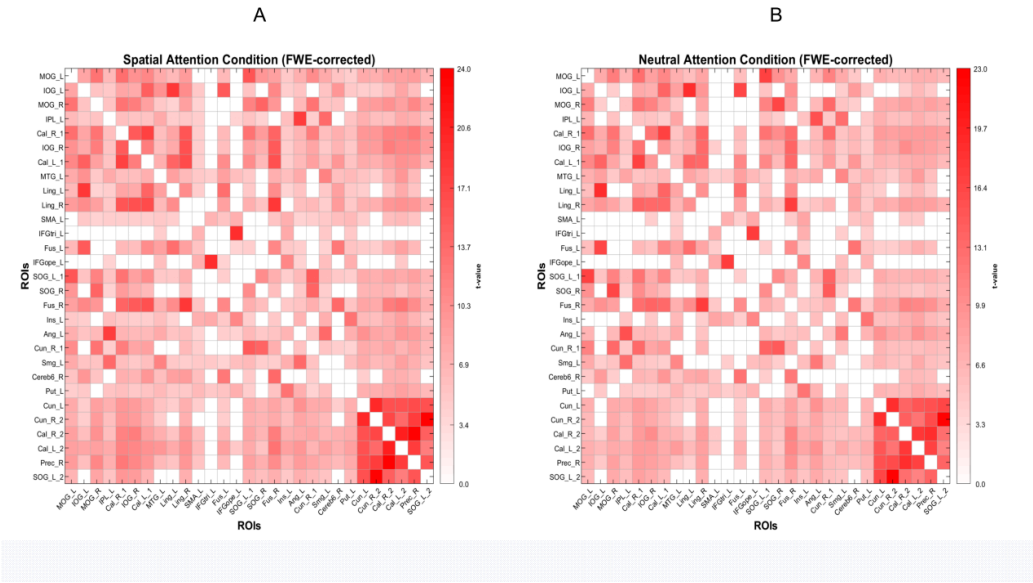


Figure 4. Shows the t-value maps from one-sample tests of the functional networks under the two conditions, with panels A and B presenting the results after Family-Wise Error (FWE) correction. Panel A shows the results where the functional connectivity among the 29 ROIs under the spatial condition significantly differs from zero; Panel B shows the results where the data under the neutral condition significantly differs from zero; zero represents the complete absence of functional connectivity

DOI: 10.54254/2753-8818/2026.AU33914

Significant functional connectivity under spatial attention and neutral attention conditions exhibited significant differences after FWE correction ($p < 0.05$), as shown in Figure 4. Under the spatial attention condition, a broader range of significant functional connectivity was observed, with 330 significant co-activations of brain regions detected. The connections covered the occipital visual cortex, parietal attention network, and frontal motor-related network, forming a visual–parietal–frontal collaborative network, consistent with the pattern of dorsal attention network regulating cross-modal processing [1, 11]. Under the neutral attention condition, 293 significant co-activations of brain regions were detected, with connections highly concentrated within the occipital visual cortex and a few involving cerebellar-related regions. No significant negative functional connectivity at the group level was observed under either condition [4, 9].

3.4. Functional network differences between spatial attention and neutral attention conditions

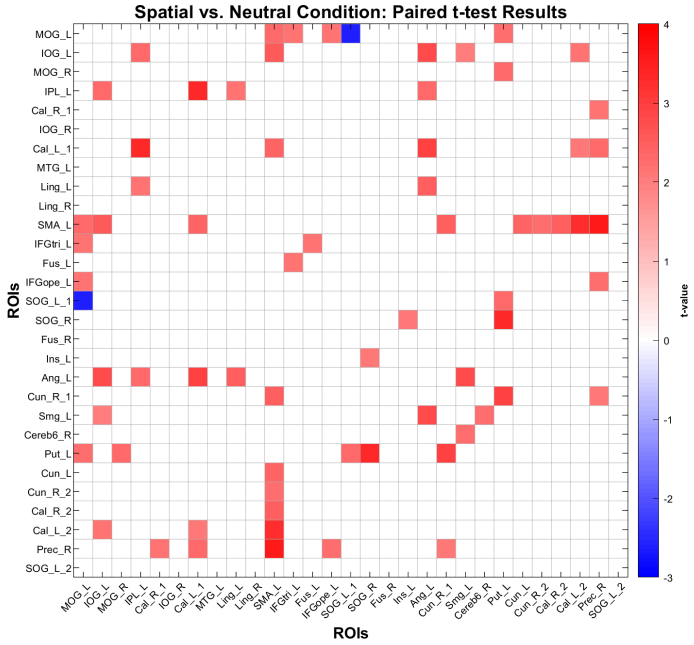


Figure 5. Shows the functional network differences between the spatial and neutral conditions, presenting uncorrected results

Paired t-tests were performed between the two conditions. Results showed no significant level after FWE correction. Local connectivity differences were only observed in uncorrected exploratory analysis ($p<0.05$) [10]. Results showed 35 unique connections with differences between conditions (corresponding to 70 non-diagonal matrix elements). Among them, the differential connection with significantly stronger spatial attention than neutral attention was most prominent in the coupling between the precuneus and supplementary motor area ($t=3.571$), followed by the connection between the right superior occipital gyrus and left putamen, and the connection between the left calcarine cortex and left inferior parietal lobule. These connections indicated enhanced connectivity between the visual cortex and subcortical nuclei, as well as between the visual cortex and parietal attention centers [1, 11]. The differential connection with significantly stronger neutral attention than spatial attention was only found in the bidirectional connection between the left superior occipital

DOI: 10.54254/2753-8818/2026.AU33914

gyrus and left middle occipital gyrus (t=-2.597), showing improved synchronization within the visual cortex [5].

4. Discussion

4.1. Differences in activation patterns between spatial attention and neutral attention align with the regulatory principles of the dorsal attention network

This study shows that spatial attention activates more brain regions than neutral attention. This result is consistent with the classic theory of the dorsal attention network [1]. Spatial attention needs top-down regulation from the inferior parietal lobule and precuneus to improve the processing efficiency of the peripheral visual field [1, 12]. Spatial attention activates the occipital cortex, parietal attention network and frontal motor-related cortex. The inferior parietal lobule codes spatial information. The frontal motor-related cortex prepares motor responses [1, 11]. Their co-activation supports the cross-modal pathway of visual cue recognition, spatial attention orientation and auditory target response.

Neutral attention does not involve spatial orientation. Participants only need basic visual processing and central fixation [13]. Activated regions are limited to the occipital visual cortex. The activation difference between the two attention patterns depends on whether the dorsal attention network and motor preparation network are recruited. This result reflects task-dependent recruitment of brain regions in visual attention processing [7].

In this study, some occipital brain regions (such as Cal.L/R, SOG.L, Cun.R) showed significant activation in both comparisons. This phenomenon may be related to the retinotopic mapping principles of the visual cortex [5, 14]: The spatial task may activate subregions for peripheral vision [15], while the neutral task may activate subregions for central vision [13]. However, since this study did not conduct independent retinotopic localization experiments, the functional attribution of these subregions remains speculative based on classic theories and requires validation through dedicated visual field mapping experiments in subsequent studies [5].

4.2. Functional connectivity characteristics of the two attention conditions reflect different organizational patterns of brain networks

This study found that both attention conditions were primarily characterized by co-activation of brain regions, with no significant negative functional connectivity detected at the group level, suggesting that the collaborative mode of brain networks in cross-modal visual cue processing is mainly characterized by enhanced synchronization of neural activity [2, 7], consistent with the conclusions of existing cross-modal attention studies [2]. However, there were significant differences in the distribution and intensity of functional connectivity between the two attention conditions, reflecting two different organizational patterns of brain networks:

Large-scale distributed collaborative mode under spatial attention: Under spatial attention, a widespread network of co-activated brain regions covering the visual-parietal-frontal areas was formed, with a significantly higher number of connections than under neutral attention. This mode is an inevitable requirement for the cross-modal processing of spatial attention—the spatial information of visual cues needs to be transmitted through the dorsal attention network to the motor preparation system [16, 17] to achieve the collaboration of "spatial orientation-response preparation." Therefore, functional coupling needs to be formed among brain regions in the visual, attention, and motor systems, expanding the network's collaborative scope to achieve multi-link information integration [7, 11].

DOI: 10.54254/2753-8818/2026.AU33914

Core region-focused mode under neutral attention: Functional connectivity under neutral attention was concentrated within the occipital visual cortex, with fewer connections but a slightly higher average connection strength in core visual brain regions than under spatial attention. Exploratory analysis found that functional connectivity within the occipital lobe was significantly stronger than under spatial attention [18, 19]. This mode suggests that the brain allocates cognitive resources to the basic visual processing network when spatial orientation is not needed [7]. This helps improve the efficiency of visual input processing and realizes effective recognition of visual cues [5].

The functional connectivity differences between the two attention modes essentially reflect the core characteristic of "on-demand adjustment and efficient allocation" of brain networks in cognitive processing [20] and also support the competitive interaction theory of visual attention [21, 22].

4.3. Collaborative characteristics of visual attention and motor preparation in cross-modal processing

Exploratory analysis found that spatial attention significantly enhanced the functional connectivity between the precuneus and the supplementary motor area, revealing a close collaboration between spatial attention and motor preparation in the cross-modal paradigm [16, 23]. The precuneus, as an important node of the dorsal attention network, is responsible for maintaining and orienting spatial attention [1, 12], while the supplementary motor area participates in the planning and initiation of motor responses [17]. The functional coupling between these two regions implies that the orientation information of spatial attention can directly regulate the motor preparation system, enabling the brain to prepare in advance for the lateral response to auditory targets after recognizing visual spatial cues.

This collaborative characteristic reflects the predictive coding mechanism of cross-modal attention processing: visual spatial cues provide predictive information about the location of auditory targets. The brain pre-activates motor pathways by enhancing connectivity between attention and motor networks [1, 11]. This mechanism speeds up responses to auditory targets. Neutral attention does not involve spatial prediction [13], so no attention-motor coupling is observed. Only internal cooperation within the visual cortex is maintained [5, 18]. The regulatory role of cross-modal spatial attention on motor preparation has been widely validated in visual-auditory paradigms [24], and functional coupling between the parietal attention network and the motor cortex is a key neural pathway for achieving this regulation [25].

5. Conclusion

This study used a visual-cue auditory-target cross-modal fMRI paradigm combined with brain functional network analysis to systematically investigate differences in brain region activation and functional connectivity patterns between spatial attention and neutral attention conditions. Results show that the occipital visual cortex serves as the core network for cross-modal visual cue processing and is stably activated under both attention conditions. Spatial attention additionally recruits the dorsal attention network and frontal motor-related cortex, forming a large-scale distributed cooperative pattern. Neutral attention activates only core occipital regions, presenting a core-region focused pattern. These two patterns reflect a dynamic trade-off between the cooperative scope and connection strength of brain networks. Exploratory analysis indicates that spatial attention enhances functional coupling between the precuneus and supplementary motor area, supporting early regulation of motor response pathways by spatial cues. Neutral attention only maintains

DOI: 10.54254/2753-8818/2026.AU33914

internal synchronization within the visual cortex. This study provides experimental evidence at the functional connectivity level for the regulatory mechanism of cross-modal visual attention. Limited by sample size and statistical power, the relevant network characteristics still require further verification under large-sample and highly rigorous statistical conditions.

References

[1] Corbetta M, Shulman G L. Control of goal-directed and stimulus-driven attention in the brain [J]. Nature Reviews Neuroscience, 2002, 3(3): 201-215.
[2] Talsma D, Senkowski D, Soto-Faraco S, et al. The multifaceted interplay between attention and multisensory integration [J]. Trends in Cognitive Sciences, 2010, 14(9): 400-410.
[3] Gong Qiyong, Wu Yin, Lu Guangming. Application of functional magnetic resonance imaging in cognitive neuroscience [J]. Chinese Journal of Radiology, 2015, 49(3): 233-237.
[4] Zang Yufeng, He Huiguang, Tian Lixia. Methodology and clinical applications of resting-state fMRI functional connectivity [J]. Science China: Life Sciences, 2012, 42(1): 1-16.
[5] Wandell B A, Dumoulin S O, Brewer A A. Visual field maps in human cortex [J]. Neuron, 2007, 56(3): 366-383.
[6] Chen Antao, Li Hong. Cognitive neural mechanisms of attention [J]. Journal of Southwest University (Social Sciences Edition), 2007, 33(4): 28-34.
[7] Gregoriou G G, Gotts S J, Zhou H, et al. Dynamic functional connectivity in the human brain during attention [J]. Cerebral Cortex, 2009, 19(1): 153-161.
[8] Fox M D, Raichle M E. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging [J]. Nature Reviews Neuroscience, 2007, 8(9): 700-711.
[9] Friston K J, Ashburner J, Kiebel S J, et al. Statistical parametric mapping: the analysis of functional images [M]. London: Academic Press, 2007.
[10] Bennett C M, Baird A A, Miller M B, et al. Nonparametric permutation tests for neuroimaging [J]. Human Brain Mapping, 2009, 30(1): 1-25.
[11] Shulman G L, Corbetta M, Buckner R L, et al. Common blood flow changes across visual tasks: II [J]. Journal of Cognitive Neuroscience, 1997, 9(5): 648-663.
[12] Bisley J W, Goldberg M E. Attention, intention, and priority in the parietal lobe [J]. Annual Review of Neuroscience, 2010, 33: 1-21.
[13] Kanwisher N, Wojciulik E. Visual attention: insights from brain imaging [J]. Nature Reviews Neuroscience, 2000, 1(1): 91-100.
[14] Silver M A, Kastner S. Topographic maps in human visual cortex [J]. Annual Review of Neuroscience, 2009, 32: 265-293.
[15] Kastner S, Pinsk M A, De Weerd P, et al. Increased activity in human visual cortex during directed attention in the absence of visual stimulation [J]. Neuron, 1999, 22(3): 751-761.
[16] Rizzolatti G, Riggio L, Dascola I, et al. Reorienting attention across the horizontal and vertical meridians [J]. Neuropsychologia, 1987, 25(1): 31-40.
[17] Moore T, Armstrong K M. Selective gating of visual processing by microstimulation in frontal cortex [J]. Nature, 2003, 421(6920): 370-373.
[18] Beck D M, Kastner S. Top-down and bottom-up mechanisms in biasing competition in the human brain [J]. Vision Research, 2009, 49(10): 1154-1165.
[19] Maunsell J H. Neuronal representations of visual attention [J]. Annual Review of Neuroscience, 1995, 18: 313-337.
[20] Honey C J, Sporns O, Cammoun L, et al. Predicting human resting-state functional connectivity from structural connectivity [J]. Proceedings of the National Academy of Sciences, 2009, 106(6): 2035-2040.
[21] Reynolds J H, Chelazzi L, Desimone R. Competitive mechanisms subserve attention in macaque areas V2 and V4 [J]. Journal of Neuroscience, 1999, 19(14): 6376-6383.
[22] Seydell-Greenwald A G, Astafiev S V, Shulman G L, et al. Human parietal cortex subserves spatial attention and eye movements [J]. Cerebral Cortex, 2012, 22(1): 11-23.
[23] Liu X, Steier J, Welch R B. Multisensory integration: current issues from the perspective of single neuron [J]. Neuron, 2011, 72(2): 261-282.
[24] Macaluso E, Driver J, Frith CD. Multisensory integration of spatial information across vision and touch. Neuron, 2000, 27(3): 521-534.
[25] Hagler DJ Jr, Sereno MI. Spatial attention modulates parietal–premotor effective connectivity in humans. Cerebral Cortex, 2006, 16(12): 1722-1733.