Behavioural Brain Research 411 (2021) 113382

Contents lists available at ScienceDirect

Behavioural Brain Research

journal homepage: www.elsevier.com/locate/bbr

Audiovisual working memory and association with resting-state regional homogeneity

Yuanjun Xie [a,b,*], Yanyan Li [a], Muzhen Guan [c], Haidan Duan [a], Xiliang Xu [a], Peng Fang [d]

[a] School of Education, Xin Yang College, Xin Yang, China
[b] Department of Radiology, Xijing Hospital, Fourth Military Medical University, Xi'an, China
[c] Department of Mental Health, Xi'an Medical University, Xi'an, China
[d] Department of Military Medical Psychology, Fourth Military Medical University, Xi'an, China

ARTICLE INFO

Keywords:
Audiovisual object
Semantic congruence
Working memory
Resting-state fMRI
Regional homogeneity
Executive control network

ABSTRACT

Multisensory processing is a prevalent research issue. However, multisensory working memory research has received inadequate attention. The present study aimed to investigate the behavioral performance of an audiovisual working memory task and its association with resting-state functional magnetic resonance imaging (fMRI) regional homogeneity (ReHo). A total of 128 healthy participants were enrolled in this study. The participants completed a modified Sternberg working memory task using complex auditory and visual objects as materials involved in different encoding conditions, including semantically congruent audiovisual, semantically incongruent audiovisual, and single modality of auditory or visual object encoding. Two subgroups received resting-state fMRI scans according to their behavioral performances. The results showed that the semantically congruent audiovisual object encoding sped up the later unisensory memory recognition in this task. Moreover, the high behavioral performance (response time, RT) group showed increased ReHo in the executive control network (ECN) and decreased ReHo in the default mode network (DMN) and saline network (SN). In addition, resting-state ReHo values in the ECN nodes (e.g., middle frontal gyrus and superior frontal gyrus) was correlated with RT. These findings suggested that semantically congruent audiovisual processing in working memory was superior to unisensory memory recognition and may be involved in the different functional networks such as ECN.

1. Introduction

We live in a multisensory world that perceives through vision, audition, and tactility. Although we simultaneously receive information from these sensory modalities, psychological research mainly focuses on studying our senses in separation. Recently, a multisensory approach to working memory has attracted the interest of researchers, which contributes to a realistic understanding of how working memory processes encoding information from different sensory inputs. One of the first studies to utilize multisensory stimuli in working memory was Paivio and Thompson [1]. In their study, participants were instructed to remember three different types of natural objects: visual (i.e., picture), auditory (i.e., sound), and audiovisual (i.e., picture-sound pairing) for a later free-recall test. The authors have found an improvement of recall for audiovisual stimuli compared to single auditory and visual stimuli. Similarly, Goolkasian and Foos have found that recall performance for the audiovisual format of picture-spoken words was superior to the double visual presentation of pictures or written words [2]. The improvement of recall memory performance in audiovisual encoding may result from integrating information from single sensory modalities rather than the redundancy of dual-format presentations [3].

Furthermore, some researchers have investigated the effects of audiovisual encoding on subsequent unisensory memory recognition in working memory tasks [4,5]. During the encoding stage, all three types of audiovisual stimuli were presented in random order: congruent audiovisual stimuli (i.e., A picture of sheep and a sound of sheep), incongruent audiovisual stimuli (i.e., a picture of cow and a sound of sheep), and neutral audiovisual stimuli (i.e., a picture of cat and white noise), and the recognition for single auditory or visual stimulus followed the encoding stage. The results showed that subsequent recognition for sound stimuli was better during congruent audiovisual stimuli than neutral audiovisual stimuli or sounds presented in isolation. Similarly,

* Corresponding author at: 7th New Avenue West, Xinyang City, Henan Province, 464099, China.
E-mail address: xieyuanj@gmail.com (Y. Xie).

https://doi.org/10.1016/j.bbr.2021.113382
Received 23 February 2021; Received in revised form 3 May 2021; Accepted 21 May 2021
Available online 24 May 2021

other studies have found that recognition for visual stimuli is improved when they are initially presented with semantically congruent sounds compared to when the stimuli are presented with semantically incongruent sounds [6,7].

These studies mentioned above may suggest that audiovisual object encoding, particularly congruent audiovisual object encoding in working memory, can facilitate subsequent recognition of objects from a single sensory modality. However, when multiple objects (i.e., visual and auditory) or multisensory objects (i.e., congruent and incongruent audiovisual) were presented simultaneously during the encoding stage, it is difficult to infer that later memory improvement is the result of initial multisensory experiences because of a mixed presentation of multiple types of objects. Furthermore, the neural basis of multisensory working memory processing is not entirely clear. Few studies have examined brain activity during audiovisual object processing in working memory. Some studies have reported that the prefrontal cortex is essential for encoding face and voice [8,9], and the posterior parietal cortex is activated while maintaining a combination of picture and sound [8].

Resting-state functional magnetic resonance imaging (fMRI) signals reflect the intrinsic functional organization of brain [9,10]. Resting brain activity is remarkably similar to functional brain activity in cognitive tasks. It is believed that inherent brain activations can predict task-related brain activations [11–13]. Therefore, resting-state fMRI is a promising tool to examine spontaneous brain activity and its relationship with cognitive processing, with widely used regional homogeneity (ReHo). ReHo measures the local synchronization of time-series of blood oxygen level-dependent (BOLD) signal and reflects the similarities of one voxel with the nearest neighboring voxels [14], which are property of neurovascular coupling [15]. Studies have shown that ReHo is correlated with working memory performance [16–18]. Therefore, the present study aimed to expand the investigation of multisensory working memory processing and its association with resting-state ReHo in healthy subjects. We adopted a Sternberg-type working memory task using complex natural auditory and visual objects to examine the effect of audiovisual encoding on sequent memory recognition. The task paradigm was widely used in cognitive neuroscience research, which involves presenting a series of memory items at the encoding stage, a delay duration during which the information must be maintained, and then a probe item is delivered to evaluate the availability of the information after the delay. Only one audiovisual combination or unisensory object was presented in the encoding stage since the capacity of multisensory processing could be limited to one item [19]. Subsequently, the subjects were divided into two subgroups according to their behavioral performance (e.g., response time), and then two subgroups received resting-state fMRI scans. Finally, ReHo values in the regions of significant differences and their relationships with behavioral performance were examined.

2. Methods

2.1. Participants

One hundred and twenty-eight participants (mean age 23 years, 55 females and 73 males, all right-handed) took part in the experiment. The inclusion criteria were participants who had normal hearing and normal or corrected-to-normal vision, to be able to complete this experiment. The exclusion criteria were participants who had any history of psychiatric disorders (screened by an experienced psychiatrist according to the non-patient version of Structured Clinical Interview for DSM-IV), neurologic diseases, pregnancy, alcohol or tobacco substance abuse, head trauma, and contraindications for MRI examination. While the experiment task was completed, according to behavioral performance (response time of congruent AV condition, see the next section for more details), those who scored below the first percentile of response times (RT) were assigned to the higher-RT group (n = 32; female = 13, male = 19). Those who scored above the third percentile of response times were assigned to the lower-RT group (n = 32; female = 15, male = 17). Then both groups participated in resting-state functional magnetic resonance (fMRI) scans. All participants gave written informed consent and were compensated after completing the tasks. This study was conducted in accordance with the declaration of Helsinki and approved by the Medical Ethics Committee of Xijing Hospital.

2.2. Behavioral tasks

In this experiment task, single visual (V), single auditory (A), congruent (cAV), and incongruent (iAV) audiovisual objects were presented in the encoding stage and investigate the effect of audiovisual encoding on later unisensory memory recognition. The test stimuli were either an auditory or visual stimulus (fifty percent each) during the encoding of the audiovisual object. The details of the tasks were shown in Table 1 and Fig. 1.

The visual objects were selected from a standard set of figure line drawings [20,21] and available online (http://www.google.com/imghp). The chosen line drawings contained an equivalent number of objects from different semantic categories (e.g., animals, tools, instruments, etc.) and were divided equally across task conditions. All pictures were 300 × 300-pixel resolution. The auditory objects were the sounds corresponding to the line drawings taken from the Internet (http://www.findsounds.com). The sounds were modified using audio editing software (Adobe Audition version 5.0) and parameters set at 16-bit, 44,100 Hz digitization with each sound lasting 600 ms.

The tasks had four blocks (V, A, cAV, iAV); each block within the task had 30 trials. All blocks within the task were presented at random to balance task difficulty levels. The order of stimulus presentation in each block was pseudorandom to prevent predictability. During each block, visual stimuli were presented in the center of a black background on a 17-inch cathode ray tube monitor, the viewing distance from the monitor was maintained at 70 cm, and the visual stimuli extended an average of 6.5° in both the vertical and horizontal planes. The auditory stimuli were presented binaurally through headphones with no level difference to best match localization with the centrally presented visual stimuli, and the sound level was adjusted to the comfort level of each participant (approximately70 dB).

Participants were asked to judge whether the test stimulus was the same as that presented in the encoding stage by pressing one of two keys. Presented and un-presented test stimuli were split equally. A practice session of six trials for each block was administered prior to the main tasks. The stimuli in the practice sessions were not used in the main task. Breaks were encouraged between the blocks of the task to ensure high concentration and prevent fatigue.

2.3. Imaging data acquisition and preprocessing

A 3-Tesla scanner (Siemens, Munich, Germany) was used to gather the MRI data for both groups (Higher-RT group and Lower-RT group). All subjects were instructed to lay down on the bench with foam pads and earplugs. During the scans, they were asked to keep their eyes closed

Table 1
Design and descriptive statistics of the experiment task.

Block	Presentation	Abbreviation	RTs (Mean ± SD, ms)	ACRs (%, Mean ± SD)
1	Visual object	V	594.80 ± 125.90	96.86 ± 0.05
2	Auditory object	A	684.40 ± 150.90	93.24 ± 0.11
3	Congruent audiovisual object	cAV	556.70 ± 121.70	96.57 ± 0.07
4	Incongruent audiovisual object	iAV	672.80 ± 139.90	96.08 ± 0.09

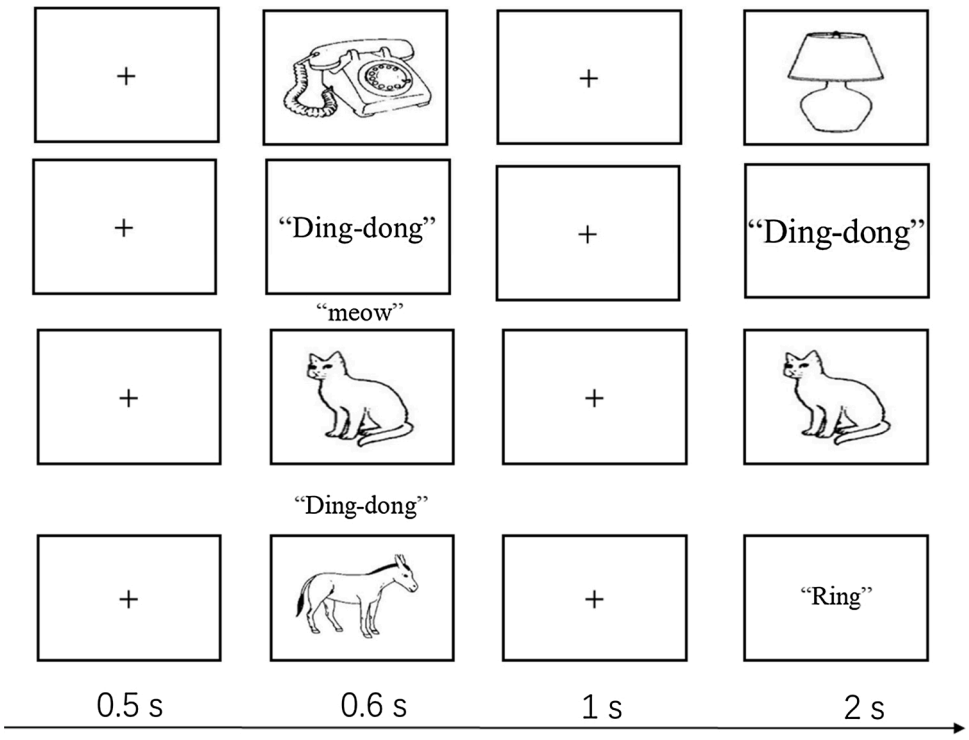

Fig. 1. Schematic diagram of the experiment task. A location fixation cross was first shown for 0.5 s, after which the unisensory or bimodal stimuli (duration of 0.6 s) were presented, a blank screen was shown as delayed for 1 s. Finally, the test stimulus appeared in 0.6, the limit of 2 s response time window. Inter-trial intervals varied randomly between 1.5 s and 3 s.

and awake. Each participant obtained resting-state fMRI scanning using an echo-planar imaging sequence with the following parameters: repletion time/echo time = 2000/30 ms, matrix = 64 × 64, the field of view (FOV) = 210 mm × 210 mm, 36 interleaved axial slices with a slice thickness of 4 mm and 240-time points which lasted about 8 min.

Resting-state fMRI data preprocessing was conducted using the DPABI (v5.1, http://rfmri.org/dpabi) toolbox, running on the MATLAB platform. The main preprocessing steps included: the first ten volumes were discarded for signal equilibrium, slice timing, and head realignment. The images were spatial normalized to a conventional Montreal Neurological Institute EPI template and resampled to 3 × 3 × 3 mm^3 isotropic voxels. Meanwhile, linear detrending and temporal bandpass filtering (0.01–0.1) were conducted. The head motion effects were further corrected with regression of realigned data on voxel-specific head motion parameters in a Fristion autoregressive strategy [22]: the 3 voxel-specific translation motion parameters in *x*, *y*, *z*, the same 3 parameters for one time point before, and the 6 corresponding squared items (voxel-specific 12-parameter model). If the head motion of participants were over 2.0 mm maximum translation in any direction or over 2.0 degrees of maximum rotation, the subjects would be excluded.

2.4. ReHo analysis

The Kendall's coefficient concordance (KCC)-ReHo calculation was performed by the DPABI toolbox. KCC values were generated by computing the consistency of the time series between a given voxel and its nearest neighbors (26 voxels). Then the individual ReHo maps were obtained by calculating individual KCC values within the whole brain voxels. Normalization of the ReHo map was conducted to decrease the effect of individual variation of KCC values. Finally, a spatial smooth with a Gaussian kernel of 6 mm was performed on the generated ReHo maps.

A two-sample *t*-test was performed to analyze the ReHo differences between the two groups on the individual smooth ReHo maps. Age, sex, educational year were used as covariates. Permutation test with threshold-free cluster enhancement (TFCE) was used to correct multiple comparisons at $P < 0.05$ by using the DPABI statistical module.

2.5. Statistical analysis

Behavioral data were analyzed using SPSS 20.0 (SPSS Inc., Chicago, IL, USA) by removing error trials and outlying data (greater than 3 SD of the mean value). Repeated measures analysis of variance (ANOVA) was used to test for differences in mean correct RT and percentages of accuracy rates (ACR) within the tasks. The Greenhouse–Geisser correction was applied to adjust for degrees of freedom. Post hoc comparisons were performed with Bonferroni correction. Parietal correlation analysis was conducted to investigate the correlation between ReHo values in the region of significant differences between the two groups and RT, with age, sex, and educational years as covariates.

3. Results

3.1. Behavioral data

RT and ACR within each block in the task were displayed in Table 1. All blocks were designed to be relatively easy. Thus, ACRs were greater than random probability. ANOVA revealed that ACRs did not differ significantly across the four conditions (F (3,125) = 2.112, $P > 0.0.5$, $\eta_p^2 = 0.046$) (Table 1 and Fig. 2A).

However, RT of memory recognition across the four conditions reached statistical significance (F (3,125) = 6.417, $P < 0.001$, $\eta_p^2 = 0.579$). Post hoc comparisons showed that RT of memory recognition during the cAV condition was faster than during single V condition (t (1,127) = 2.037, $P < 0.05$), A condition (t (1, 127) = 6.748, $P < 0.001$), and iAV condition (t (1,127) = 6.139, $P < 0.001$) (Table 1 and Fig. 2).

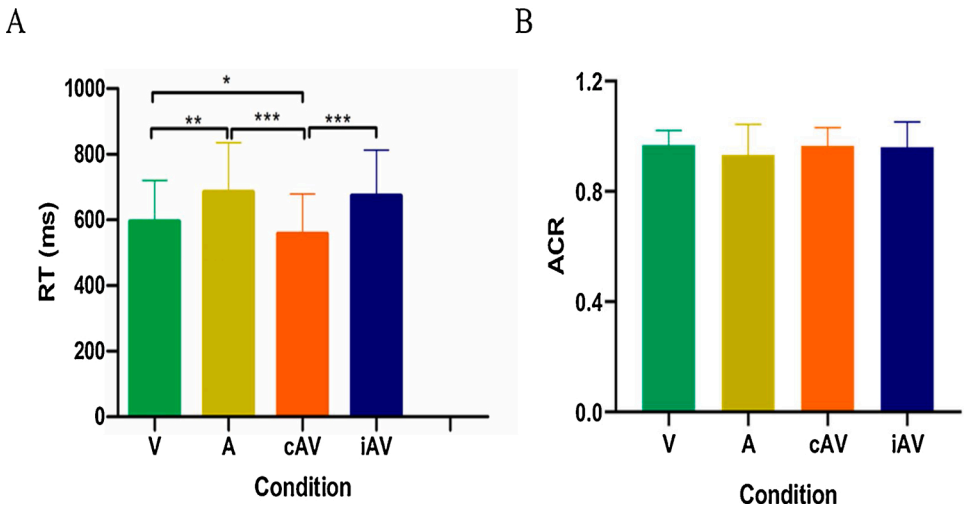

Fig. 2. A) Response time (RT) comparisons of memory recognition among the four conditions. B) Accuracy rate (ACR) comparisons of memory recognition among the four conditions. Error bars depict the standard deviation of the mean. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

3.2. ReHo differences

Compared with the higher-RT group, the lower-RT group had increased ReHo in the middle frontal gyrus (BA 46) and superior frontal gyrus (BA 9), which could be classified to executive control network (ECN) [23–25], while decreased ReHo in the left insula and supramarginal gyrus, which could be categorized to salience network (SN) [26–28], and in the left medial frontal gyrus (BA 32), right superior parietal lobule (BA 7), and left angular gyrus (BA 40), which could be assigned to default mode network (DMN) [29–31]. All cluster thresholds were corrected by TFCE ($P < 0.05$, cluster size > 50). Details are displayed in Table 2 and Fig. 3.

3.3. Correlations between ReHo and behavioral data

The partial correlations showed ReHo values in the left middle frontal gyrus ($r = -0.265$, $P = 0.035$) and left superior frontal gyrus ($r = -0.299$, $P = 0.017$) were negatively correlated with RT, using age, gender, and educational years as covariates. The ReHo values in the other brain areas, such as locating in the SN and DMN circuit, did not significantly correlate with RT ($P > 0.05$). Details are displayed in Fig. 4.

Table 2
Significant differences of ReHo between the Lower-RT group and higher-RT group (lower-RT group vs. higher-RT group).

Network	Location	BA	Peak coordinate			Cluster	*t* Value
			x	*y*	*z*		
ECN	Left middle frontal gyrus	46	−33	18	36	83	3.65
	Left superior frontal gyrus	9	−24	42	39	63	3.88
SN	Left insula	48	−45	3	6	81	−4.96
	Left superiormarginal gyrus	48	−51	−24	27	46	−2.86
DMN	Left medial frontal gyrus	32	−6	60	39	69	−3.56
	Right superior parietal lobule	7	33	−51	60	61	−3.30
	Left angular gyrus	40	−42	−54	42	286	−4.55

4. Discussion

The present study investigated the audiovisual working memory processing and its association with resting-state fMRI ReHo. The results showed that later unisensory memory recognition is faster during congruent audiovisual stimuli encoding than during the encoding of a single visual object, auditory object, or incongruent audiovisual object. During the congruent audiovisual object condition, the higher performance group showed increased ReHo in the ECN and decreased ReHo in the SN and DMN compared to the lower performance group. The ReHo values in major regions of ECN were correlated with the behavioral performance. These results suggested that audiovisual working memory was superior to memory recognition and may be associated with ECN.

In the field of cognitive neuroscience in working memory, researchers have concentrated on resolving whether the information is memorized in the form of modality-specific or integrated representations [32]. Atkinson and Shiffrin have proposed that incoming sensory information is processed by modality-specific sensory registers and transferred into an amodal "working memory" system (short-term store) [33]. Baddeley and Hitch have declared a multiple-component working memory model and emphasize that information is processed and stored in two modalities-specific subsystems (e.g., the phonological loop and the visuospatial sketchpad) [34]. Recently, an episodic buffer as the additional component has been added in the multiple-component model, which is responsible for combining memory traces from different modalities or subsystems into a coherent perceptual representation [35]. Therefore, we could speculate that semantically congruent audiovisual objects may be integrated into unified object representation during the encoding stage and further be stored in a specific subsystem such as episode buffer. The Bayes causal inference model contributes to elucidate the speculation. This model has demonstrated that the nervous system must depend on causal relationships to determine which signals can be integrated or not [36]. If the two sensory signals conveying the information have a common cause (i.e., they belong to the same object), they should be integrated; otherwise, they should be processed separately. The semantically congruent audiovisual objects represented an equivalent agent at an abstract level with different sensory inputs, so they could be fused as a unified representation. While incongruent audiovisual objects figured different representations at different abstract characteristics, then they could not be combined. Our previous study has also provided primary evidence to support this speculation [37].

Furthermore, a larger body of literature has demonstrated that multisensory integration shows behavioral benefits, including improved

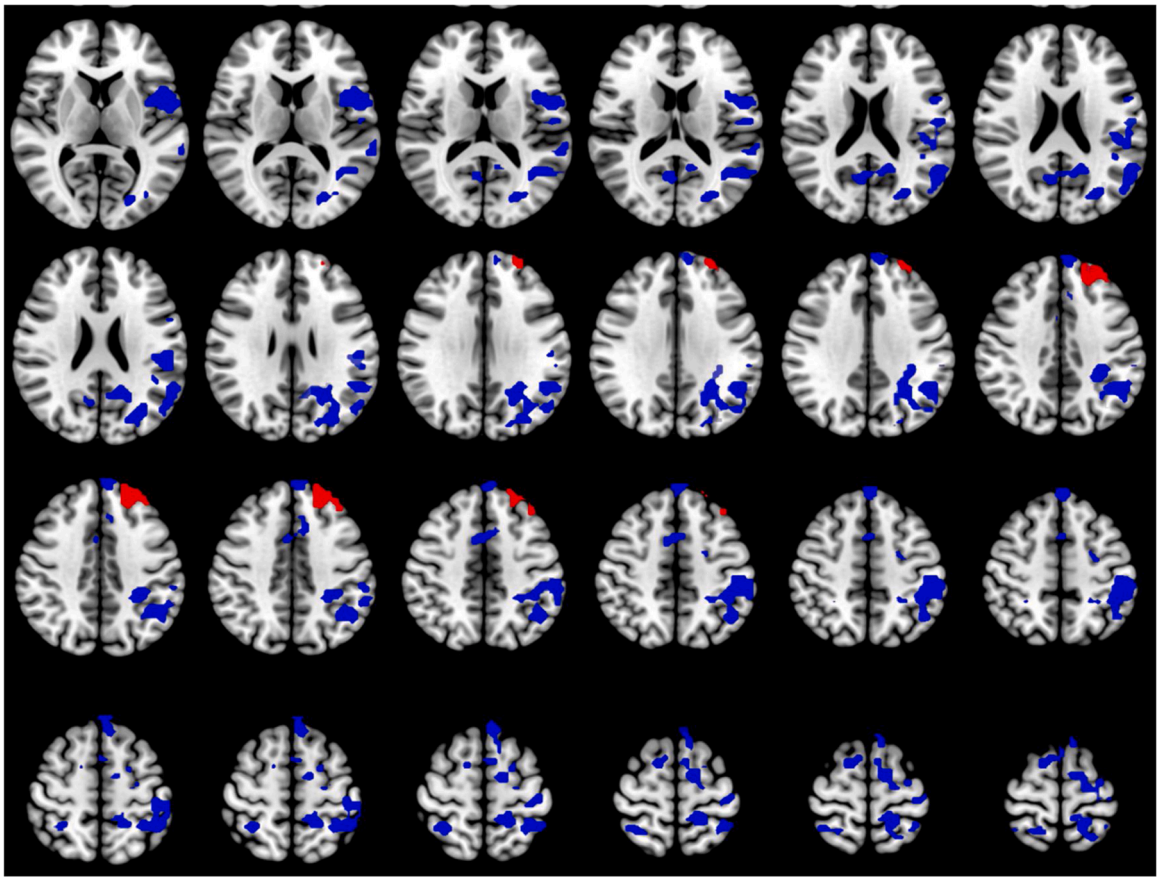

Fig. 3. The ReHo differences between the higher-RT group and lower-RT group. The warm color indicates increased ReHo, and the cold color indicates decreased ReHo (lower-RT group vs. higher-RT group) (TFCE correction, $P < 0.05$, cluster size > 50).

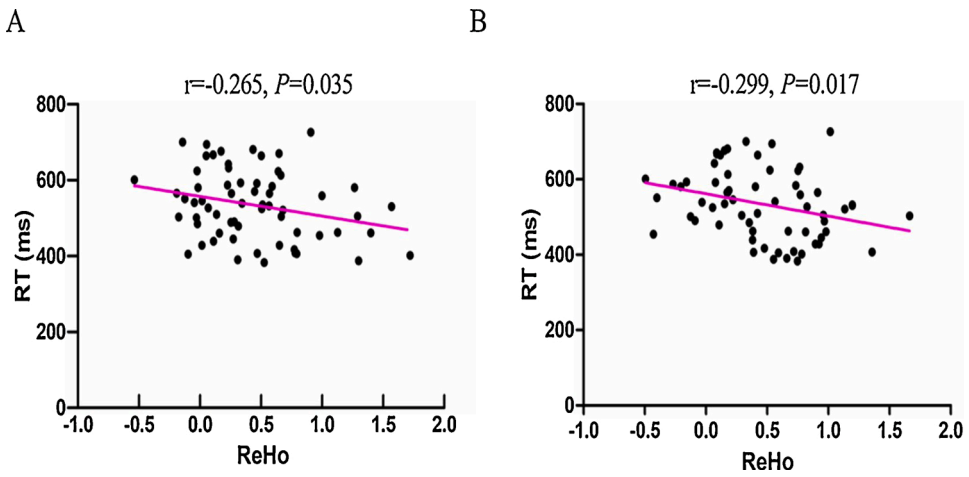

Fig. 4. The partial correlation between the ReHo values in the left middle frontal gyrus (A) and left superior frontal gyrus (B) and RT of congruent audiovisual object encoding condition in the whole-group ($P < 0.05$).

performance accuracy (Lehmann & Murray, 2005), reducing response time [38]. Thus, it is not surprising to find the semantically congruent audiovisual object encoding during working memory speeding up the later memory recognition. We also found that the later visual recognition during visual object encoding was faster than auditory recognition during auditory object encoding in the task. The result was similar to the previous study using a multisensory object recognition task [39], which reported that the semantic access for a visual object (e.g., picture) was faster than an auditory object (e.g., natural sound). Because the meaning of a picture can be retrieved about 100 ms after its onset [40], while it takes about 150 ms to extract the meaning of a natural sound after its onset [41].

The electrical neuroimaging studies provide evidence for multisensory brain activity [6,42,43]. Several studies reported the source modeling at the posterior parietal, temporal, and lateral occipital cortices during a continuous recognition task involve presenting pictures and sounds [6,44,45]. The finding has implicated these areas in perceptual advantage for multisensory object recognition. Recently, it

has been suggested that resting-state fMRI is a promising tool to examine the relation between spontaneous brain activity and behavioral performance [46–48]. Furthermore, ReHo is a promising index to measure the local neuronal synchronization of spontaneous fMRI BOLD signal [14]. Correlation analysis of ReHo and behavior could be a valuable tool for investigating the neural basis of individual behavioral differences. Studies have shown that there is a significant relationship between ReHo and cognitive performance (e.g., RT) [49–51], since ReHo might reflect common mechanisms of neurovascular coupling for both synchronized spontaneous and task-related activities [52]. In this study, we observed that the higher-RT group showed enhanced ReHo in the ECN nodes, including the middle frontal gyrus and superior frontal gyrus, compared to the lower-RT group, and higher ReHo values in these regions were correlated with faster RT of the audiovisual working memory task. The results suggest that the prefrontal cortex is critical for multisensory working memory processing. The prefrontal cortex has been known to implicate working memory [53,54], and the damage of this cortex is associated with deficit of working memory [55]. In addition, the prefrontal cortex has been thought to involve in multisensory processing and represent an organizational unit for multisensory representations [56–58]. Thus, the correlation between the ReHo value of the prefrontal cortex and RT of audiovisual working memory may suggest that this cortex is implicated with complex audiovisual object processing in working memory, since the ECN circuit is associated with demanding cognitive processes [26].

It has been shown that the major regions of DMN (e.g., medial prefrontal cortex and angular gyrus) and SN (e.g., insula) have higher ReHo than global mean during resting state [59,60], which suggests that spontaneous brain activities within the two networks have strong intra-regions synchronization. The high intra-region synchronization of spontaneous activity could be considered as a model to increase the efficiency of transferring information within nodes in the network [59]. While we found that the higher-RT group had decreased ReHo in the major nodes of DMN and SN compared with those of the lower-RT group. Studies have shown that there was reduced functional connectivity between the DMN and SN [61–63], although ReHo does not measure the synchronization of neuronal activity in remote regions. The results implied that participants with high-capability of multisensory processing had lower intra-regional synchronization of spontaneous activity within the networks. The characteristic could reflect an inhibition of information transition within the nodes in DMN and SN circuits to increase information processing efficiency in other networks such as the ECN, as the interactions of these networks appear fundament to higher-level cognition [64]. Particularly, the DMN and ECN exist anti-correlation [65], and loss of the correlation between the DMN and ECN has been associated with more inferior learning and memory [66,67]. All the results illustrated that spontaneous resting-state brain activities (e.g., ReHo) are coherent in multiple brain networks and could reflect neuronal information processing and be asscoatied with cognitive task performance such as audiovisual working memory.

In summary, we explored the audiovisual working memory processing of the healthy adults using ReHo analysis of resting-state fMRI data, expanding multisensory working memory research. The results showed that congruent audiovisual encoding sped up the later memory recognition in working memory. The resting-state fMRI ReHo could be used to characterize the neural correlations of multisensory working memory processing. Our findings reveal a possible relationship between the ECN circuit based on ReHo and multisensory processing in working memory during the wake for the first time.

Funding sources

This study was funded by the Humanities and Social Sciences Fund Youth Project of the Ministry of Education in China (No. 18YJC190026) and the National Nature Science Foundation of China (No.61806210).

Author contributions

Conception and design: XYJ. Collection, analysis, and data interpretation: XYJ, LYY, GMZ, DHD, XXL, FP. Drafting the article: XYJ. Revising the article: XYJ, GMZ, and FP. All authors approved the final version of manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

References

[1] V.A. Thompson, A. Paivio, Memory for pictures and sounds: independence of auditory and visual codes, Can. J. Exp. Psychol. 48 (3) (1994) 380–398.
[2] P. Goolkasian, P.W. Foos, Bimodal format effects in working memory, Am. J. Psychol. 118 (1) (2005) 61–78.
[3] A. Paivio, Dual coding theory: retrospect and current status, Can. J. Psychol. 45 (3) (1991) 255–287.
[4] J. Heikkila, K. Alho, H. Hyvonen, K. Tiippana, Audiovisual semantic congruency during encoding enhances memory performance, Exp. Psychol. 62 (2) (2015) 123–130.
[5] J. Heikkila, K. Alho, K. Tiippana, Semantically congruent visual stimuli can improve auditory memory, Multisens. Res. 30 (7–8) (2017) 639–651.
[6] A. Thelen, C. Cappe, M.M. Murray, Electrical neuroimaging of memory discrimination based on single-trial multisensory learning, NeuroImage 62 (3) (2012) 1478–1488.
[7] A. Thelen, D. Talsma, M.M. Murray, Single-trial multisensory memories affect later auditory and visual object discrimination, Cognition 138 (2015) 148–160.
[8] Y.J. Xie, Y.Y. Li, B. Xie, Y.Y. Xu, L. Peng, The neural basis of complex audiovisual objects maintenances in working memory, Neuropsychologia 133 (9) (2019) 1–9.
[9] M.D. Fox, M.E. Raichle, Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging, Nat. Rev. Neurosci. 8 (9) (2007) 700–711.
[10] M.E. Raichle, Two views of brain function, Trends Cognit. Sci. 14 (4) (2010) 180–190.
[11] M.W. Cole, T. Ito, D.S. Bassett, D.H. Schultz, Activity flow over resting-state networks shapes cognitive task activations, Nat. Neurosci. 19 (12) (2016) 1718–1726.
[12] J. Connolly, J.P. McNulty, L. Boran, R.A.P. Roche, D. Delany, A.L.W. Bokde, Identification of resting state networks involved in executive function, Brain Connect. 6 (5) (2016) 365–374.
[13] S.M. Smith, P.T. Fox, K.L. Miller, D.C. Glahn, P.M. Fox, C.E. Mackay, N. Filippini, K. E. Watkins, R. Toro, A.R. Laird, C.F. Beckmann, Correspondence of the brain's functional architecture during activation and rest, Proc. Natl. Acad. Sci. USA 106 (31) (2009) 13040–13045.
[14] Y. Zang, T. Jiang, Y. Lu, Y. He, L. Tian, Regional homogeneity approach to fMRI data analysis, NeuroImage 22 (1) (2004) 394–400.
[15] R. Yuan, X. Di, E.H. Kim, S. Barik, B. Rypma, B.B. Biswal, Regional homogeneity of resting-state fMRI contributes to both neurovascular and task activation variations, Magn. Reson. Imaging 31 (9) (2013) 1492–1500.
[16] H. Cheng, G. Sun, M. Li, M. Yin, H. Chen, Neuron loss and dysfunctionality in hippocampus explain aircraft noise induced working memory impairment: a resting-state fMRI study on military pilots, BioSci. Trends 13 (5) (2019) 430–440.
[17] X. Chen, Q. Zhang, J. Wang, J. Liu, W. Zhang, S. Qi, H. Xu, C. Li, J. Zhang, H. Zhao, S. Meng, D. Li, H. Lu, M. Aschner, B. Li, H. Yin, J. Chen, W. Luo, Cognitive and neuroimaging changes in healthy immigrants upon relocation to a high altitude: a panel study, Hum. Brain Mapp. 38 (8) (2017) 3865–3877.
[18] P. Wang, R. Li, J. Yu, Z. Huang, J. Li, Frequency-dependent brain regional homogeneity alterations in patients with mild cognitive impairment during working memory state relative to resting state, Front. Aging Neurosci. 8 (3) (2016) 1–12.
[19] E. Van der Burg, E. Awh, C.N.L. Olivers, The capacity of audiovisual integration is limited to one item, Psychol. Sci. 24 (3) (2013) 345–351.
[20] M.C. Sanfeliu, A. Fernandez, 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 (4) (1996) 537–555.
[21] J.G. Snodgrass, M. Vanderwart, A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity, J. Exp. Psychol.: Hum. Learn. Memory 6 (2) (1980) 174–215.
[22] K.J. Friston, S. Williams, R. Howard, R.S.J. Frackowiak, R. Turner, Movement-related effects in fMRI time-series, Magn. Reson. Med. 35 (3) (1996) 346–355.
[23] Q. Zhao, H. Lu, H. Metmer, W.X.Y. Li, J. Lu, Evaluating functional connectivity of executive control network and frontoparietal network in Alzheimer's disease, Brain Res. 1678 (2018) 262–272.
[24] R.C. McIntosh, R. Hoshi, J.S. Nomi, M. Di Bello, Z.T. Goodman, S. Kornfeld, L. Q. Uddin, C. Ottaviani, Neurovisceral integration in the executive control network: a resting state analysis, Biol. Psychol. 157 (11) (2020) 1–9.
[25] K.-K. Shen, T. Welton, M. Lyon, A.N. McCorkindale, G.T. Sutherland, S. Burnham, J. Fripp, R. Martins, S.M. Grieve, Structural core of the executive control network:

a high angular resolution diffusion MRI study, Hum. Brain Mapp. 41 (5) (2020) 1226–1236.
[26] W.W. Seeley, V. Menon, A.F. Schatzberg, J. Keller, G.H. Glover, H. Kenna, A. L. Reiss, M.D. Greicius, Dissociable intrinsic connectivity networks for salience processing and executive control, J. Neurosci. 27 (9) (2007) 2349–2356.
[27] M. Pievani, L. Pini, C. Ferrari, F.B. Pizzini, I. Boscolo Galazzo, C. Cobelli, M. Cotelli, R. Manenti, G.B. Frisoni, Coordinate-based meta-analysis of the default mode and salience network for target identification in non-invasive brain stimulation of alzheimer's disease and behavioral variant frontotemporal dementia networks, J. Alzheimer's Dis. 57 (2) (2017) 825–843.
[28] J.A. Muthulingam, T.M. Hansen, A.M. Drewes, S.S. Olesen, J.B. Frøkjær, Disrupted functional connectivity of default mode and salience networks in chronic pancreatitis patients, Clin. Neurophysiol. 131 (5) (2020) 1021–1029.
[29] M.E. Raichle, A.M. MacLeod, A.Z. Snyder, W.J. Powers, D.A. Gusnard, G. L. Shulman, A default mode of brain function, Proc. Natl. Acad. Sci. USA 98 (2) (2001) 676–682.
[30] B.J. Harrison, J. Pujol, M. López-Solà, R. Hernández-Ribas, J. Deus, H. Ortiz, C. Soriano-Mas, M. Yücel, C. Pantelis, N. Cardoner, Consistency and functional specialization in the default mode brain network, Proc. Natl. Acad. Sci. USA 105 (28) (2008) 9781–9786.
[31] M.D. Fox, A.Z. Snyder, J.L. Vincent, M. Corbetta, D.C. Van Essen, M.E. Raichle, The human brain is intrinsically organized into dynamic, anticorrelated functional networks, Proc. Natl. Acad. Sci. USA 102 (27) (2005) 9673–9678.
[32] M. Quak, R.E. London, D. Talsma, A multisensory perspective of working memory, Front. Hum. Neurosci. 9 (4) (2015) 197–208.
[33] R.C. Atkinson, R.M. Shiffrin, Human memory: a proposed system and its control processes, in: K.W. Spence, J.T. Spence (Eds.), Psychology of Learning and Motivation, Academic Press, 1968, pp. 89–195.
[34] A.D. Baddeley, G. Hitch, Working memory, in: G.H. Bower (Ed.), Psychology of Learning and Motivation, Academic Press, 1974, pp. 47–89.
[35] A. Baddeley, The episodic buffer: a new component of working memory? Trends Cognit. Sci. 4 (11) (2000) 417–423.
[36] O. Sporns, K.P. Körding, U. Beierholm, W.J. Ma, S. Quartz, J.B. Tenenbaum, L. Shams, Causal inference in multisensory perception, PLoS ONE 2 (9) (2007) e943.
[37] Y.J. Xie, Y.Y. Xu, C. Bian, M. Li, Semantic congruent audiovisual integration during the encoding stage of working memory: an ERP and sLORETA study, Sci. Rep. 7 (6) (2017) 1–10.
[38] B. Forster, C. Cavina-Pratesi, S.M. Aglioti, G. Berlucchi, Redundant target effect and intersensory facilitation from visual-tactile interactions in simple reaction time, Exp. Brain Res. 143 (4) (2002) 480–487.
[39] S. Molholm, W. Ritter, D.C. Javitt, J.J. Foxe, Multisensory visual–auditory object recognition in humans: a high-density electrical mapping study, Cereb. Cortex 14 (4) (2004) 452–465.
[40] M. Fabre-Thorpe, The characteristics and limits of rapid visual categorization, Front. Psychol. 2 (3) (2011) 243–254.
[41] M.M. Murray, L. Spierer, Auditory spatio-temporal brain dynamics and their consequences for multisensory interactions in humans, Hear. Res. 258 (1) (2009) 121–133.
[42] P.J. Matusz, M.T. Wallace, M.M. Murray, A multisensory perspective on object memory, Neuropsychologia 105 (2017) 243–252.
[43] M.M. Murray, J.F. John, G.R. Wylie, The brain uses single-trial multisensory memories to discriminate without awareness, Neuroimage 27 (2) (2005) 473–478.
[44] A. Thelen, P.J. Matusz, M.M. Murray, Multisensory context portends object memory, Curr. Biol. 24 (16) (2014) R734–R735.
[45] M.M. Murray, C.M. Michel, R.G. de Peralta, S. Ortigue, D. Brunet, S.G. Andino, A. Schnider, Rapid discrimination of visual and multisensory memories revealed by electrical neuroimaging, Neuroimage 21 (1) (2004) 125–135.
[46] M. Hampson, N.R. Driesen, P. Skudlarski, J.C. Gore, R.T. Constable, Brain connectivity related to working memory performance, J. Neurosci. 26 (51) (2006) 13338–13343.
[47] A.M.C. Kelly, L.Q. Uddin, B.B. Biswal, F.X. Castellanos, M.P. Milham, Competition between functional brain networks mediates behavioral variability, NeuroImage 39 (1) (2008) 527–537.
[48] M. Song, Y. Zhou, J. Li, Y. Liu, L. Tian, C. Yu, T. Jiang, Brain spontaneous functional connectivity and intelligence, NeuroImage 41 (3) (2008) 1168–1176.
[49] L. Tian, J. Ren, Y. Zang, Regional homogeneity of resting state fMRI signals predicts Stop signal task performance, NeuroImage 60 (1) (2012) 539–544.
[50] S. Diciotti, S. Orsolini, E. Salvadori, A. Giorgio, N. Toschi, S. Ciulli, A. Ginestroni, A. Poggesi, N. De Stefano, L. Pantoni, D. Inzitari, M. Mascalchi, Resting state fMRI regional homogeneity correlates with cognition measures in subcortical vascular cognitive impairment, J. Neurol. Sci. 373 (2017) 1–6.
[51] H.-H. Lee, S. Hsieh, Resting-state fMRI associated with stop-signal task performance in healthy middle-aged and elderly people, Front. Psychol. 8 (5) (2017) 1–11.
[52] R. Yuan, X. Di, E.H. Kim, S. Barik, B. Rypma, B.B. Biswal, Regional homogeneity of resting-state fMRI contributes to both neurovascular and task activation variations, Magn. Reson. Imaging 31 (9) (2013) 1492–1500.
[53] S. Funahashi, Prefrontal cortex and working memory processes, Neuroscience 139 (1) (2006) 251–261.
[54] C.E. Curtis, M. D'Esposito, Persistent activity in the prefrontal cortex during working memory, Trends Cognit. Sci. 7 (9) (2003) 415–423.
[55] A.K. Barbey, M. Koenigs, J. Grafman, Dorsolateral prefrontal contributions to human working memory, Cortex 49 (5) (2013) 1195–1205.
[56] J.S. Anderson, M.A. Ferguson, M. Lopez-Larson, D. Yurgelun-Todd, Topographic maps of multisensory attention, Proc. Natl. Acad. Sci. USA 107 (46) (2010) 20110–20114.
[57] T. Chen, L. Michels, K. Supekar, J. Kochalka, S. Ryali, V. Menon, Role of the anterior insular cortex in integrative causal signaling during multisensory auditory–visual attention, Eur. J. Neurosci. 41 (2) (2015) 264–274.
[58] J.A. Johnson, A.P. Strafella, R.J. Zatorre, The role of the dorsolateral prefrontal cortex in bimodal divided attention: two transcranial magnetic stimulation studies, J. Cognit. Neurosci. 19 (6) (2007) 907–920.
[59] X.-Y. Long, X.-N. Zuo, V. Kiviniemi, Y. Yang, Q.-H. Zou, C.-Z. Zhu, T.-Z. Jiang, H. Yang, Q.-Y. Gong, L. Wang, K.-C. Li, S. Xie, Y.-F. Zang, Default mode network as revealed with multiple methods for resting-state functional MRI analysis, J. Neurosci. Methods 171 (2) (2008) 349–355.
[60] D. Liu, C. Yan, J. Ren, L. Yao, V. Kiviniemi, Y. Zang, Using coherence to measure regional homogeneity of resting-state fMRI signal, Front. Syst. Neurosci. 4 (6) (2010) 1–9.
[61] W. Chiong, S.M. Wilson, M. D'Esposito, A.S. Kayser, S.N. Grossman, P. Poorzand, W.W. Seeley, B.L. Miller, K.P. Rankin, The salience network causally influences default mode network activity during moral reasoning, Brain 136 (6) (2013) 1929–1941.
[62] F. Orliac, M. Naveau, M. Joliot, N. Delcroix, A. Razafimandimby, P. Brazo, S. Dollfus, P. Delamillieure, Links among resting-state default-mode network, salience network, and symptomatology in schizophrenia, Schizophr. Res. 148 (1) (2013) 74–80.
[63] H. Chen, Y. Li, Q. Liu, Q. Shi, J. Wang, H. Shen, X. Chen, J. Ma, L. Ai, Y.M. Zhang, Abnormal interactions of the salience network, central executive network, and default-mode network in patients with different cognitive impairment loads caused by leukoaraiosis, Front. Neural Circuits 13 (6) (2019) 1–8.
[64] K.K. Ng, J.C. Lo, J.K.W. Lim, M.W.L. Chee, J. Zhou, Reduced functional segregation between the default mode network and the executive control network in healthy older adults: a longitudinal study, NeuroImage 133 (3) (2016) 321–330.
[65] H. Gu, Y. Hu, X. Chen, Y. He, Y. Yang, Regional excitation-inhibition balance predicts default-mode network deactivation via functional connectivity, NeuroImage 185 (2019) 388–397.
[66] K. Mevel, B. Landeau, M. Fouquet, R. La Joie, N. Villain, F. Mézenge, A. Perrotin, F. Eustache, B. Desgranges, G. Chételat, Age effect on the default mode network, inner thoughts, and cognitive abilities, Neurobiol. Aging 34 (4) (2013) 1292–1301.
[67] J.R. Andrews-Hanna, A.Z. Snyder, J.L. Vincent, C. Lustig, D. Head, E. Marcus, R. L. Raichle, Buckner, Disruption of large-scale brain systems in advanced aging, Neuron 56 (5) (2007) 924–935.