

# Visual association learning induces global network reorganization

Mengxia Yu <sup>a</sup>, Xueting Li <sup>b, \*\*</sup>, Yiyong Song <sup>c, \*</sup>, Jia Liu <sup>d</sup>

<sup>a</sup> Bilingual Cognition and Development Lab, Center for Linguistics and Applied Linguistics, Guangdong University of Foreign Studies, Guangzhou, 510420, China

<sup>b</sup> Department of Psychology, Renmin University of China, Beijing, 100872, China

<sup>c</sup> Beijing Key Laboratory of Applied Experimental Psychology, Faculty of Psychology, Beijing Normal University, Beijing, 100875, China

<sup>d</sup> Department of Psychology & Tsinghua Laboratory of Brain and Intelligence, Tsinghua University, Beijing, 100084, China

## ARTICLE INFO

### Keywords:

Visual association learning  
Multivariate connectivity pattern analysis  
Network reorganization  
Top-down processing

## ABSTRACT

It has been proposed that visual learning is accomplished not only by neural plasticity in the visual cortex, but also by complex interactions between bottom-up and top-down processes that may induce global network reorganization. Here, we applied a multivariate analysis to functional connectivity (FC) patterns across the brain to investigate how visual association learning was achieved through large-scale network reorganization. Participants were trained to associate a set of artificial line-drawing objects with English letters. After three consecutive days of training, participants underwent a functional magnetic resonance imaging scan in which they were presented with the trained stimuli, untrained stimuli, and English words. By calculating pairwise FC between 189 nodes of 10 well-established networks across the brain, we found that the visual association learning induced changes in the global FC pattern when viewing the trained stimuli, rendering it more similar to the FC pattern when viewing English words. Critically, the learning-induced global FC pattern differences were mainly driven by the FC related to the high-level networks involved in attention and cognitive control, suggesting the modification of top-down processes during learning. In sum, our study provides one of the first evidence revealing global network reorganization induced by visual learning and sheds new light on the network mechanisms of top-down influences in learning.

## 1. Introduction

Extensive findings indicate that visual learning can improve behavior performance in the training tasks and modify the visual system of humans and animals (Li, 2016; Watanabe and Sasaki, 2015). For example, training to discriminate basic visual properties (e.g., grating orientation, contrast, motion direction) or more complex visual objects modifies neuronal responses in the early or higher-order visual cortical areas (Baker et al., 2007; de Beeck et al., 2006; Golby et al., 2001; Kobatake et al., 1998; Logothetis et al., 1995; A. Schoups, Vogels, Qian and Orban, 2001; A. A. Schoups, Vogels and Orban, 1995; Shiu and Pashler, 1992; Sigman et al., 2005). However, it is increasingly acknowledged that visual learning is a complex and constructive process and is not mediated by local changes in the visual cortex alone.

Accumulating evidence indicates that top-down influences play a pivotal role in visual learning. Behaviorally, psychophysics studies demonstrate that the learning effect of discriminating grating orientations or motion directions can transfer to untrained location or stimulus

properties with double-training procedures, suggesting the involvement of high-level cognitive processes (Wang et al., 2016; Wang et al., 2012; Xiao et al., 2008). The role of top-down influences in visual learning is also supported by physiological and neuroimaging studies showing task-dependent modulation of responses in the visual areas when two different tasks are trained for an identical set of stimuli (Li et al., 2004; Song et al., 2010b). More direct evidence for the involvement of top-down processes in visual learning come from neuroimaging studies showing that training of visual tasks induced activation changes in high-level fronto-parietal areas responsible for attentional control and decision making, in addition to changes in the visual areas (Kahnt et al., 2011; Lewis et al., 2009; Mukai et al., 2007; Sigman et al., 2005). Visual learning also strengthens functional connectivity (FC) between the visual areas and fronto-parietal areas (Lewis et al., 2009; Mukai et al., 2007). Therefore, convergent evidence suggests that visual learning induces complex interactions between bottom-up sensory processing and top-down cognitive control, which may manifest as large-scale network reorganization in addition to local changes. Here we tested

\* Corresponding author. Faculty of Psychology, Beijing Normal University, Beijing, 100875, China.

\*\* Corresponding author. Department of Psychology, Renmin University of China, Beijing, 100872, China.

E-mail addresses: [lixueting02@gmail.com](mailto:lixueting02@gmail.com) (X. Li), [songyiyong@bnu.edu.cn](mailto:songyiyong@bnu.edu.cn) (Y. Song).

this hypothesis by exploring the plasticity of global FC patterns across the brain with a multivariate approach (Dresler et al., 2017; Tambini et al., 2017).

To do this, participants were trained to associate a set of artificial line-drawing objects with English letters for three consecutive days, and underwent an fMRI scan after learning. Then, we calculated pairwise FC between 189 nodes from 10 well-established networks across the whole brain (Cole et al., 2013; Power et al., 2011) when participants viewed English words, the trained and untrained stimuli respectively. Instead of examining univariate changes in the FC for the trained stimuli, we used an FC pattern similarity analysis (Dresler et al., 2017; Tambini et al., 2017) to compare the multivariate FC pattern for the trained stimuli to a template provided by the FC pattern for English words. We tested whether the global FC pattern for the trained objects became more similar to the learning template and, more importantly, whether the learning-induced changes in global FC pattern were driven by the connectivity related to the high-level control networks or low-level sensory networks. Although changes in local activation have usually been observed in low-level visual system, we predicted that the training would induce distributed differences in the FC pattern of the high-level control networks if top-down influences indeed play a critical role in visual learning.

## 2. Materials and methods

### 2.1. Participants

Twelve college students (six females, aged 21–28) with normal or corrected-to-normal vision participated in the study. All participants were native Chinese speakers who have studied English for at least 10 years. None of the participants had any history of neurological or psychiatric disorders. The fMRI protocol was approved by the IRB of the Institute of Biophysics, Chinese Academy of Sciences. All participants provided written informed consent and were paid for their participation. Part of the dataset was reported in our previous study with analyses of regional activation (Song et al., 2010a).

### 2.2. Behavioral training

Participants were trained with a visual association task to learn 16 paired associates, in which each line-drawing figure was paired with either a consonant (e.g., b, Fig. 1A) or two vowels/one vowel + one consonant (e.g., ie, ar, Fig. 1A). This design allowed us to create stimuli that corresponded to three-letter English words (e.g., bar) by combining

two trained figures (Fig. 1C) for the fMRI scanning session. Participants were first shown all 16 paired associates to be familiar with them before the training task. Then, they performed a two-alternative forced choice task to determine whether the presented stimulus pair was one of the correct associates. Each trial started with a blank screen for 800 ms, followed by a stimulus pair, which was presented until a response was made (Fig. 1B). Auditory feedback was given to indicate whether the response was correct (high pitch) or incorrect (low pitch). Half the trial contained the correct paired associates, whereas the other half contained the incorrect ones. There were 480 trials in each training session and the training ended when the reaction time (RT) reached asymptote (i.e., no significant decrease in at least three consecutive sessions). On average, all participants completed at least ten training sessions, which took about 3–4 h for two to three successive days. One participant's behavioral data was accidentally lost due to improper saving, and there were eleven participants left for behavioral data analysis.

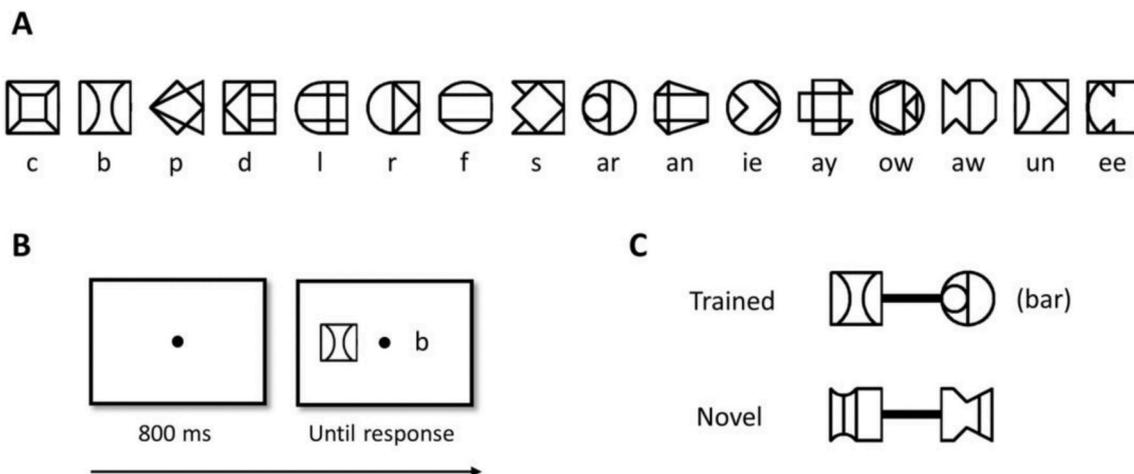
### 2.3. fMRI scanning

After behavioral training, each participant completed an MRI scanning session, consisting of (1) three blocked-design functional localizer runs (three of the participants participated in two runs) and (2) three blocked-design experimental runs. There were five stimuli categories in the localizer runs, including English words, Chinese characters, frontal-view faces, line-drawing objects, and scrambled line-drawing objects. The experimental runs consisted of four sets of dumbbell-shaped stimuli, each created by combining two trained or untrained figures with a connection bar. One set of dumbbell-shaped stimuli were formed by trained figures, which corresponded to English words (Trained, Fig. 1C), and another set of stimuli were formed by 16 untrained figures (Novel, Fig. 1C) to serve as the baseline. The English words corresponding to the trained stimuli were different from those used in the localizer runs. The other two sets of dumbbell-shaped stimuli were formed by trained stimuli too, but did not correspond to English words.

Each localizer or experimental run consisted of four 15-s blocks for each stimuli type and five 15-s blocks of fixation. Participants were asked to press a button whenever they saw two identical stimuli in a row (for more details on the paradigm, see (Song et al., 2010a)).

### 2.4. MRI data acquisition

Images were acquired using a 3T Siemens Trio scanner with an eight-channel phased-array head coil at Beijing Imaging Center for Brain Research in the Institute of Biophysics, Chinese Academy of Sciences.



**Fig. 1. Stimuli and behavioral training.** A, Participants learned 16 paired associates in which each line-drawing figure was paired with either an English consonant or two vowels. B, Visual association learning procedure. Participants conducted a two-alternative forced-choice task to determine whether a stimulus pair was one of the actual paired associates. C, Examples of the stimuli in fMRI experimental runs. Dumbbell-shaped stimuli were created by combining two trained or novel figures.

Thirty 2.3-mm-thick (20% skip) near axial slices were collected (in-plane resolution =  $1.4 \times 1.4$  mm) and oriented parallel to each subject's temporal cortex. A T2\*-weighted gradient-echo echo-planar-imaging (EPI) sequence was used (TR = 3000 ms, TE = 32 ms, flip angle =  $90^\circ$ ). In addition, three-dimensional (3D) structural images were acquired with MPRAGE, an inversion prepared gradient echo sequence (TR/TE/TI = 2730/3.44/1000 ms, flip angle =  $7^\circ$ , voxel size =  $1.1 \times 1.1 \times 1.9$  mm<sup>3</sup>).

## 2.5. fMRI data preprocessing

The functional images were preprocessed with the FMRI Expert Analysis Tool (FEAT) of the Oxford Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB) Software Library (FSL, <http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/>). The preprocessing was conducted separately on each run for each participant and included the following steps: motion correction, brain extraction, spatial smoothing with a 4-mm FWHM Gaussian kernel, intensity normalization, and high-pass temporal filtering (0.01 Hz). To reduce further the physiological noise, such as fluctuations caused by motion, cardiac and respiratory cycles, nuisance signals from cerebrospinal fluid, white matter, motion correction parameters, and first derivatives of these signals were regressed out using the methods described in previous studies (Biswal et al., 2010; Fox et al., 2005). Registration of each participant's fMRI images to the structural images was carried out using FMRIB's linear image registration tool (FLIRT). Registration of each participant's structural images to the Montreal Neurological Institute (MNI) space ( $2 \times 2 \times 2$  mm<sup>3</sup>) was accomplished using FMRIB's nonlinear image registration tool (FNIRT) (Jenkinson et al., 2002; Jenkinson and Smith, 2001).

## 2.6. Definition of nodes and networks

We used the nodes defined in Power et al. (2011) to construct the whole brain networks, which have been shown to provide high test-retest reliability of global network properties (Cao et al., 2014) and demonstrated to be representative for global networks (Gordon et al., 2014; Mohr et al., 2016; Power et al., 2011). Of the original 264 nodes in Power et al. (2011), 227 nodes have been assigned to 10 networks well-established in previous studies, comprising the sensorimotor, visual, auditory, salience, subcortical networks, default mode network (DMN), ventral and dorsal attention networks (VAN and DAN), and frontal-parietal and cingulo-opercular task control networks (FPN and CON) (Cole et al., 2013; Mohr et al., 2016). We then defined 5-mm-radius spheres centered on the coordinates of all the 227 nodes reported by Power et al. (2011) and projected them to the brain of each of our participants. Several nodes were discarded due to incomplete coverage of parietal and frontal lobes during functional data acquisition across participants, with 189 remaining nodes distributed in the 10 networks (sensorimotor: 24/35, visual: 28/31, auditory: 13/13, salience: 18/18, subcortical: 13/13, DMN: 44/58, VAN: 8/9, DAN: 9/10, FPN: 21/25 and CON: 10/14).

## 2.7. FC calculation

Subsequently, time courses of each non-fixation condition and the fixation condition were concatenated from all runs for each node. This was done for the localizer and experimental runs separately. We extracted the mean time series for each node and calculated Pearson correlation between the time series of all nodes, producing a  $189 \times 189$  FC matrix of r-values for each condition and each participant. Figure S1 (Supporting Information) displayed the FC matrices for the English condition in the localizer runs, the trained and novel conditions in the experimental runs, and their corresponding fixation conditions for an exemplar participant. The r-values were then transformed to z-scores using Fisher's z-transformation. Previous research has shown that the

functional network structure is highly similar across a range of task states and resting state (Cole et al., 2013; Geerligts et al., 2015; Krienen et al., 2014), which was also observed in our results (Supporting Information, Figure S1). To remove the common network structure underlying the task and resting states, we performed paired-t tests on the z-score FC matrices between each non-fixation condition and their corresponding fixation condition across participants, producing a  $189 \times 189$  t-score FC matrix for each non-fixation condition for the following analyses (Fig. 2A and B).

## 2.8. FC pattern similarity analysis

The t-score FC matrix during the presentation of English Words (vs. fixation) in the localizer runs was used as a template FC pattern (Fig. 2A). Pattern similarity was evaluated by calculating spatial correlation between the template FC matrix and the FC matrix of each condition in the experimental runs (Fig. 2B and C). The learning effect was examined by comparing the pattern similarity between the trained stimuli and English words ( $r(E,T)$ ) with that between the novel stimuli and English words ( $r(E,N)$ ). Statistical significance of the pattern similarity difference was examined with a permutation test that constructed a null distribution for the difference between the two r-values (Fig. 2D). We randomly permuted the FC matrices of the trained and novel conditions respectively for each participant; the permuted FC matrices for the trained and novel conditions were correlated with the FC matrix of English words and then the two r values were subtracted. The permutation was replicated for 5000 times, resulting in a null distribution for the difference between the two r values. The p value was estimated by assessing the proportion of r-value differences in the null distribution that was higher than the real r-value differences.

## 3. Results

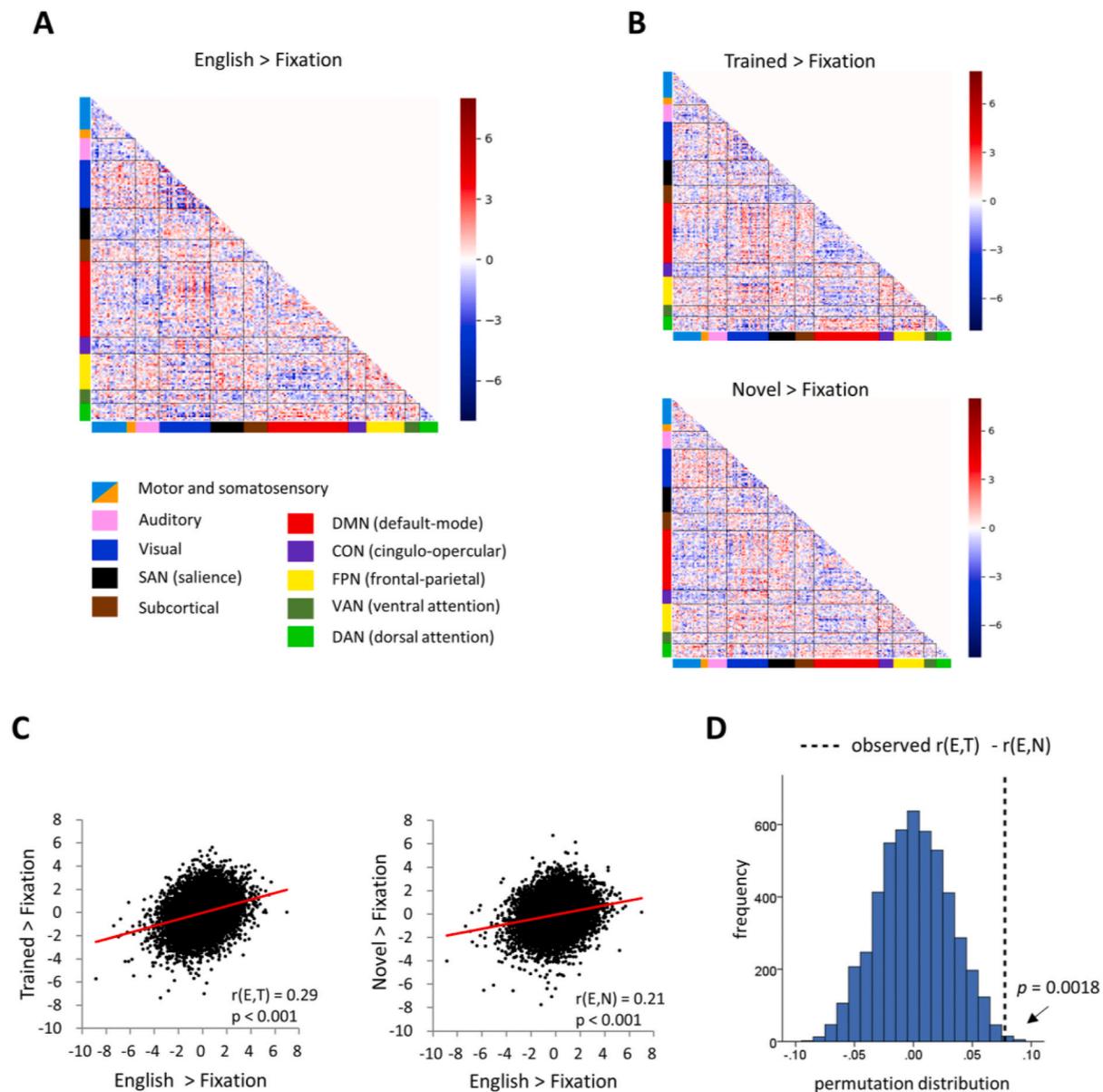
### 3.1. Behavioral results

Participants were trained to learn the paired associates between novel line-drawing objects and English letters in a two-alternative forced-choice (2AFC) task for three consecutive days (Fig. 1A and B). As expected, the training greatly improved the behavioral performance of the participants. The reaction times (RTs) in association judgment decreased monotonically from session 1 to session 6 ( $F_{(5,50)} = 13.38, p < 0.001$ ) and then reached an asymptote from session 7 to session 10 ( $F_{(3,30)} < 1$ ). As the participants had been familiarized with the paired associates before the 2AFC task, the mean accuracy remained above 90% during the entire training process and did not differ across training sessions ( $F_{(9,90)} < 1$ ).

### 3.2. Learning induced differences in global organization of brain networks

During the fMRI scanning after learning, a set of dumbbell-shaped objects were presented by combining two trained figures with a connection bar, which corresponded to English words (trained, Fig. 1C). Another set of dumbbell-shaped objects were created in a similar way by combining two untrained figures (novel, Fig. 1C) to serve as the baseline.

First, we examined whether the visual association training modulated brain network organization by rendering the global FC pattern for the trained stimuli more similar to the template (i.e., FC pattern for the English words) after learning. To do this, we calculated pairwise FC between 189 nodes belonging to ten networks across the brain to construct the FC matrices. We used the t-score FC matrix for English words in the localizer runs (English words vs. fixation, Fig. 2A) as the template, and calculated the spatial correlation between the template and the t-score FC matrices for the trained (vs. fixation, Fig. 2B) and novel conditions (vs. fixation, Fig. 2B) in the experimental runs respectively. We found that the FC matrices for both the trained and



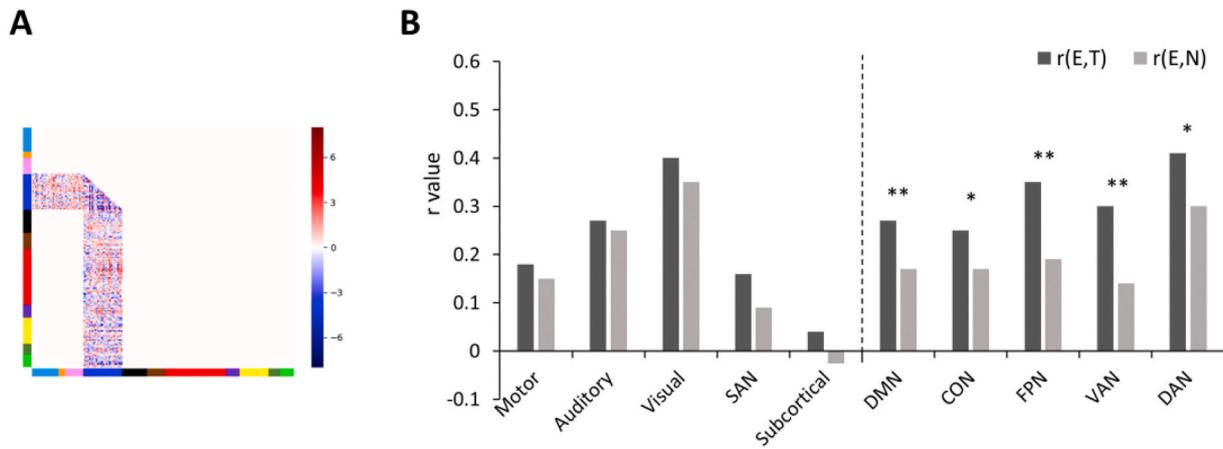
**Fig. 2.** Learning-induced differences in global FC pattern. **A**, T-score FC matrix for English words in the localizer runs (English vs. fixation). **B**, T-score FC matrices for the trained (vs. fixation, upper) and novel stimuli (vs. fixation, lower) in the experimental runs. **C**, Scatterplots of the correlations between the FC matrices for English words and the trained ( $r(E,T)$ , left) or novel ( $r(E,N)$ , right) stimuli. **D**, A permutation test was used to examine the statistical significance of the difference between  $r(E,T)$  and  $r(E,N)$ , and the histogram shows the null distribution of the 5000 permutations.

novel conditions showed significant correlation to the FC matrix for English words (Fig. 2C;  $r(E,T) = 0.29$ ,  $p < 0.001$ ;  $r(E,N) = 0.21$ ,  $p < 0.001$ ), implicating some shared network patterns generally involved in viewing English words and line-drawing objects. Critically, the correlation between the FC matrices for English words and the trained stimuli was greater than that between English words and the novel stimuli (Fig. 2D; permutation  $p = 0.002$ ), indicating that FC pattern for the trained stimuli was more similar to English words after training compared with that for the novel stimuli. This result suggested that learning to associate objects with English words induced the global network organization for the trained stimuli to resemble that for English words.

### 3.3. Learning modulated the interactions between high- and low-level networks

Next, we investigated whether the difference in global FC pattern

between the trained and untrained stimuli mainly resulted from modulation of the connectivity related to the high-level networks involved in attention and cognitive control or low-level networks involved in sensory processing. To do this, we examined the learning-induced FC changes for each of the ten networks. Specifically, for each network, we collected all the FCs relating to the nodes in the network, comprising all the FCs within this network and those between this network and the other nine networks (e.g., Fig. 3A). FC pattern similarity analyses were performed for each network in the same manner as for the whole-brain FC patterns. Interestingly, of the ten networks, the results showed a clear distinction between high- and low-level networks. That is, we found greater FC pattern similarity between English words and the trained stimuli than the novel stimuli for five high-level networks, including the cognitive control networks (cingulo-opercular network (CON) and fronto-parietal network (FPN)), dorsal and ventral attention networks (DAN and VAN), and the default mode network (DMN) (permutation  $p < 0.05$ , Fig. 3B). In contrast, no significant learning effect was found for

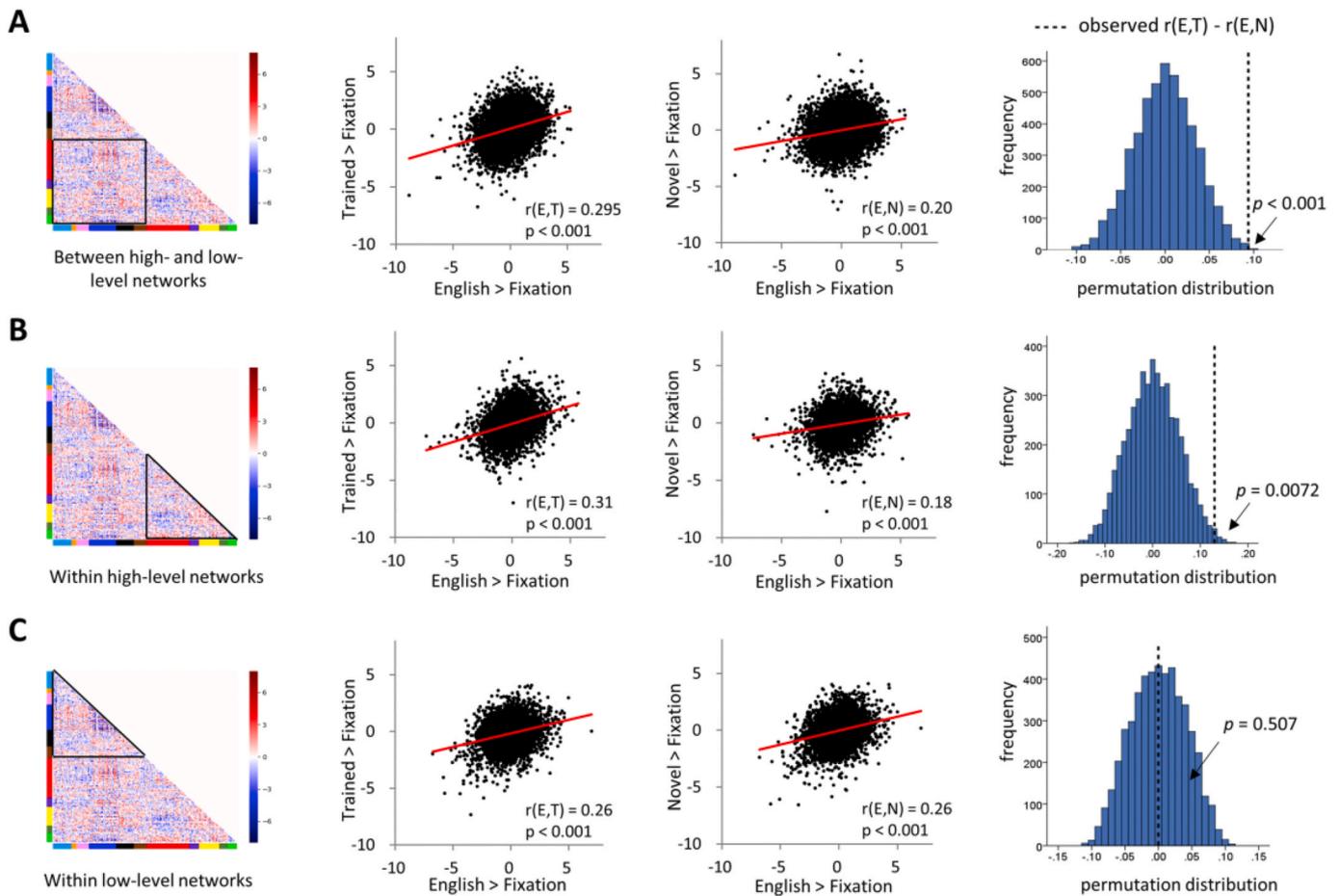


**Fig. 3. Learning-induced differences in FC pattern of each network.** **A**, The FC pattern of an exemplar network (visual) that included the FC within this network and that between this network and all the other networks. **B**, The correlations between the FC patterns for English words and the trained ( $r(E,T)$ ) or novel ( $r(E,N)$ ) stimuli for each network.  $*p < 0.05$ ,  $**p < 0.01$ , permutation test.

the remaining five low-level networks, including the visual, auditory, sensorimotor, subcortical, and salience (SAN) networks (permutation  $ps > 0.05$ , Fig. 3B). These results suggested that the learning-induced network reorganization was mainly driven by the connectivity relating to the high-level networks involved in cognitive control and attention deployment.

Since the FC pattern changes relating to high-level networks

observed above contained both the FCs within high-level networks and those between high- and low-level networks (e.g., Fig. 3A), we next asked whether learning induced modifications of the interactions between high- and low-level networks, or communications within high-level networks, or both. Based on the above results, we merged the five high-level networks (i.e., CON, FPN, DAN, VAN, and DMN, in total 93 nodes), as well as the five low-level networks (i.e., visual, auditory,



**Fig. 4. Learning-induced differences in FC patterns relating to high-level networks.** The left column illustrate the FC patterns between high- and low-level networks (A), within high-level networks (B), and within low-level networks (C). Scatterplots in the middle columns show correlations between the FC patterns for English words and the trained or novel condition. Histograms in the right column show the results of the permutation tests.

sensorimotor, subcortical and SAN, in total 96 nodes), and examined the learning effect for all the FCs between the high- and low-level networks (Fig. 4A, left), within the high-level networks (Fig. 4B, left), and within the low-level networks (Fig. 4C, left) respectively.

First, we examined the FC pattern between high- and low-level networks (Fig. 4A). We found that the trained stimuli showed greater similarity with English words than the novel stimuli ( $r(E,T) = 0.295, p < 0.001$ ;  $r(E,N) = 0.20, p < 0.001$ ; permutation  $p < 0.001$ ; Fig. 4A), indicating modification of the interactions between high- and low-level networks induced by visual association learning. Of the five low-level networks, the visual and sensorimotor networks were more directly involved in the training task. We further inspected the FC patterns between individual networks to test whether the training altered the interactions between high-level networks and the visual and sensorimotor networks. We found greater similarity between English words and the trained stimuli than the novel stimuli for the FC pattern between the VAN and sensorimotor network (permutation  $p = 0.006$ , FDR corrected, Supporting Information, Figure S2), and a similar trend was observed for the FC pattern between the FPN and visual network (permutation  $p = 0.055$ , Supporting Information, Figure S2). Taken together, these results suggested an interplay between top-down and bottom-up processes during visual association learning.

Next, we examined the FC pattern within the high-level networks, and the result showed higher similarity between English words and the trained stimuli than the novel stimuli ( $r(E,T) = 0.31, p < 0.001$ ;  $r(E,N) = 0.18, p < 0.001$ ; permutation  $p = 0.006$ ; Fig. 4B). This result indicated reorganization among high-level networks induced by visual association learning. Further inspection of FC patterns between individual networks revealed significant learning effect ( $r(E,T) > r(E,N)$ ) for the FC patterns of FPN-VAN, FPN-DAN and FPN-DMN (permutation  $ps < 0.05$ , FDR corrected, Supporting Information, Figure S3), indicating modification of interactions among the cognitive control and attention networks.

Finally, consistent with the above results, there was no learning effect within the low-level networks ( $r(E,T) = 0.26, p < 0.001$ ;  $r(E,N) = 0.26, p < 0.001$ ; permutation  $p = 0.507$ ; Fig. 4C). The absence of learning effect within the low-level networks might result from some networks less involved in the training task, such as auditory and subcortical networks. Therefore, we examined the learning effect for all the FCs among the nodes in the visual network only (28 nodes). Still, there was no difference in the FC pattern similarity with English words between the trained and novel stimuli ( $r(E,T) = 0.73, p < 0.001$ ;  $r(E,N) = 0.70, p < 0.001$ ; permutation  $p = 0.40$ ). Taken together with the learning effect observed between the FPN and visual network, these results suggested that visual association learning might rely more on adjustment of modulations from high-level cognitive control areas to visual areas than on modification of interactions among the visual regions.

In addition, we examined whether the observed FC pattern difference between the trained and novel stimuli was attributed to the general effect of familiarity. Participants were presented with stimuli corresponded to English pseudo-words (PW, e.g., cie) during scanning, which were also composed of trained figures and thus familiar to participants. However, we found that the PW condition showed lower similarity with English words than RW for both the FC pattern between high- and low-level networks ( $r(E,PW) = 0.22, p < 0.001$ ; permutation  $p = 0.014$ ) and that within high-level networks ( $r(E,PW) = 0.243, p < 0.001$ ; permutation  $p = 0.041$ ), and did not differ from novel condition (permutation  $ps > 0.05$ ). These results indicated that the observed learning effect of RW could not be fully accounted for by familiarity.

### 3.4. Transfer of the learning effects to other language

Next, we asked whether the learning effect of associating novel objects with English letters was specific to English or could generalize to other language familiar to the participants (i.e. Chinese). We performed similar FC pattern analyses for Chinese characters as for English words,

using the FC matrix when participants viewing Chinese characters as the template. We found that the trained stimuli showed greater similarity with Chinese characters than the novel stimuli ( $r(C,T) = 0.28, p < 0.001$ ;  $r(C,N) = 0.22, p < 0.001$ , Fig. 5A; permutation  $p = 0.015$ , Fig. 5B). This result indicated that the whole-brain network organization for the trained stimuli became more similar not only to English, but also to Chinese. Meanwhile, no significant difference was found between the trained and novel stimuli in the FC pattern similarity with face or line-drawing object condition (permutation  $ps > 0.05$ ), indicating that the trained stimuli were not processed similarly as other line-drawing objects or faces. Taken together, these results suggested that the network organization pattern for the trained stimuli resembled that for language processing.

Further analyses were conducted for the FC patterns within high-level, within low-level, and between high- and low-level networks separately. Replicating the results for English words, we found greater resemblance between the trained stimuli and Chinese characters than the novel stimuli for the FC pattern within high-level networks (permutation  $p = 0.041$ ; Fig. 5C) and that between high- and low-level networks (permutation  $p = 0.005$ ; Fig. 5C), but not for the FC pattern within low-level networks (permutation  $p = 0.252$ ; Fig. 5C). Taken together, the generalization of the learning effects to Chinese supported the role of top-down modulations in visual association learning, that is, the artificial objects were associated with general language information, rather than visual properties of English letters per se, during learning.

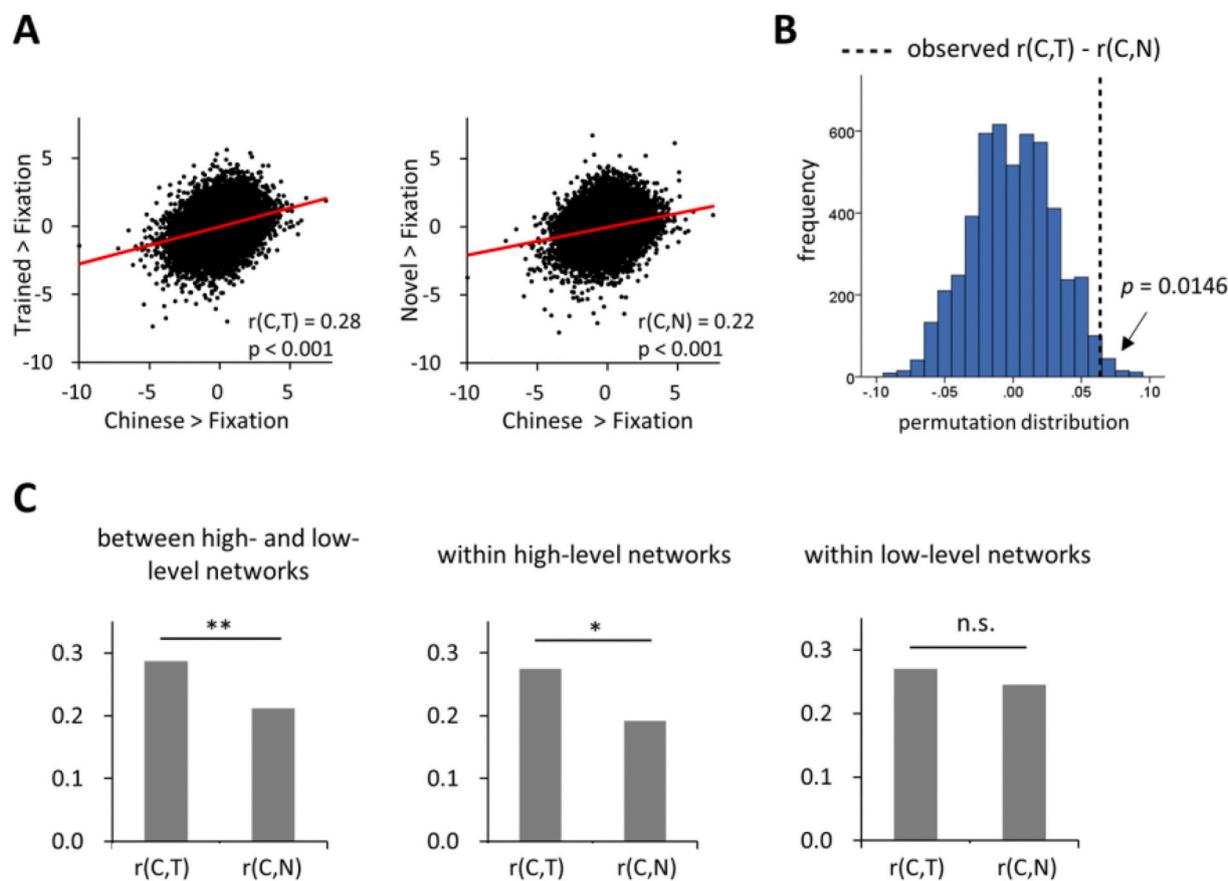
## 4. Discussion

In the current study, we applied the multivariate analysis to the whole-brain FC profiles to investigate how the learning goal of visual association training was achieved through large-scale network reorganization. We found that the training rendered the global FC pattern when viewing the trained stimuli more similar to the FC pattern when viewing English words. More importantly, further analyses showed that the learning-induced differences in global FC pattern arose from the FC between high-level cognitive control and attention networks and low-level visual and sensorimotor networks, as well as the FC among high-level networks. These findings suggested interactions of bottom-up and top-down processes and highlighted the pivotal role of top-down modulation in visual learning.

Our study provided the first evidence that visual association learning was supported by changes in whole-brain FC profiles that engaged the majority of the ten networks examined here, displaying a clear picture of global network reorganization in visual learning. Our study extended previous findings by showing learning-induced connectivity changes not only among a small number of regions, but widely distributed across multiple networks (Chen et al., 2015; Lewis et al., 2009; Mukai et al., 2007). More importantly, while previous studies with univariate approach revealed learning effects as either increase or decrease of FC among different regions which are difficult to interpret, our study used the multivariate pattern similarity analysis to reveal the FC changes that contribute directly to the achievement of the learning goal.

Further, the changes in global FC pattern were mainly driven by the connectivity related to the high-level networks, including the FC between high- and low-level networks as well as that among high-level networks. In contrast, no network reorganization was found in low-level networks. Therefore, although changes in neural activation are usually observed in sensory and perceptual cortical regions (i.e., low-level cortical regions) in visual learning, the learning may actually occur in high-level networks, and then propagate back to low-level cortical regions. In line with this speculation, a recent physiological study shows that visual-avoidance association learning enhances top-down projection from high-level area to the mouse V1 and adjusts the balance between bottom-up and top-down inputs in the V1 (Makino and Komiyama, 2015).

More specifically, the high-level networks modified by learning were



**Fig. 5. Transfer of learning effects to Chinese.** **A**, Scatterplots showing the correlations between the FC matrices for Chinese characters and the trained ( $r(C,T)$ ) or novel stimuli ( $r(C,N)$ ). **B**, Result of the permutation test for the difference between  $r(C,T)$  and  $r(C,N)$ . **C**, The correlations between the FC patterns for Chinese characters and the trained ( $r(C,T)$ ) or novel stimuli ( $r(C,N)$ ) for the FC between high- and low-level networks, within high-level networks, and within low-level networks. \* $p < 0.05$ , \*\* $p < 0.01$ , n.s. not significant, permutation test.

those engaged in task control and attention deployment (i.e., FPN, VAN, DAN, and DMN). Thus, our study revealed how the top-down influences of the task and attention factors might be implemented through network reorganization during learning. The influence of task context has been indicated by task specificity of visual learning, that is, the learning effect can hardly transfer across tasks even if the trained stimuli are identical (Ahissar and Hochstein, 1993; Saffell and Matthews, 2003; Shiu and Pashler, 1992). Such behavioral task specificity corresponds well to the task-dependent neural response changes in visual areas for identical trained stimuli (Li et al., 2004; Song et al., 2010b). Top-down attention plays a role in task specificity by selectively attending to task-relevant features and ignoring irrelevant features (Vidnyánszky and Sohn, 2005), and the importance of attention is also indicated by the findings that passive exposure to the stimulus can hardly produce learning effects (Li et al., 2006). In particular, our previous study found that learning of the same set of line-drawing objects increases activation in the VWFA with a visual association task but produces greater activation in regions processing general objects with a discrimination task, indicating the modulation of task context and attention in learning (Song et al., 2010b). Critically, here we further revealed the network reorganization mechanisms that may mediate these top-down modulation effects of the task and attention factors.

Different models have been proposed to account for the mechanisms of visual learning. While the input selection models assume that learning refines neural representation of the trained stimuli in the low-level sensory areas (Fahle, 2004), the readout reweighting models postulate that learning improves readout of visual inputs through reweighting connections between low-level sensory signals and high-level decision unit (Doshier and Lu, 1998; Kahnt et al., 2011; Law and Gold, 2008;

Petrov et al., 2005). The altered FC pattern between high- and low-level networks revealed in the present study fits nicely with the reweighting models, especially those proposing that visual learning mainly occur in high-level areas since the reweighting rules are learned by the decision unit in high-level areas (Zhang et al., 2010). Notably, the mechanisms of visual representation refinement and connectivity reweighting are not mutually exclusive. The visual association task used in our study not only modifies FC pattern between high- and low-level networks, but also alters the responses for the trained stimuli in the visual cortex (Song et al., 2010a). That is, the responses in the visual area may reflect an association-based representation resulting from top-down modulation during learning. In addition, our results are also in agreement with the predictive coding models in which learning leads to the generation and enhancement of top-down predictions that modulate bottom-up processing to minimize predictive errors (Makino and Komiyama, 2015).

Finally, the large-scale network reorganization revealed in visual learning may be a general mechanism shared by various types of learning. For example, a long-term motor learning over six weeks induces reconfiguration of distributed networks involving the motor, visual, and cognitive control networks (Bassett et al., 2015). Even the rapid practice of novel task rules over less than 2 min is associated with large-scale network reorganization among multiple high-level networks including the FPN, DAN, CON, and DMN (Mohr et al., 2016). These findings suggest that interplay between top-down and bottom-up processes is generally responsible for learning from different modalities. Future studies are invited to reveal the global network reorganization mechanisms for more learning tasks and compare them across various learning types to deepen our understanding of brain plasticity in general. In addition to the implications for learning studies, global network

reorganization has also broad implications for visual cognition (Bola and Sabel, 2015) and clinical situations (Bola et al., 2014; Bola et al., 2015). For example, EEG studies have shown global network reorganization in patients with optic nerve damage which is related to their vision loss and restoration (Bola et al., 2014; Bola et al., 2015). Future studies are invited to reveal the global network reorganization mechanisms for more learning tasks and variant clinical situations to deepen our understanding of brain plasticity in general.

### Credit author statement

M.Y., Y.S. and J.L. designed research. Y.S. and X.L. performed research. M.Y. and Y.S. analyzed the data. M.Y., X.L., Y.S. and J.L. wrote the paper.

### Declaration of competing interest

The authors declare no competing financial interests.

### Acknowledgements

This study was funded by the National Natural Science Foundation of China (31861143039 and 31872786), the National Basic Research Program of China (2018YFC0810602).

### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.neuropsychologia.2021.107789>.

### References

- Ahissar, M., Hochstein, S., 1993. Attentional control of early perceptual learning. *Proc. Natl. Acad. Sci. Unit. States Am.* 90 (12), 5718–5722.
- Baker, C.L., Liu, J., Wald, L.L., Kwong, K.K., Benner, T., Kanwisher, N., 2007. Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proc. Natl. Acad. Sci. Unit. States Am.* 104 (21), 9087–9092.
- Bassett, D.S., Yang, M., Wymbs, N.F., Grafton, S.T., 2015. Learning-induced autonomy of sensorimotor systems. *Nat. Neurosci.* 18 (5), 744–751.
- Biswal, B.B., Mennes, M., Zuo, X.-N., Gohel, S., Kelly, C., Smith, S.M., Colcombe, S., 2010. Toward discovery science of human brain function. *Proc. Natl. Acad. Sci. Unit. States Am.* 107 (10), 4734–4739.
- Bola, M., Gall, C., Moewes, C., Fedorov, A., Hinrichs, H., Sabel, B.A., 2014. Brain functional connectivity network breakdown and restoration in blindness. *Neurology* 83 (6), 542–551.
- Bola, M., Gall, C., Sabel, B.A., 2015. Disturbed temporal dynamics of brain synchronization in vision loss. *Cortex* 67, 134–146.
- Bola, M., Sabel, B.A., 2015. Dynamic reorganization of brain functional networks during cognition. *Neuroimage* 114, 398–413.
- Cao, H., Plichta, M.M., Schäfer, A., Haddad, L., Grimm, O., Schneider, M., Tost, H., 2014. Test-retest reliability of fMRI-based graph theoretical properties during working memory, emotion processing, and resting state. *Neuroimage* 84, 888–900.
- Chen, N., Bi, T., Zhou, T., Li, S., Liu, Z., Fang, F., 2015. Sharpened cortical tuning and enhanced cortico-cortical communication contribute to the long-term neural mechanisms of visual motion perceptual learning. *Neuroimage* 115, 17–29.
- Cole, M.W., Reynolds, J.R., Power, J.D., Repovs, G., Anticevic, A., Braver, T.S., 2013. Multi-task connectivity reveals flexible hubs for adaptive task control. *Nat. Neurosci.* 16 (9), 1348–1355.
- de Breeck, H.P.O., Baker, C.L., DiCarlo, J.J., Kanwisher, N.G., 2006. Discrimination training alters object representations in human extrastriate cortex. *J. Neurosci.* 26 (50), 13025–13036.
- Doshier, B.A., Lu, Z.-L., 1998. Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. *Proc. Natl. Acad. Sci. Unit. States Am.* 95 (23), 13988–13993.
- Dresler, M., Shirer, W.R., Konrad, B.N., Müller, N.C., Wagner, I.C., Fernández, G., Greicius, M.D., 2017. Mnemonic training reshapes brain networks to support superior memory. *Neuron* 93 (5), 1227–1235.
- Fahle, M., 2004. Perceptual learning: a case for early selection. *J. Vis.* 4 (10), 4.
- Fox, M.D., Snyder, A.Z., Vincent, J.L., Corbetta, M., Van Essen, D.C., Raichle, M.E., 2005. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc. Natl. Acad. Sci. U.S.A.* 102 (27), 9673–9678.
- Geerligs, L., Rubinov, M., Henson, R.N., 2015. State and trait components of functional connectivity: individual differences vary with mental state. *J. Neurosci.* 35 (41), 13949–13961.
- Golby, A.J., Gabrieli, J.D., Chiao, J.Y., Eberhardt, J.L., 2001. Differential responses in the fusiform region to same-race and other-race faces. *Nat. Neurosci.* 4 (8), 845–850.
- Gordon, E.M., Laumann, T.O., Adeyemo, B., Huckins, J.F., Kelley, W.M., Petersen, S.E., 2014. Generation and evaluation of a cortical area parcellation from resting-state correlations. *Cerebr. Cortex* 26 (1), 288–303.
- Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 17 (2), 825–841.
- Jenkinson, M., Smith, S., 2001. A global optimisation method for robust affine registration of brain images. *Med. Image Anal.* 5 (2), 143–156.
- Kahnt, T., Grueschow, M., Speck, O., Haynes, J.-D., 2011. Perceptual learning and decision-making in human medial frontal cortex. *Neuron* 70 (3), 549–559.
- Kobatake, E., Wang, G., Tanaka, K., 1998. Effects of shape-discrimination training on the selectivity of inferotemporal cells in adult monkeys. *J. Neurophysiol.* 80 (1), 324–330.
- Krienen, F.M., Yeo, B.T., Buckner, R.L., 2014. Reconfigurable task-dependent functional coupling modes cluster around a core functional architecture. *Phil. Trans. R. Soc. B* 369 (1653), 20130526.
- Law, C.-T., Gold, J.I., 2008. Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. *Nat. Neurosci.* 11 (4), 505–513.
- Lewis, C.M., Baldassarre, A., Committer, G., Romani, G.L., Corbetta, M., 2009. Learning sculpt the spontaneous activity of the resting human brain. *Proc. Natl. Acad. Sci. Unit. States Am.* 106 (41), 17558–17563.
- Li, W., 2016. Perceptual learning: use-dependent cortical plasticity. *Annual Review of Vision Science* 2, 109–130.
- Li, W., Piéch, V., Gilbert, C.D., 2004. Perceptual learning and top-down influences in primary visual cortex. *Nat. Neurosci.* 7 (6), 651–657.
- Li, W., Piéch, V., Gilbert, C.D., 2006. Contour saliency in primary visual cortex. *Neuron* 50 (6), 951–962.
- Logothetis, N.K., Pauls, J., Poggio, T., 1995. Shape representation in the inferior temporal cortex of monkeys. *Curr. Biol.* 5 (5), 552–563.
- Makino, H., Komiyama, T., 2015. Learning enhances the relative impact of top-down processing in the visual cortex. *Nat. Neurosci.* 18 (8), 1116–1122.
- Mohr, H., Wolfensteller, U., Betzel, R.F., Mišić, B., Sporns, O., Richiardi, J., Ruge, H., 2016. Integration and segregation of large-scale brain networks during short-term task automatization. *Nat. Commun.* 7, 13217.
- Mukai, I., Kim, D., Fukunaga, M., Japee, S., Marrett, S., Ungerleider, L.G., 2007. Activations in visual and attention-related areas predict and correlate with the degree of perceptual learning. *J. Neurosci.* 27 (42), 11401–11411.
- Petrov, A.A., Doshier, B.A., Lu, Z.-L., 2005. The dynamics of perceptual learning: an incremental reweighting model. *Psychol. Rev.* 112 (4), 715.
- Power, J.D., Cohen, A.L., Nelson, S.M., Wig, G.S., Barnes, K.A., Church, J.A., Schlaggar, B.L., 2011. Functional network organization of the human brain. *Neuron* 72 (4), 665–678.
- Saffell, T., Matthews, N., 2003. Task-specific perceptual learning on speed and direction discrimination. *Vis. Res.* 43 (12), 1365–1374.
- Schoups, A., Vogels, R., Qian, N., Orban, G., 2001. Practising orientation identification improves orientation coding in V1 neurons. *Nature* 412 (6846), 549–553.
- Schoups, A.A., Vogels, R., Orban, G.A., 1995. Human perceptual learning in identifying the oblique orientation: retinotopy, orientation specificity and monocularity. *J. Physiol.* 483 (3), 797–810.
- Shiu, L.-P., Pashler, H., 1992. Improvement in line orientation discrimination is retinally local but dependent on cognitive set. *Atten. Percept. Psychophys.* 52 (5), 582–588.
- Sigman, M., Pan, H., Yang, Y., Stern, E., Silbersweig, D., Gilbert, C.D., 2005. Top-down reorganization of activity in the visual pathway after learning a shape identification task. *Neuron* 46 (5), 823–835.
- Song, Y., Bu, Y., Hu, S., Luo, Y., Liu, J., 2010a. Short-term language experience shapes the plasticity of the visual word form area. *Brain Res.* 1316, 83–91.
- Song, Y., Hu, S., Li, X., Li, W., Liu, J., 2010b. The role of top-down task context in learning to perceive objects. *J. Neurosci.* 30 (29), 9869–9876.
- Tambini, A., Rimmele, U., Phelps, E.A., Davachi, L., 2017. Emotional brain states carry over and enhance future memory formation. *Nat. Neurosci.* 20 (2), 271.
- Vidnyánszky, Z., Sohn, W., 2005. Learning to suppress task-irrelevant visual stimuli with attention. *Vis. Res.* 45 (6), 677–685.
- Wang, R., Wang, J., Zhang, J.-Y., Xie, X.-Y., Yang, Y.-X., Luo, S.-H., Li, W., 2016. Perceptual learning at a conceptual level. *J. Neurosci.* 36 (7), 2238–2246.
- Wang, R., Zhang, J.-Y., Klein, S.A., Levi, D.M., Yu, C., 2012. Task relevancy and demand modulate double-training enabled transfer of perceptual learning. *Vis. Res.* 61, 33–38.
- Watanabe, T., Sasaki, Y., 2015. Perceptual learning: toward a comprehensive theory. *Annu. Rev. Psychol.* 66, 197–221.
- Xiao, L.-Q., Zhang, J.-Y., Wang, R., Klein, S.A., Levi, D.M., Yu, C., 2008. Complete transfer of perceptual learning across retinal locations enabled by double training. *Curr. Biol.* 18 (24), 1922–1926.
- Zhang, T., Xiao, L.-Q., Klein, S.A., Levi, D.M., Yu, C., 2010. Decoupling location specificity from perceptual learning of orientation discrimination. *Vis. Res.* 50 (4), 368–374.