NeuroImage 247 (2022) 118864

Contents lists available at ScienceDirect

NeuroImage

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

An awareness-dependent mapping of saliency in the human visual system

Lijuan Wang [a,b,c], Ling Huang [a,b,c], Mengsha Li [a,b,c], Xiaotong Wang [a,b,c], Shiyu Wang [a,b,c], Yuefa Lin [a,b,c], Xilin Zhang [a,b,c,d,e,*]

[a] Philosophy and Social Science Laboratory of Reading and Development in Children and Adolescents (South China Normal University), Ministry of Education, Guangzhou, Guangdong 510631, China
[b] Key Laboratory of Brain, Cognition and Education Sciences (South China Normal University), Ministry of Education, Guangzhou, Guangdong 510631, China
[c] School of Psychology, South China Normal University, Guangzhou, Guangdong 510631, China
[d] Center for Studies of Psychological Application, South China Normal University, Guangzhou, Guangdong 510631, China
[e] Guangdong Provincial Key Laboratory of Mental Health and Cognitive Science, South China Normal University, Guangzhou, Guangdong 510631, China

ARTICLE INFO

Keywords:
Saliency map
Awareness
Graded manner
Non-graded manner
fMRI

ABSTRACT

The allocation of exogenously cued spatial attention is governed by a saliency map. Yet, how salience is mapped when multiple salient stimuli are present simultaneously, and how this mapping interacts with awareness remains unclear. These questions were addressed here using either visible or invisible displays presenting two foreground stimuli (whose bars were oriented differently from the bars in the otherwise uniform background): a high salience target and a distractor of varied, lesser salience. Interference, or not, by the distractor with the effective salience of the target served to index a graded or non-graded nature of salience mapping, respectively. The invisible and visible displays were empirically validated by a two-alternative forced choice test (detecting the quadrant of the target) demonstrating subjects' performance at or above chance level, respectively. By combining psychophysics, fMRI, and effective connectivity analysis, we found a graded distribution of salience with awareness, changing to a non-graded distribution without awareness. Crucially, we further revealed that the graded distribution was contingent upon feedback from the posterior intraparietal sulcus (pIPS, especially from the right pIPS), whereas the non-graded distribution was innate to V1. Together, this awareness-dependent mapping of saliency reconciles several previous, seemingly contradictory findings regarding the nature of the saliency map.

1. Introduction

Attentional selection is the mechanism by which a subset of incoming information is processed preferentially. Numerous studies have demonstrated that this attentional selection can either be executed voluntarily by top-down signals derived from goals, such as when directing gaze to an interesting book (Baluch and Itti, 2011; Corbetta and Shulman, 2002; Kanwisher and Wojciulik, 2000; Kastner and Ungerleider, 2000; Serences and Yantis, 2006; Zhang et al., 2016; Zhang et al., 2018) or automatically by bottom-up signals from salient stimuli, such as a vertical bar among horizontal bars (Corbetta and Shulman, 2002; Hegdé and Felleman, 2003; Jonides, 1981; Koch and Ullman, 1985; Nakayama and Mackeben, 1989). Throughout this study, we use the term 'salience' to refer to this bottom-up attraction of attentional selection. While a 'saliency map' is defined as a topographical map to describe and predict the distribution of this bottom-up attraction based on a visual input (Itti and Koch, 2001; Koch and Ullman, 1985). Although the bottom-up attraction is typically quick and potent (Jonides, 1981; Nakayama and Mackeben, 1989), little is known about how this bottom-up attraction will be distributed in the human visual system when multiple salient stimuli are presented simultaneously and instantaneously.

A dominant model of the saliency map developed by Itti and Koch (2001) presumed that the bottom-up attention is sequentially allocated, in a winner-take-all (WTA) manner, to the most salient location from multiple salient regions, which was then suppressed by the inhibition-of-return mechanism (Klein, 2000) so that the bottom-up attention can focus onto the next most salient location (Koch and Ullman, 1985; Wolfe, 1994), and repeating this process generates attentional scanpaths in the stimulus. Accordingly, when a stimulus contains two salient regions with different levels of saliency (for example, 25° and 90° orientation contrasts, Fig. 1 **A**); the bottom-up attention will be captured by the more salient (i.e., the 90° orientation contrast) region at any given moment in time. To date, however, this model lacks

* Corresponding author at: Philosophy and Social Science Laboratory of Reading and Development in Children and Adolescents (South China Normal University), Ministry of Education, Guangzhou, Guangdong 510631, China.
E-mail address: xlzhang@m.scnu.edu.cn (X. Zhang).

https://doi.org/10.1016/j.neuroimage.2021.118864.
Received 2 March 2021; Received in revised form 20 December 2021; Accepted 25 December 2021
Available online 26 December 2021.

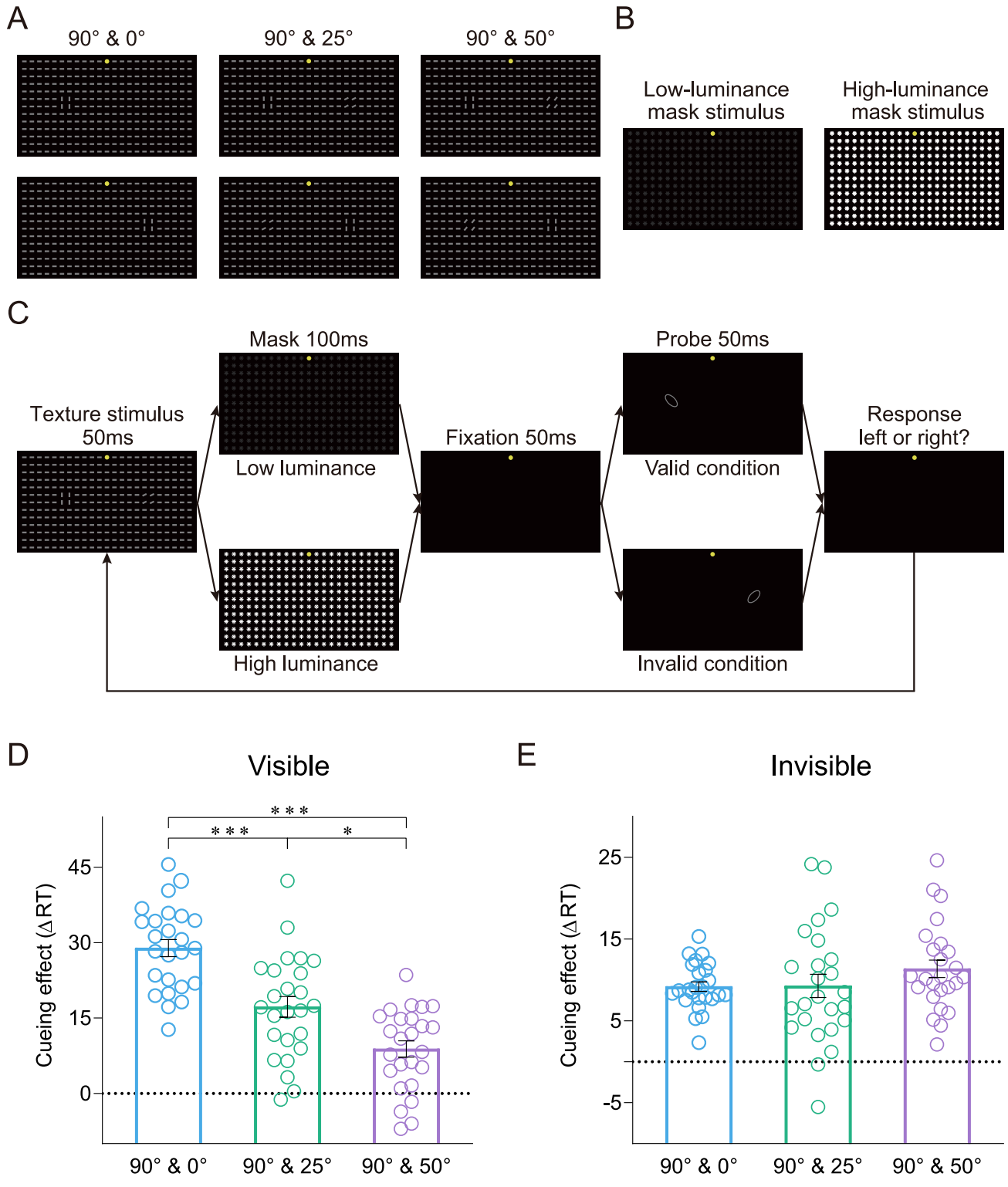

Fig. 1. Stimuli, psychophysical protocol and data. ***A*** Three types of low-luminance texture stimuli (90° & 0°, 90° & 25°, and 90° & 50°) presented in the lower visual field (top: 90°-foreground in the left visual field, bottom: 90°-foreground in the right visual field, and the yellow dots indicate the fixation point). Each texture stimulus contained a pair of salient foregrounds: a high salience target and a distractor of varied, lesser salience. ***B*** Low- (left) and high-luminance (right) mask stimuli used in the Visible and Invisible conditions, respectively. ***C*** Psychophysical protocol. A texture stimulus was presented for 50 ms, followed by a 100-ms mask and another 50-ms fixation interval. Then an ellipse probe was presented at randomly either the target location (e.g., the 90°-foreground location, valid cue condition) or its contralateral counterpart (the distractor location, i.e., invalid cue condition) with equal probability. The ellipse probe was orientated at 45° or 135° away from the vertical. Subjects were asked to press one of two buttons as rapidly and correctly as possible to indicate the orientation of the ellipse probe (45° or 135°). The cueing effect of the 90°-foreground (i.e., the target) for 90° & 0°, 90° & 25°, and 90° & 50° texture stimuli in Visible (***D***) and Invisible (***E***) conditions. Each cueing effect was quantified as the difference between the reaction time of the probe task performance in the invalid cue condition and that in the valid cue condition. Error bars denote 1 SEM calculated across subjects and colored dots denote the data from each subject.

empirical support since most previous studies used the stimulus with a single salient region only (Burrows and Moore, 2009; Buschman and Miller, 2007; Chen et al., 2016; Geng and Mangun, 2009; Katsuki and Constantinidis, 2012; Serences and Yantis, 2007, but see Bogler et al., 2011; Knudsen, 2011, 2018; White et al., 2017a).

In addition, there has been a longstanding debate about the neural loci of the saliency map. Evidence from numerous neurophysiological and imaging studies have shown that the superior colliculus (SC, Fecteau and Munoz, 2006; Kustov and Robinson, 1996; White et al., 2017a; 2017b), pulvinar (Shipp, 2004), substantia nigra (Basso and Wurtz 2002), parietal cortex (Bisley and Goldberg, 2010; Bogler et al., 2011; Buschman and Miller, 2007; Geng and Mangun, 2009; Gottlieb et al., 1998; Serences et al., 2005), V4 (Burrows and Moore, 2009; Mazer and Gallant, 2003), ventral attention network (Asplund et al., 2010; Corbetta and Shulman, 2002), frontal eye fields (Bogler et al., 2011; Serences and Yantis, 2007; Thompson and Bichot, 2005), and the dorsolateral prefrontal cortex (Katsuki and Constantinidis, 2012; Squire et al., 2013) could realize the saliency map. Most of these studies were consistent with the dominant view by Itti and Koch (2001), which proposes that saliency results from pooling different visual features, being independent of whether the feature distinction making a location salient is in color, orientation, or other features (Koch and Ullman, 1985; Wolfe, 1994). Accordingly, higher cortical areas, particularly the parietal and frontal cortex, whose neurons are less selective to specific visual features, are more likely to be possible candidates that realize the saliency map. By contrast, Li (1999, 2002) proposed that primary visual cortex (V1) creates the saliency map via intra-cortical interactions that are manifest in contextual influences (Allman et al., 1985). The saliency of a location is monotonically related to the highest neural response among all the V1 cells that cover that location with their spatial receptive fields (relative to the V1 responses to the other locations), regardless of the preferred feature of the most responsive neuron. This theory has also been supported by several psychophysical (Koene and Zhaoping, 2007; Zhaoping, 2008; Zhaoping and May 2007; Zhaoping and Zhe, 2015), neurophysiological (Kastner et al., 1997; Nothdurft et al., 1999; Yan et al., 2018), and brain imaging (Chen et al., 2016; Zhang et al., 2012) studies.

An important reason of this controversy is that most of the observed neural substrates for the saliency map are also involved in top-down attentional selection, so that the saliency map is easily and inadvertently contaminated with the top-down signals, such as feature perception, object recognition, and subjects' intentions (Zhang et al., 2012). Here, to investigate how the bottom-up attraction will be distributed among multiple salient regions, on the one hand, it is thus important to probe bottom-up attractions free from top-down influences. One way to obviate this is to use the backward masking paradigm in which the salient stimuli are presented so briefly and followed by a high luminance mask that they are invisible to subjects. On the other hand, the salient stimulus, visible versus invisible, offers a unique opportunity to reveal how its saliency map interacts with awareness.

As such stimuli, we used both visible (Experiment 1) and invisible (Experiment 2) textures made from bars (Fig. 1 **A**). Each texture stimulus contained two foreground regions (whose bars were oriented differently from the bars in the otherwise uniform background): a high salience target and a distractor of varied, lesser salience. Interference, or not, by the distractor with the effective salience of the target served to index a graded or non-graded nature of salience mapping, respectively. To examine this interference, both the Posner cueing effect and fMRI blood oxygen level-dependent (BOLD) signals in retinotopically organized areas evoked by the target were measured. We also performed a whole-brain group analysis to identify potential cortical or subcortical area(s) that showed a similar interference as those retinotopically organized areas, as well as interregional correlation and intrinsic connectivity analyses to examine the neural loci of saliency map with and without awareness.

2. Materials and methods

2.1. Subjects

A total of 25 human subjects (4 male, 19–26 years old) were involved in the study. All of them participated in the psychophysical experiment. Twenty-one of them participated in the fMRI experiment. One subject in the fMRI experiment was excluded because of large head motion (> 3 mm). The sample size was based on those used in previous studies on saliency map in our lab (Huang et al., 2020; Wang et al., 2021; Zhang et al., 2012) and a priori power calculation using G*Power program (Faul et al., 2009). The power analysis indicated that a sample of 25 and 20 in our psychophysical and fMRI experiments, respectively, would be sufficient to detect a medium-size effect ($f = 0.25$) in a within-subjects analysis of ANOVA with the power of 0.8. They were naïve to the purpose of the study. They were right-handed, reported normal or corrected-to-normal vision, and had no known neurological or visual disorders. They gave written, informed consent, and our procedures and protocols were approved by the human subjects review committee of School of Psychology at South China Normal University.

2.2. Stimuli

Each texture stimulus (Fig. 1 **A**) had a regular Manhattan grid of 13 × 25 low luminance bars (1.38 cd/m^2), presented in the lower visual field on a dark screen (0.007 cd/m^2). Each bar was a rectangle of 0.0625° × 0.5° in visual angle. The center-to-center distance between the bars was 0.75°. All bars were identically oriented except for a foreground region of 2 × 2 bars with another orientation. The orientation of the background bars was randomly chosen from 0° to 180° on each trial. There were four different foregrounds with 0°, 25°, 50°, and 90° orientation contrasts between the foreground bars and the background bars (Fig. 1 **A**). In each texture stimulus, a pair of foregrounds was centered in the lower left and lower right quadrants at 5.83° eccentricity. There were five possible types of texture stimuli: 90° & 0°, 90° & 25°, 90° & 50°, 25° & 0°, and 50° & 0° Each type of texture stimuli contained two possible pairs of foregrounds; one for the high salient foreground was in the left visual field (i.e., 90° + 0°, 90° + 25°, 90° + 50°, 25° + 0°, and 50° + 0°, respectively) and the other one for the high salient foreground was in the right visual field (i.e., 0° + 90°, 25° + 90°, 50° + 90°, 0° + 25°, and 0° + 50°, respectively). Low- (0.016 cd/m^2) and high- (78.675 cd/m^2) luminance masks, which had the same grid as the texture stimuli, rendered the whole texture stimulus visible (Experiment 1) and invisible (Experiment 2, confirmed by a 2AFC test, i.e., Experiment 3) to subjects, respectively. Each element of the mask contained 12 intersecting bars oriented from 0° to 165° at every 15° interval. The bars in the mask had the same size and shape as those in the texture stimuli.

2.3. Psychophysical experiments

Visual stimuli were displayed on an IIYAMA color graphic monitor (model: HM204DT; refresh rate: 60 Hz; resolution: 1280 × 1024; size: 22 inches) at a viewing distance of 57 cm. Subjects' head position was stabilized using a chin rest. A yellow fixation point was always present at the center of the monitor. Psychophysical experiments consisted of three experiments. Experiments 1 (Visible) and 2 (Invisible) investigated whether the mapping of salience depended on the visibility of texture stimuli. Experiment 3 checked the effectiveness of the awareness manipulation in Experiments 1 and 2, and was always preceded them.

2.3.1. Experiments 1 and 2

Subjects participated in Experiments 1 and 2 on two different days, and the order of the two experiments was counterbalanced across subjects. In both Experiments 1 and 2, we used a modified version of the Posner paradigm (Posner et al., 1980) to measure the spatial cueing effect induced by the high salient foreground (the target, Fig. 1 **C**).

Namely, the 90°-foreground served as the target for the 90° & 0°, 90° & 25°, and 90° & 50° texture stimuli, the other low salient foreground (i.e., 25° and 50°) that presented at its contralateral counterpart, served as the distractor. Note that there was no distractor for the 90° & 0° texture stimuli since the 0°-foreground region would always contain background bars. Similarly, the 25°- and 50°-foregrounds were the target for the 25° & 0° and 50° & 0° texture stimuli without the distractor, respectively.

Each trial began with the fixation. A texture stimulus was presented for 50 ms, followed by a 100 ms mask (low- and high-luminance in Experiments 1 and 2, respectively) and another 50-ms fixation interval. For each type of texture stimuli, the two possible pairs of foregrounds were presented randomly and equiprobably. The defined target in each type of texture stimulus served as a cue to attract spatial attention. Then an ellipse probe was presented for 50-ms at randomly either the target location (e.g., the 90°-foreground location, valid cue condition) or its contralateral counterpart (the distractor location, i.e., invalid cue condition) with equal probability (Fig. 1 **C**). The ellipse probe was orientated at 45° (left) or 135° (right) away from the vertical. Subjects were asked to press one of two buttons as rapidly and correctly as possible to indicate the orientation of the ellipse probe (45° or 135°). For each condition, a leftward response to a 45° ellipse was (arbitrarily) considered to be a hit, a leftward response to a 135° ellipse was considered to be a false alarm, and a rightward response to a 45° ellipse was considered to be a miss. Each experiment consisted of 12 blocks, 6 for the 25°-distractor and 6 for the 50°-distractor. The order of these two different distractors was counterbalanced across subjects. Each block had 96 trials, from randomly interleaving 32 trials from three conditions (25°-distractor: 90° & 0°, 90° & 25°, and 25° & 0°; 50°-distractor: 90° & 0°, 90° & 50°, and 50° & 0°). The cueing effect for each texture stimulus was quantified as the difference between the reaction time of the probe task performance in the invalid cue condition and that in the valid cue condition.

2.3.2. Experiment 3

In Experiment 3, all subjects underwent a 2AFC task to determine whether the masked foreground was visible or invisible in a criterion-free way. The stimuli and procedure in this 2AFC experiment were the same as those in Experiments 1 and 2, except that no probe was presented (Fig. S1**A**). After the presentation of a masked texture stimulus, subjects were asked to make a forced choice response regarding which side (lower left or lower right) from the fixation they thought the high salient foreground (i.e., the defined target) appeared. Their performances were significantly higher or not statistically different from chance for all possible texture stimuli, providing an objective confirmation that the foreground was indeed visible or invisible to subjects, respectively.

2.4. fMRI experiments

Using a block design, the experiment consisted of 10 functional runs. Each run consisted of 14 stimulus blocks of 10 s, interleaved with 14 blank intervals of 12 s. There were 14 different stimulus blocks: 3 (texture stimulus: 90° & 0°, 90° & 25°, and 90° & 50°) × 2 (visual field: left/right) × 2 (awareness: visible/invisible), and 2 mask-only blocks: low- and high-luminance masks. Each stimulus block was randomly presented once in each run, and consisted of 5 trials. On each trial in the texture stimulus and mask-only blocks, a texture stimulus or a fixation was presented for 50 ms, respectively, followed by a 100-ms mask (low- and high-luminance for Visible and Invisible conditions, respectively) and 1850 ms fixation (Fig. 2 **C**). In the Invisible condition, on each trial during both the texture stimulus and mask-only blocks, subjects were asked to press one of two buttons to indicate the location of the 90°-foreground (the target), which was left of fixation in one half of blocks and right of fixation in the other half at random (i.e., the 2AFC task). Note that although the salient foreground was not presented in the mask-only blocks, subjects also indicated the location of the target since in the invisible condition; they were unaware whether the target was present or absent. In the Visible condition, on each trial during the texture stimulus block, subjects needed to perform the same 2AFC task of the target; whereas during the mask-only block, subjects were asked to press one of two buttons randomly since in the visible condition, they were easy to perceive the absence of targets.

Retinotopic visual areas (SC and V1–V4) were defined by a standard phase-encoded method developed by Sereno et al. (1995) and Engel et al. (1997), in which subjects viewed rotating wedge and expanding ring stimuli that created traveling waves of neural activity in visual cortex. An independent block-design scan was used to localize the ROIs in SC and V1–V4 corresponding to the pair of foreground regions. The scan consisted of 12 12-s stimulus blocks, interleaved with 12 12-s blank intervals. In a stimulus block, subjects passively viewed images of colorful natural scenes, which had the same size as the foreground regions in texture stimuli and were presented at locations of the pair of foreground regions. Images appeared at a rate of 8 Hz.

2.4.1. MRI data acquisition

MRI data were collected using a 3T Siemens Trio scanner with a 32-channel phase-array coil at the Center for MRI Research at South China Normal University. In the scanner, the stimuli were back-projected via a video projector (refresh rate: 60 Hz; spatial resolution: 1024 × 768) onto a translucent screen placed inside the scanner bore. Subjects viewed the stimuli through a mirror located above their eyes. The viewing distance was 90 cm. Blood oxygen level-dependent (BOLD) signals were measured with an echo-planar imaging sequence (TE: 30 ms; TR: 2000 ms; FOV: 192 × 192 mm^2; matrix: 64 × 64; flip angle: 90; slice thickness: 3 mm; gap: 0 mm; number of slices: 32, slice orientation: axial). A high-resolution 3D structural data set (3D MPRAGE; 1 × 1 × 1 mm^3 resolution; TR: 2600 ms; TE: 3.02 ms; FOV: 256 × 256 mm^2; flip angle: 8; number of slices: 176; slice orientation: sagittal) was collected in the same session before the functional scans. Subjects underwent three sessions, one for the retinotopic mapping and ROI localization, and the other two for the main experiment.

2.4.2. MRI data analysis

Note that, the MRI data analysis, whole-brain group analysis, and DCM of this study closely followed those used by our previous studies (Zhang et al., 2014, 2016, 2018) and therefore, for consistency, we largely reproduce that description here, noting differences as necessary. The anatomical volume for each subject in the retinotopic mapping session was transformed into a brain space that was common for all subjects (Talairach and Tournoux, 1988) and then inflated using BrainVoyager QX. Functional volumes in both sessions for each subject were preprocessed, including 3D motion correction, linear trend removal, and high-pass (0.015 Hz, Smith et al., 1999) filtering using BrainVoyager QX. Head motion within any fMRI session was < 3 mm for all subjects. The images were then aligned to the anatomical volume in the retinotopic mapping session and transformed into Talairach space (Talairach and Tournoux, 1988). The first 8-s of BOLD signals were discarded to minimize transient magnetic saturation effects.

A general linear model (GLM) procedure was used for the ROI analysis. The ROIs in SC and V1–V4 were defined as areas that responded more strongly to the natural scene images than blank screen ($P < 0.05$, corrected by FDR correction, Genovese et al., 2002). The block-design BOLD signals were extracted from these ROIs and then averaged according to each type of trials. For each stimulus block, the 2-s preceding the block served as a baseline, and the mean BOLD signal from 5-s to 10-s after stimulus onset was used as a measure of the response amplitude.

2.4.3. Whole-brain group analysis

In the whole-brain group analysis, a fixed-effects general linear model (FFX-GLM) was performed for each subject on the spatially non-smoothed functional data in Talairach space. The 1st-level regressors

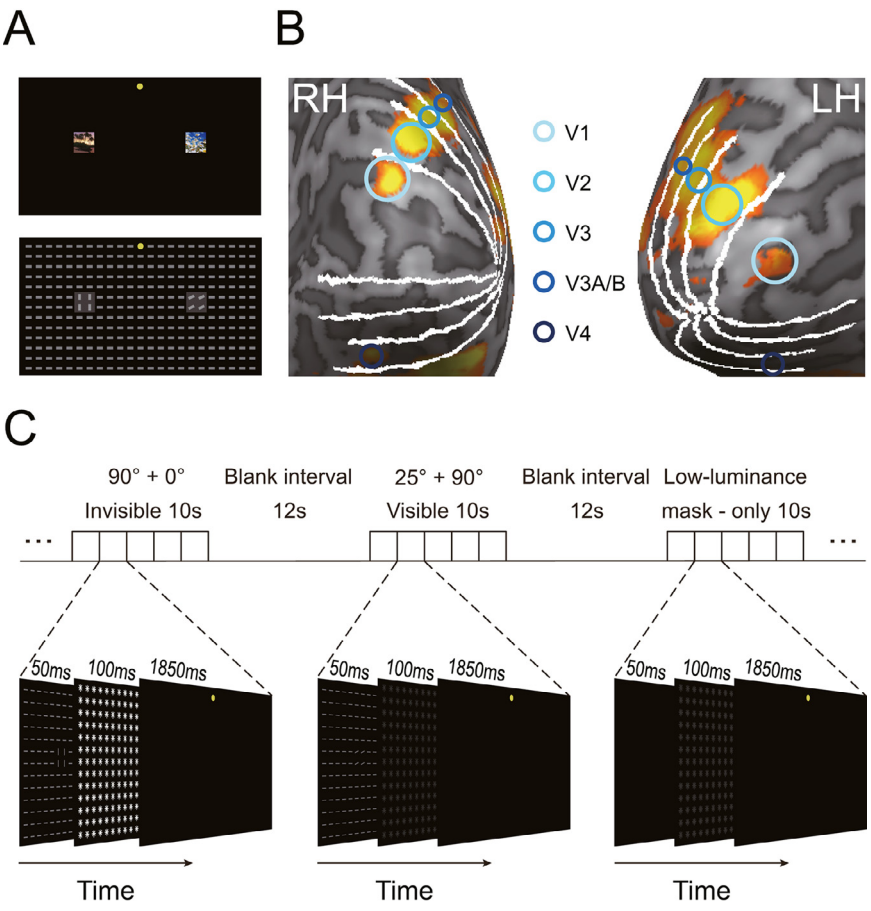

Fig. 2. fMRI stimuli and protocol. ***A*** ROI definition. The natural scenes in the top panel were used to define ROIs corresponding to the pair of foregrounds in texture stimuli. The transparent squares show the size and location of the natural scenes relative to texture stimuli. ***B*** ROIs on an inflated cortical surface of a representative subject. The ROIs in V1–V4 were defined as the cortical regions responding to the foreground. The boundaries among V1–V4, defined by retinotopic mapping, are indicated by the white lines. ***C*** Block design fMRI procedure. On each trial in the texture stimulus and mask-only blocks, a texture stimulus and the fixation were presented for 50 ms, respectively, followed by a 100-ms mask (low- and high-luminance for Visible and Invisible conditions, respectively) and 1,850-ms fixation interval. In the Invisible condition, on each trial during both the texture stimulus and mask-only blocks, subjects were asked to press one of two buttons to indicate the location of the 90°-foreground, which was left of fixation in one half of blocks and right of fixation in the other half at random (i.e., the 2AFC task). In the Visible condition, on each trial during the figure block, subjects needed to perform the same 2AFC task of the 90°-foreground; whereas during the mask-only block, subjects were asked to press one of two buttons randomly.

were created by convolving the onset of each stimuli block with the default BrainVoyager QX's two-gamma hemodynamic response function. Six additional parameters resulting from 3D motion correction (x, y, z rotation and translation) were included in the model. In both Visible and Invisible conditions, for each subject, we first calculated fixed effects analyses for each texture stimulus separately (90° & 0° texture stimulus block vs. mask-only block, 90° & 25° texture stimulus block vs. mask-only block, and 90° & 50° texture stimulus block vs. mask-only block) (Fig. 2 **C**). Next, these FFX-GLM estimates were entered into a second-level group analysis (random-effects) of variance (ANOVA) and the statistical F-map was generated expressing the main effect of the three types of texture stimuli (90° & 0°, 90° & 25°, and 90° & 50°). Statistical maps were thresholded at $P < 0.05$ and corrected by FDR correction (Genovese et al., 2002).

2.4.4. Effective connectivity analysis

In the Visible condition, our results showed that the BOLD signal in both V1 and frontoparietal cortical areas could be modulated by the three types of texture stimuli (90° & 0°, 90° & 25°, and 90° & 50°), indicating a graded manner of saliency map. To further examine which area is a potential source of this graded manner of saliency map, we applied DCM analysis (Friston et al., 2003) in SPM12 to examine interregional intrinsic connectivity change among these three types of texture stimuli. For each subject and each hemisphere, using BrainVoyager QX, V1 voxels were identified as those activated by the foreground at a significance level of $p < 0.01$. All of the pIPS, anterior intraparietal sulcus (aIPS), and frontal eye field (FEF) voxels were identified as those activated by the stimulus block at a significance level of $p < 0.01$. The mean Talairach coordinates of V1, pIPS, aIPS, and FEF, and their standards errors across subjects were [-10 ± 0.99, -95 ± 0.96, 0 ± 1.74], [-25 ± 1.04, -65 ± 1.04, 40 ± 0.98], [-39 ± 1.70, -44 ± 1.87, 38 ± 4.71], and [-46 ± 0.82, 2 ± 0.94, 31 ± 0.90] for the left hemisphere, and [7 ± 0.98, -93 ± 0.63, 3 ± 1.57], [25 ± 0.81, -64 ± 1.46, 40 ± 1.79], [31 ± 1.16, -45 ± 1.06, 39 ± 0.90], and [44 ± 0.57, 3 ± 0.57, 24 ± 1.01], for the right hemisphere, respectively. For each subject and each hemisphere, these Talairach coordinates were converted to Montreal Neurological Institute (MNI) coordinates using the tal2mni conversion utility (http://imaging.mrc-cbu.cam.ac.uk/downloads/MNI2tal/tal2mni.m). In SPM, for each of these areas, we extracted voxels within a 6-mm sphere centered on the most significant voxel and used their time series for the DCM analysis. The estimated DCM parameters were later averaged across left and right V1 using the Bayesian model averaging method (Penny et al., 2004).

Given the extrinsic visual input into both V1T (i.e., the ROI in V1 that was evoked by the target: 90°-foreground) and V1D (i.e., the ROI in V1 that was evoked by the distractor), bidirectional connections were hypothesized to exist among the pIPS, aIPS, FEF, V1T, and V1D (Fig. 5), and these intrinsic connections could be modulated by the three types of texture stimuli (90° & 0°, 90° & 25°, and 90° & 50°). In addition, given the well-known hemispheric asymmetries in the frontoparietal attention networks (Bartolomeo and Malkinson, 2019), we constructed two families of models with the left hemisphere ROIs only (Fig. 5 **A**) and the right hemisphere ROIs only (Fig. 5 **E**) in the pIPS, aIPS, and FEF.

3. Results

3.1. Psychophysical experiments

In psychophysical experiments, there were four possible foregrounds with 0°, 25°, 50°, and 90° orientation contrasts between the foreground bars and the background bars. In each texture stimulus, a pair of foregrounds was centered in the lower left and lower right quadrants at 5.83°

eccentricity (Fig. 1 **A**). There were five possible pairs of foregrounds: 90° & 0°, 90° & 25°, 90° & 50°, 25° & 0°, and 50° & 0° Low- and high-luminance masks (Fig. 1 **B**) rendered the whole texture stimulus visible (Experiment 1) or invisible (Experiment 2) to subjects, respectively, confirmed by two-alternative forced choice (2AFC, Experiment 3, Fig. S1). In both Experiments 1 and 2, we used a modified version of the Posner paradigm (Posner et al., 1980) to measure the cueing effect induced by the high salient foreground (the target) in each pair of foregrounds, as shown in Fig. 1 **C**. Namely, the 90°-foreground served as the target for the 90° & 0°, 90° & 25°, and 90° & 50° texture stimuli, the other low salient foreground (i.e., 25° and 50°) that presented at its contralateral counterpart, served as the distractor. Note that there was no distractor for the 90° & 0° texture stimuli since the 0°-foreground region would always contain background bars. Similarly, the 25°- and 50°-foregrounds were the target for the 25° & 0° and 50° & 0° texture stimuli without the distractor, respectively. In our modified Posner paradigm, a valid cue condition was defined as a match of quadrant between the defined target and the ellipse probe (for example, both the 90°-foreground and probe were presented in the lower left quadrant); an invalid cue condition was defined as a mismatch (Fig. 1 **C**). Subjects were asked to press one of two buttons as rapidly and correctly as possible to indicate the orientation of the ellipse probe (45° or 135°). For each condition, a leftward response to a 45° ellipse was (arbitrarily) considered to be a hit, a leftward response to a 135° ellipse was considered to be a false alarm, and a rightward response to a 45° ellipse was considered to be a miss. There was no significant difference in the false alarm rate, miss rate, or removal rate (i.e., correct reaction times shorter than 200 ms and beyond three standard deviations from the mean reaction time in each condition were removed) across conditions (all $P > 0.05$, partial eta squared, $\eta_p{}^2 < 0.151$, Fig. S2). The cueing effect for each texture stimulus was quantified as the difference between the reaction time of the probe task performance in the invalid cue condition and that in the valid cue condition.

We found that salience mapping, as measured by the cueing effect, was indeed altered by the tested combinations of distractor salience and stimulus visibility (Fig. 1 **A**). In both Visible and Invisible conditions, using the one-sample *t*-test, we found that the 90°-foreground of all the texture stimuli exhibited a positive cueing effect (Visible condition: 90° & 0°: $t(24) = 17.090$, $P < 0.001$, Cohen's $d = 4.834$; 90° & 25°: $t(24) = 8.338$, $P < 0.001$, Cohen's $d = 2.358$; 90° & 50°: $t(24) = 5.448$, $P < 0.001$, Cohen's $d = 1.541$; Invisible condition: 90° & 0°: $t(24) = 16.038$, $P < 0.001$, Cohen's $d = 4.536$; 90° & 25°: $t(24) = 6.517$, $P < 0.001$, Cohen's $d = 1.843$; 90° & 50°: $t(24) = 10.622$, $P < 0.001$, Cohen's $d = 3.004$), indicating that the attention of the subject was attracted more to the 90°-foreground location, allowing them to perform more proficiently in the valid than the invalid cue condition of the probe task (Fig. 1 **D** and 1 **E**). Then, these cueing effects were submitted to a repeated measures ANOVA with awareness (Visible and Invisible) and texture stimulus (90° & 0°, 90° & 25°, and 90° & 50°) as within-subjects factors. The main effect of awareness ($F(1,24) = 41.734$, $P < 0.001$, $\eta_p{}^2 = 0.635$), the main effect of texture stimulus ($F(2,48) = 24.258$, $P < 0.001$, $\eta_p{}^2 = 0.503$), and the interaction between these two factors ($F(2,48) = 37.159$, $P < 0.001$, $\eta_p{}^2 = 0.608$) were all significant. Thus, these data were submitted to a further simple effect analysis. Post hoc paired *t* tests showed that the cueing effect in the Visible condition was significantly greater than that in the Invisible condition for both 90° & 0° ($t(24) = 12.145$, $P < 0.001$, Cohen's $d = 3.123$) and 90° & 25° ($t(24) = 3.396$, $P = 0.002$, Cohen's $d = 0.896$) texture stimuli, but not for the 90° & 50° texture stimulus ($t(24) = -1.334$, $P = 0.195$, Cohen's $d = -0.369$). For the Invisible condition, the main effect of the texture stimulus was not significant ($F(2,48) = 1.252$, $P = 0.287$, $\eta_p{}^2 = 0.050$, Fig. 1 **E**), indicating that the cueing effect of 90°-foreground could not be interfered by the distractor (i.e., 25°- and 50°-foregrounds, Fig. 1 **A**). However, the null interference here could be explained by the low saliency distractor lacking the capacity to automatically attract subjects' attention to its location. Accordingly, to address this issue, we analyzed the cueing effect of the 25°- and 50°-foreground within the 25° & 0° and 50° & 0° texture stimuli, respectively. The results argue against this explanation by showing that the cueing effect of 25°- and 50°-foregrounds were both significantly above zero (25°-foreground: $t(24) = 3.003$, $P = 0.006$, Cohen's $d = 0.849$; 50°-foreground: $t(24) = 4.185$, $P < 0.001$, Cohen's $d = 1.184$, Fig. S3**A**). These results thus indicated that bottom-up salience was mapped as a non-graded manner without awareness. For the Visible condition, the main effect of the texture stimulus was significant ($F(2,24) = 47.262$, $P < 0.001$, $\eta_p{}^2 = 0.663$); post hoc paired *t* tests revealed that the cueing effect of 90° & 25° texture stimulus was significantly lower than that of the 90° & 0° texture stimulus ($t(24) = 6.113$, $P < 0.001$, Cohen's $d = -1.240$) but significantly higher than that of the 90° & 50° texture stimulus ($t(24) = 3.268$, $P = 0.01$, Cohen's $d = 0.903$). These results demonstrated that the cueing effect of 90°-foreground can be interfered by the distractor and, notably, the degree of this interference increased with the orientation contrast of the distractor (i.e., the graded distribution, Fig. 1 **D**).

Notably, our inference of the graded or non-graded mapping of salience with or without awareness, respectively, depends upon the results of the cueing effect (calculated as the difference in reaction time between the invalid and valid cue conditions) among 90° & 0°, 90° & 25°, and 90° & 50° texture stimuli. To further clarify this awareness-dependent salience mapping, for each condition, we analyzed the reaction time in the valid and invalid trials separately and proposed four potential mechanisms of the salience mapping (Fig. S4**A**): First, the graded mechanism: exogenous attention is proportionally allocated to the high- and low-salient foregrounds based on their levels of saliency. This mechanism is consistent with many traditional capacity limitation models of attention (Cheal et al., 1994; Eriksen and Yeh, 1985; Posner, 1980), the valid cue would produce improved performance at cued location (benefit), which would be accompanied by corresponding impairments in the invalid cue locations (cost). Such a trade-off between the benefit and cost could predict the increased and decreased reaction time with the saliency of distractor in the valid and invalid cue conditions, respectively. Second, the priming mechanism (i.e., the cost-without-benefit mechanism): the attention allocation is dependent upon the localised priming of the forward processing and subjects' reaction times were changed in the invalid cue (cost) condition only, i.e., showing the decreased reaction time with the saliency of distractor. Remarkably, this priming or cost-without-benefit effect has been demonstrated by a number of previous studies (Folk and Hoyer, 1992; Hawkins et al., 1990; Shiu and Pashler, 1994, 1995, see Mulckhuyse and Theeuwes, 2010 for a review). Third, the relative-WTA mechanism (i.e., the benefit-without-cost mechanism): the attention allocation is dependent on the confidence associated with the identification of the locus of highest salience and subjects' reaction times were changed in the valid cue (benefit) condition only (Friesen and Kingstone, 1998; Posner and Snyder, 1975), i.e., showing the increased reaction time with the saliency of distractor. Finally, the WTA-like mechanism: exogenous attention is completely captured by the most salient foreground (90°-foreground) within the texture stimulus, as proposed by V1 saliency hypothesis (Li, 1999, 2002), whereby saliency of a visual location is determined by its highest evoked V1 response relative to those evoked by other locations. This mechanism would predict non-significant differences in the reaction time among the three types of texture stimuli in both valid and invalid conditions. Accordingly, in both Visible and Invisible conditions, the reaction times were submitted to a repeated measures ANOVA with cue validity (valid and invalid) and texture stimulus (90° & 0°, 90° & 25°, and 90° & 50°) as within-subjects factors and the results indicated that bottom-up salience was mapped as the graded (Fig. S4**B**) and WTA-like (Fig. S4**C**) mechanisms with and without awareness, respectively.

Although our results suggested a WTA-like mechanism for the non-graded mapping of salience without awareness, we cannot deny a potential contribution from the priming mechanism in this processing. First, the distractor in our study was always presented in an opposite visual field to the target, producing the bilateral competition. A number

of previous studies have demonstrated that the competition between stimulus items is more marked within the same hemifield than that across hemifields (Cavanagh and Alvarez, 2005; Franconeri et al., 2013; Shipp, 2011), and more importantly, V1 is well known to lack the interhemispheric connectivity required to achieve this bilateral competition (Wandell et al., 2007). Therefore, the priming mechanism that does not invoke the bilateral competition between the target and distractor may alternatively underlie our non-graded mapping of salience. Second, it has been proved and widely accepted that rendering a stimulus invisible could maximally (although not completely, Boly et al., 2017) reduce various top-down contaminations (Zhang et al., 2012) and in this case, the residual and heavily diminished cueing effect must be more dependent upon the localised priming of the forward processing pathway, sensitizing it to subsequent probe at the same location. Finally, although the priming mechanism predicted the graded rather than the non-graded mapping of salience (Fig. S4**A**), if the low-salient distractor lacks the capacity to attract subjects' attention to its location, then the priming mechanism is capable of producing the same results with the WTA-like mechanism. Given a little weak (although significant, Fig. S3) effect of the distractor in our study, we thus cannot deny a potential contribution from the priming mechanism in our non-graded mapping of salience. Further work is needed to use the more salient distractor (e.g., 90° & 70°) to parse the relative contributions of the WTA-like and priming mechanisms to the non-graded mapping of salience without awareness.

In addition, it could be argued that the awareness-dependent mapping of saliency could depend on the strength of cueing effect. To examine this issue, in both Visible and Invisible conditions, the subjects who scored in the top and bottom 48% (i.e., $n = 12$) of the sample's cueing effect distribution were assigned to High- and Low-cueing effect groups, respectively. Their cueing effects were submitted to a mixed ANOVA with group (High and Low) as the between-subjects factor and texture stimulus (90° & 0°, 90° & 25°, and 90° & 50°) as the within-subjects factor. In both Visible and Invisible conditions, the results showed the same qualitative conclusion and insignificant interactions between the two factors (Fig. S5), further confirming that, despite the strength of cueing effect, the bottom-up salience was mapped as graded or non-graded manners with or without awareness, respectively.

3.2. fMRI experiments

Using a block design, the functional magnetic resonance imaging (fMRI) experiment consisted of ten functional runs. Each run consisted of 14 stimulus blocks of 10 s, interleaved with 14 blank intervals of 12 s. There were 14 different stimulus blocks, including 12 different texture stimulus blocks: 3 (texture stimulus: 90° & 0°, 90° & 25°, and 90° & 50°) × 2 (visual field: left and right) × 2 (awareness: Visible and Invisible), and 2 mask-only blocks: low- and high-luminance masks (Fig. 2 **C**). Each stimulus block was randomly presented once in each run, and consisted of the same 5 trials. On each trial in the texture stimulus and mask-only blocks, a texture stimulus and the fixation were presented for 50 ms, respectively, followed by a 100-ms mask (low- and high-luminance for Visible and Invisible conditions, respectively) and 1850-ms fixation. In the Invisible condition, on each trial during both the texture stimulus and mask-only blocks, subjects were asked to press one of two buttons to indicate the location of the 90°-foreground, which was left of fixation in one half of blocks and right of fixation in the other half at random (i.e., the 2AFC task). In the Visible condition, on each trial during the figure block, subjects needed to perform the same 2AFC task of the figure; whereas during the mask-only block, subjects were asked to press one of two buttons randomly (Fig. 2 **C**). Behavioral data showed that, our awareness manipulation was effective for both Visible and Invisible conditions (Fig. S1**C**).

3.2.1. Region of interest analysis

Regions of interest (ROIs) in SC and V1–V4 were defined as the cortical regions responding significantly to each pair of foregrounds (Fig. 2 **B**). For each foreground, BOLD signals were extracted from the contralateral ROIs in these retinotopically organized areas and then averaged according to the stimulus (90° & 0°, 90° & 25°, and 90° & 50° texture stimuli, and the mask-only) and awareness (Visible and Invisible). For each stimulus block, the 2-s preceding the block served as a baseline, and the mean BOLD signal from 5-s to 10-s after stimulus onset was used as a measure of the response amplitude. To isolate the texture stimulus signal, the BOLD amplitudes of the low- and high-luminance mask-only blocks were subtracted from those of the Visible and Invisible texture stimulus blocks, respectively, The BOLD signal difference of the 90°-foreground for each condition is shown in Fig. 3 , and all of them were significantly above zero (all $t(19) > 2.130$, $P < 0.046$, Cohen's $d > 0.674$). Sequentially, in both Visible and Invisible condition, these BOLD signal differences of the 90°-foreground were submitted to a repeated-measures ANOVA with texture stimulus (90° & 0°, 90° & 25°, and 90° & 50°) and cortical area (SC and V1–V4) as within-subjects factors.

In the Invisible condition, the main effect of the texture stimulus ($F(2,38) = 1.100$, $P = 0.342$, $\eta_p^2 = 0.055$), the main effect of cortical area ($F(5,95) = 0.528$, $P = 0.689$, $\eta_p^2 = 0.027$), and the interaction between these two factors ($F(10,190) = 0.709$, $P = 0.641$, $\eta_p^2 = 0.036$) were not significant (Fig. 3 **B**). These results indicated that the BOLD signal difference of 90°-foreground was not influenced by the distractor (i.e., 25°- and 50°-foregrounds). Similar to the psychophysical cueing effect (Fig. 1 **E**), the null interference here could also be explained by the low saliency distractor lacking the capacity to automatically attract subjects' attention to its location. Accordingly, to address this issue, we analyzed the BOLD signal difference of the ROIs evoked by the 25°- and 50°-foregrounds. The results argue against this explanation by showing that the BOLD signal difference of 25°- and 50°-foregrounds were both significantly above zero (25°-foreground, SC and V1–V4: all $t(19) > 2.194$, $P < 0.041$, Cohen's $d > 0.694$; 50°-foreground, SC and V1–V4: all: $t(19) > 2.188$, $P < 0.041$, Cohen's $d > 0.692$, Fig. S3**B**, right). In addition, to examine which cortical area's activities closely mirrored the psychophysical cueing effect, we calculated the correlation coefficients between the cueing effect and BOLD signal difference of 90°-foreground across individual subjects. The cueing effect was significantly correlated with the BOLD signal difference of the 90°-foreground in V1 (90° & 0°: $r = 0.467$, $P = 0.038$; 90° & 25°: $r = 0.633$, $P = 0.003$; 90° & 50°: $r = 0.445$, $P = 0.049$, Fig. 3 **D**, left), but not in SC or V2–V4 (90° & 0°: all $r < 0.426$, $P > 0.061$; 90° & 25°: all $r < 0.343$, $P > 0.139$; 90° & 50°: all $r < 0.409$, $P > 0.073$, Fig. 3 **D**, right). Furthermore, pooling data across all three texture stimuli, their mean cueing effect was also significantly correlated with their mean BOLD signal difference of the 90°-foreground in V1 ($r = 0.640$, $P = 0.002$, Fig. S6**A**), but not in SC or V2–V4 (all $r < 0.315$, $P > 0.176$, Fig. S6**B**). Moreover, using the Fisher's z-transformation (Fisher, 1921), all these correlation coefficient *rs* were converted into *zs*, and Z-tests revealed that the correlation coefficient in V1 was significantly larger than those in SC ($P = 0.040$) and V4 ($P = 0.023$). Together, the physiological correlates of the limited cueing effect attained under backward masking conditions appeared to be restricted to area V1.

In the Visible condition, the main effect of the texture stimulus was significant ($F(2,38) = 5.859$, $P = 0.008$, $\eta_p^2 = 0.236$), demonstrating that the BOLD signal difference of 90°-foreground decreased with the orientation contrast of the distractor, namely, increasing the interference of the distractor (the graded manner). We also found a significant main effect of cortical area ($F(5,95) = 6.366$, $P < 0.001$, $\eta_p^2 = 0.251$) and a significant interaction between texture stimulus and cortical area ($F(10, 190) = 2.357$, $P = 0.047$, $\eta_p^2 = 0.110$). Hence, the interference of the distractor decreased gradually from lower to higher cortical areas. This was confirmed in further analysis which showed that the main effect of texture stimulus was significant in V1 ($F(2,38) = 24.904$, $P < 0.001$, $\eta_p^2 = 0.567$), but not in SC or V2–V4 (all $F(2,38) < 2.658$, $P > 0.086$, $\eta_p^2 < 0.123$) (Fig. 3 **A**). These findings revealed that neural activities in V1 were parallel to the psychophysical cueing effect (Fig. 1 **D**). To further evaluate a close relationship between the V1 activities

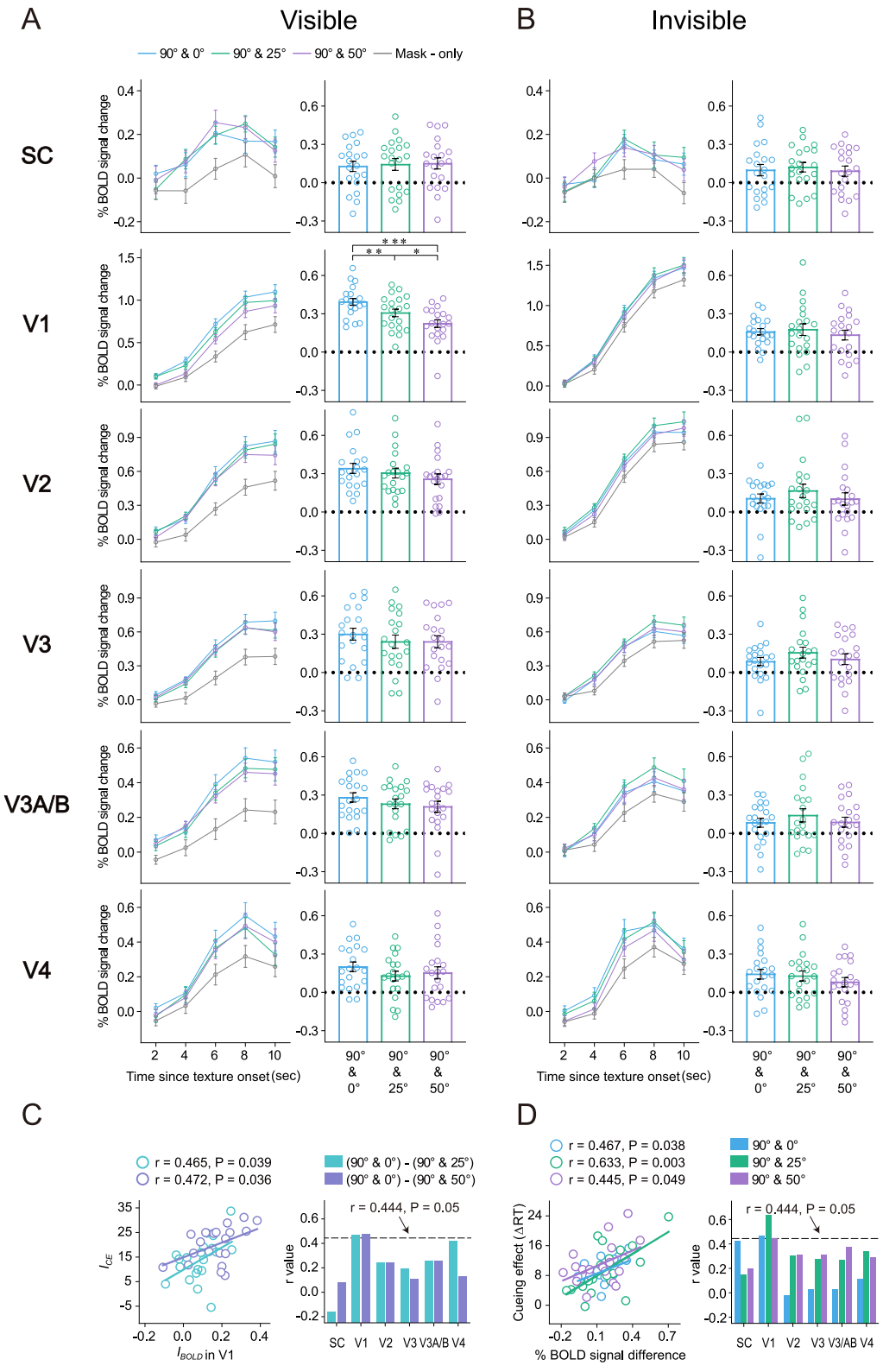

Fig. 3. fMRI results. Left: Blocked BOLD signals averaged across subjects of the ROIs in SC and V1–V4 evoked by the 90˚-foreground for three types of texture stimuli and the mask-only, during the Visible (***A***) and Invisible (***B***) conditions. Error bars denote 1 SEM calculated across subjects at each time point. Right: BOLD signal differences (i.e., the BOLD signals in the texture stimulus block - the BOLD signals in the mask-only block) of the 90˚-foreground in SC and V1–V4 during the Visible (***A***) and Invisible (***B***) conditions. Error bars denote 1 SEM calculated across subjects and colored dots denote the data from each subject. ***C*** Correlations between the I_{CE} and the I_{BOLD} in V1 (left), and correlation coefficients (r values) between the I_{CE} and the I_{BOLD} in SC and V1–V4 (right), across individual subjects during the Visible condition. ***D*** Correlations between the cueing effect and the BOLD signal difference of 90˚-foreground in V1 (left), and correlation coefficients (r values) between the cueing effect and the BOLD signal difference of 90˚-foreground in SC and V1–V4 (right), across individual subjects during the Invisible condition.

and our psychophysical cueing effect, we computed an interference of the distractor to quantify how much the cueing effect (I_{CE}) and BOLD signal (I_{BOLD}) of the 90°-foreground changed in both the 90° & 25° and 90° & 50° texture stimuli relative to that in the 90° & 0° texture stimulus. The interference was calculated as follows: $I_{CE\ (90°\ \&\ 25°)} = CE_{90°\ \&\ 0°} - CE_{90°\ \&\ 25°}$ and $I_{CE(90°\ \&\ 50°)} = CE_{90°\ \&\ 0°} - CE_{90°\ \&\ 50°}$, where $CE_{90°\ \&\ 0°}$, $CE_{90°\ \&\ 25°}$, and $CE_{90°\ \&\ 50°}$ are the cueing effect of 90°-foreground for the 90° & 0°, 90° & 25°, and 90° & 50° texture stimuli, respectively. Similarly, for each ROI, $I_{BOLD\ (90°\ \&\ 25°)} = BOLD_{90°\ \&\ 0°} - BOLD_{90°\ \&\ 25°}$ and $I_{BOLD(90°\ \&\ 50°)} = BOLD_{90°\ \&\ 0°} - BOLD_{90°\ \&\ 50°}$, where $BOLD_{90°\ \&\ 0°}$, $BOLD_{90°\ \&\ 25°}$, and $BOLD_{90°\ \&\ 50°}$ are the BOLD signal difference of 90°-foreground for the 90° & 0°, 90° & 25°, and 90° & 50° texture stimuli, respectively. Subsequently, for each ROI, we calculated the correlation coefficients between the I_{CE} and the I_{BOLD} across individual subjects. The results showed that, the $I_{CE\ (90°\ \&\ 25°)}$ and $I_{CE\ (90°\ \&\ 50°)}$ correlated significantly with the $I_{BOLD\ (90°\ \&\ 25°)}$ and $I_{BOLD\ (90°\ \&\ 50°)}$ in V1, respectively (90° & 25°: $r = 0.465$, $P = 0.039$; 90° & 50°: $r = 0.472$, $P = 0.036$, Fig. 3 **C**, left), but not in other cortical areas (Fig. 3 **C**, right). Taken together, these results further indicate a close relationship between the V1 activities and psychophysical cueing effect in both visible and invisible conditions.

3.2.2. Whole-brain group analysis

In both Visible and Invisible conditions, to identify potential cortical or subcortical area(s) whose activities were modulated by the three types of texture stimuli (90° & 0°, 90° & 25°, and 90° & 50°), we performed a group analysis and did a whole-brain search with a general linear model (GLM) procedure (Friston et al., 1994) for cortical and subcortical area(s) that exhibited a differential response to these three conditions, once their respective mask signals were subtracted (note that the data of each type of texture stimuli from the left [i.e., 90° + 0°] and right [i.e., 0° + 90°] visual fields were pooled together for analysis as there was no significant difference between contralateral and ipsilateral sides to the 90°-foreground in pIPS, aIPS, or FEF, Fig. S7). Statistical maps were thresholded at $p < 0.05$ and corrected by FDR correction (Genovese et al., 2002). The results showed that, in the Visible condition, the bilateral pIPS (lpIPS: $F(2,38) = 3.661$, $P = 0.038$, $\eta_p^2 = 0.162$; rpIPS: $F(2,38) = 3.749$, $P = 0.045$, $\eta_p^2 = 0.165$), aIPS (laIPS: $F(2,38) = 4.578$, $P = 0.018$, $\eta_p^2 = 0.194$; raIPS: $F(2,38) = 7.521$, $P = 0.003$, $\eta_p^2 = 0.284$), and FEF (lFEF: $F(2,38) = 4.829$, $P = 0.022$, $\eta_p^2 = 0.203$; rFEF: $F(2,38) = 4.155$, $P = 0.024$, $\eta_p^2 = 0.179$) demonstrated a significant difference among the three types of texture stimuli (Fig. 4 **A**). Post hoc paired t tests revealed that the BOLD signal difference of 90° & 50° and 90° & 25° texture stimuli were significantly higher than that of the 90° & 0° texture stimuli in lpIPS ($t(19) = 3.043$, $P = 0.02$, Cohen's $d = 0.394$) and rpIPS ($t(19) = 3.708$, $P = 0.004$, Cohen's $d = 0.426$), respectively, and no significant differences were found for other comparisons. For the bilateral aIPS, only the BOLD signal difference of 90° & 50° texture stimuli was significantly higher than that of the 90° & 0° texture stimuli (laIPS: $t(19) = 2.737$, $P = 0.039$, Cohen's $d = 0.410$; raIPS: $t(19) = 4.331$, $P = 0.039$, Cohen's $d = 0.475$). For the lFEF, there was no significant difference in the BOLD signal difference between 90° & 25° and 90° & 50° texture stimuli ($t(19) = -0.357$, $P = 1.000$, Cohen's $d = -0.058$). The BOLD signal difference of 90° & 0° texture stimuli was significantly lower than that of the 90° & 50° texture stimuli ($t(19) = 3.809$, $P = 0.004$, Cohen's $d = 0.380$), but not of the 90° & 25° texture stimuli ($t(19) = 2.513$, $P = 0.063$, Cohen's $d = 0.338$). For the rFEF, however, only the BOLD signal difference of 90° & 50° texture stimuli was significantly higher than that of the 90° & 0° texture stimuli ($t(19) = 2.663$, $P = 0.046$, Cohen's $d = 0.359$). Together, these results indicated that the relative levels of activation in pIPS, aIPS, and FEF across the three types of texture stimuli showed a reverse pattern to that observed in V1–V4 (Fig. 3 **A**). Furthermore, for both 90° & 25° and 90° & 50° texture stimuli, we calculated the correlation coefficients between their I_{BOLD} in V1 and those in the bilateral pIPS, aIPS, and FEF across individual subjects. We found that both the $I_{BOLD\ (90°\ \&\ 25°)}$ and $I_{BOLD\ (90°\ \&\ 50°)}$ in V1 correlated significantly and negatively with those in rpIPS (90° & 25°: $r = -0.468$, $P = 0.038$; 90° & 50°: $r = -0.603$, $P = 0.005$, Fig. 4 **B**), but not in either lpIPS or the bilateral aIPS and FEF (Fig. 4 **C**). These results indicate a potential involvement of rpIPS in the graded manner of saliency map in V1. In the Invisible condition, however, no such cortical or subcortical areas were found, further supporting the idea that the non-graded manner of saliency map was constructed in V1.

3.2.3. Effective connectivity analysis

In the Visible condition, our results showed that the BOLD signal in both V1 and frontoparietal cortical areas was modulated by the three types of texture stimuli (90° & 0°, 90° & 25°, and 90° & 50°), indicating a graded manner of saliency map. To further examine which area is a potential source of this graded manner of saliency map, we applied dynamic causal modeling (DCM) analysis (Friston et al., 2003) in SPM12 to examine interregional intrinsic connectivity change among these three types of texture stimuli. Given the extrinsic visual input into both V1T (i.e., the ROI in V1 that was evoked by the target: 90°-foreground) and V1D (i.e., the ROI in V1 that was evoked by the distractor), we posited a network of bidirectional connections linking pIPS, aIPS, FEF, V1T, and V1D, and assessed the degree of modulation of each connection across the three types of texture stimuli.

Modulation of connectivity was determined almost exclusively for backward connections: congruously, for both the left (Fig. 5 **A**) and right (Fig. 5 **E**) hemispheres, the backward connection was decreased from pIPS to both V1T (left hemisphere: $F(2,38) = 21.578$, $P < 0.001$, $\eta_p^2 = 0.532$; right hemisphere: $F(2,38) = 13.035$, $P < 0.001$, $\eta_p^2 = 0.407$) and V1D (left hemisphere: $F(2,38) = 9.114$, $P = 0.001$, $\eta_p^2 = 0.324$; right hemisphere: $F(2,38) = 8.305$, $P = 0.001$, $\eta_p^2 = 0.304$), but not from either aIPS or FEF. Post hoc paired t tests revealed that, for lpIPS → V1T (Fig. 5 **B**), the 90° & 25° texture stimuli was significantly lower than the 90° & 0° texture stimuli ($t(19) = -4.039$, $P = 0.002$, Cohen's $d = -1.068$) but significantly higher than the 90° & 50° texture stimuli ($t(19) = 3.289$, $P = 0.012$, Cohen's $d = 0.655$); for rpIPS → V1T (Fig. 5 **F**), there was no significant difference between the 90° & 25° and 90° & 50° ($t(19) = 1.496$, $P = 0.453$, Cohen's $d = 0.377$) texture stimuli, and both were significantly lower than the 90° & 0° texture stimuli (90° & 25° versus 90° & 0°: $t(19) = -3.167$, $P = 0.015$, Cohen's $d = -0.940$; 90° & 50° versus 90° & 0°: $t(19) = -5.623$, $P < 0.001$, Cohen's $d = -1.373$). For lpIPS → V1D (Fig. 5 **B**), there was no significant difference between the 90° & 0° and 90° & 25° ($t(19) = 1.014$, $P = 0.970$, Cohen's $d = 0.268$) texture stimuli, and both were significantly higher than the 90° & 50° texture stimuli (90° & 0° versus 90° & 50°: $t(19) = 3.663$, $P = 0.005$, Cohen's $d = 0.989$; 90° & 25° versus 90° & 50°: $t(19) = 3.414$, $P = 0.009$, Cohen's $d = 0.980$); for rpIPS → V1D (Fig. 5 **F**), only the 90° & 50° texture stimuli was significantly lower than the 90° & 0° texture stimuli ($t(19) = -4.012$, $P = 0.002$, Cohen's $d = -0.140$). In addition, there was a significant change in backward connection from aIPS to pIPS for both hemispheres (left hemisphere: $F(2,38) = 8.077$, $P = 0.002$, $\eta_p^2 = 0.298$; right hemisphere: $F(2,38) = 8.985$, $P < 0.001$, $\eta_p^2 = 0.321$). Post hoc paired t tests revealed that, for both hemispheres, there was no significant difference between the 90° & 25° and 90° & 50° (left hemisphere: $t(19) = -1.700$, $P = 0.316$, Cohen's $d = 0.444$; right hemisphere: $t(19) = -1.663$, $P = 0.338$, Cohen's $d = -0.549$) texture stimuli. The 90° & 0° texture stimuli was significantly lower than the 90° & 50° texture stimuli (left hemisphere: $t(19) = 3.489$, $P = 0.007$, Cohen's $d = 1.175$; right hemisphere: $t(19) = 4.138$, $P < 0.001$, Cohen's $d = 1.311$), but was not significantly lower than the 90° & 25° texture stimuli (left hemisphere: $t(19) = 2.551$, $P = 0.059$, Cohen's $d = 0.754$; right hemisphere: $t(19) = 2.603$, $P = 0.052$, Cohen's $d = 0.803$). Differently, for the left hemisphere (Fig. 5 **B**), we found a significant feedforward connection from V1D to FEF ($F(2,38) = 5.323$, $P = 0.012$, $\eta_p^2 = 0.219$), but not to either pIPS or aIPS. Post hoc paired t tests revealed that, only the 90° & 50° texture stimuli was significantly higher than the 90° & 0° texture stimuli ($t(19) = 2.801$, $P = 0.034$, Cohen's $d = 0.840$). For the right hemi-

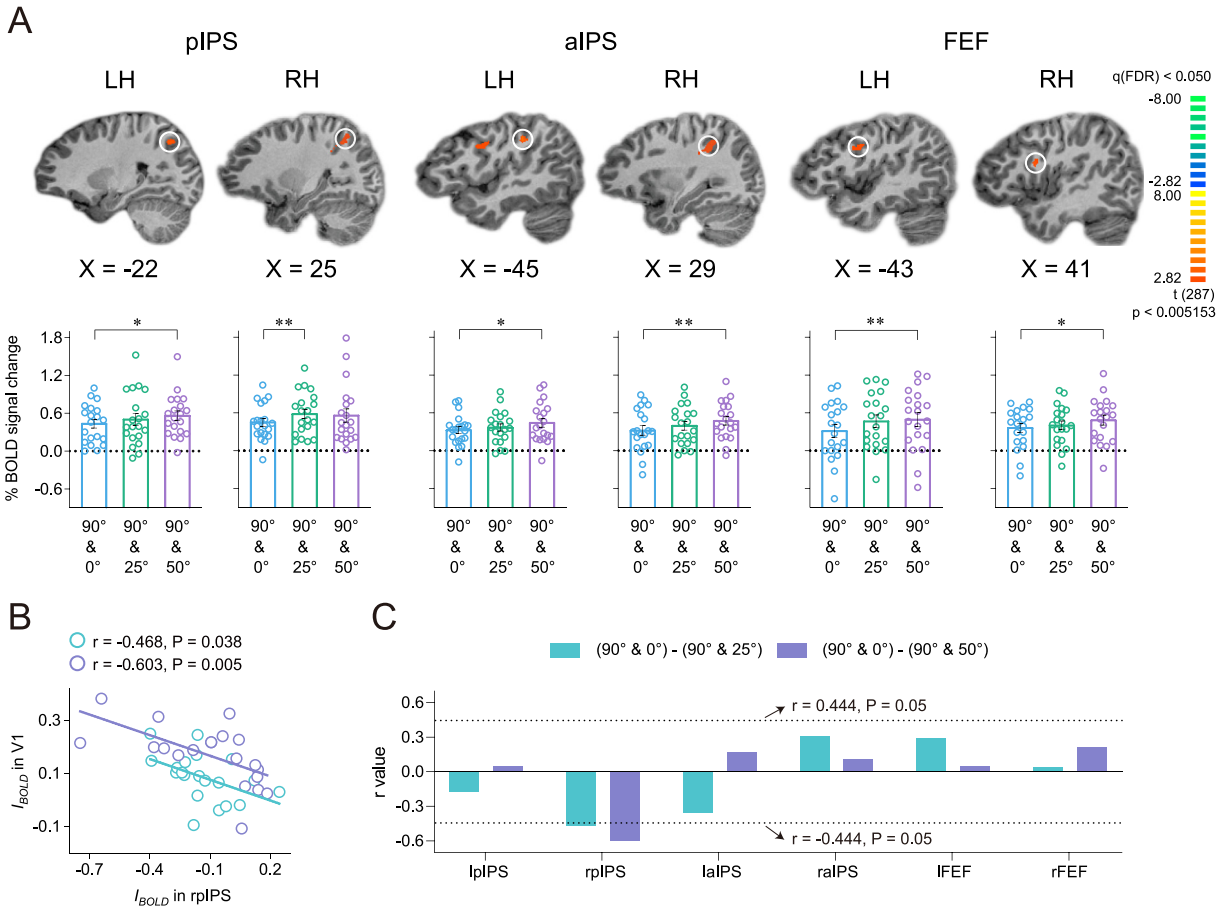

Fig. 4. Results of whole-brain group analysis. ***A*** Whole-brain search for the bilateral pIPS (left), aIPS (middle), and FEF (right), with all showing a significant difference in the BOLD signal change among the three types of texture stimuli (90° & 0°, 90° & 25°, and 90° & 50°) during the Visible condition. Note that no such cortical or subcortical areas were found in the Invisible condition (*$P < 0.05$; **$P < 0.01$). Error bars denote 1 SEM calculated across subjects and colored dots denote the data from each subject (LH: left hemisphere; RH: right hemisphere). ***B*** Correlations between the I_{BOLD} in V1 and that in rpIPS across individual subjects during the Visible condition. ***C*** Correlation coefficients (r values) between the I_{BOLD} in V1 and that in other cortical areas across individual subject during the Visible condition.

sphere (Fig. 5 **F**), however, we found a significant backward connection from FEF to pIPS ($F(2,38) = 10.990$, $P = 0.001$, $\eta_p^2 = 0.366$). Post hoc paired *t* tests revealed that there was no significant difference between the 90° & 0° and 90° & 25° ($t(19) = 0.446$, $P = 1.000$, Cohen's $d = 0.146$) texture stimuli, and both were significantly lower than the 90° & 50° texture stimuli (90° & 0° versus 90° & 50°: $t(19) = -3.822$, $P = 0.003$, Cohen's $d = -1.026$; 90° & 25° versus 90° & 50°: $t(19) = -6.971$, $P < 0.001$, Cohen's $d = -1.069$).

To further evaluate the role of these forward and backward connections in the mapping of salience, we computed an interference of the distractor to quantify how much the intrinsic connections (I_{IC}) changed in both the 90° & 25° and 90° & 50° texture stimuli relative to that in the 90° & 0° texture stimuli. The interference was calculated as follows: $I_{IC\ (90^\circ\ \&\ 25^\circ)} = IC_{90^\circ\ \&\ 0^\circ} - IC_{90^\circ\ \&\ 25^\circ}$ *and* $I_{IC(90^\circ\ \&\ 50^\circ)} = IC_{90^\circ\ \&\ 0^\circ} - IC_{90^\circ\ \&\ 50^\circ}$, where $IC_{90^\circ\ \&\ 0^\circ}$, $IC_{90^\circ\ \&\ 25^\circ}$, and $IC_{90^\circ\ \&\ 50^\circ}$ are the interregional intrinsic connections for the 90° & 0°, 90° & 25°, and 90° & 50° texture stimuli, respectively. Subsequently, for each interregional intrinsic connection and each interference, we calculated the correlation coefficients between the I_{IC} and the I_{BOLD}, and between the I_{IC} and the I_{CE} across individual subjects. Results showed that the $I_{IC\ (90^\circ\ \&\ 25^\circ)}$ and $I_{IC\ (90^\circ\ \&\ 50^\circ)}$ for the bilateral pIPS → V1T correlated significantly with the $I_{BOLD\ (90^\circ\ \&\ 25^\circ)}$ (lpIPS → V1T: $r = 0.477$, $P = 0.033$; rpIPS → V1T: $r = 0.516$, $P = 0.020$) and $I_{BOLD\ (90^\circ\ \&\ 50^\circ)}$ (lpIPS → V1T: $r = 0.570$, $P = 0.009$; rpIPS → V1T: $r = 0.457$, $P = 0.043$) in V1, respectively, but not for other cortical areas (Fig. 5 **C** and 5 **G**). The $I_{IC\ (90^\circ\ \&\ 25^\circ)}$ and $I_{IC\ (90^\circ\ \&\ 50^\circ)}$ for rpIPS → V1T correlated significantly with the $I_{CE\ (90^\circ\ \&\ 25^\circ)}$ ($r = 0.451$, $P = 0.046$) and $I_{CE\ (90^\circ\ \&\ 50^\circ)}$ ($r = 0.468$, $P = 0.037$), respectively, but not for either lpIPS or other cortical areas (Fig. 5 **D** and 5 **H**). Together, our results implied that the graded manner of saliency map in V1 was derived by feedback from pIPS rather than from aIPS or FEF, and further indicated a slight right pIPS advantage in controlling this graded mapping.

4. Discussion

We examined how the mapping of the saliency interacts with awareness when multiple salient stimuli are presented simultaneously and found the following psychophysical and neuroimaging results. First, we found support for previous neurophysiological (Kastner et al., 1997; Nothdurft et al., 1999; White et al., 2017a; 2017b; Yan et al., 2018), psychophysical (Koene and Zhaoping, 2007; Zhaoping, 2008; Zhaoping and May 2007; Zhaoping and Zhe, 2015), and brain imaging (Chen et al., 2016; Zhang et al., 2012) studies, indicating that both visible and invisible salient foregrounds could attract bottom-up attention in behavior and evoked greater BOLD signals relative to the background in the texture stimuli. Second, however, there was a critical distinction between their dependence on awareness, with visible and invisible salient foregrounds guiding bottom-up attention in a graded or non-graded manner, respectively. Finally, a plausible account or our results is that the graded manner of saliency map in V1 was derived by feedback from pIPS (in particular, the right pIPS), whereas the non-graded manner of saliency map was created in V1.

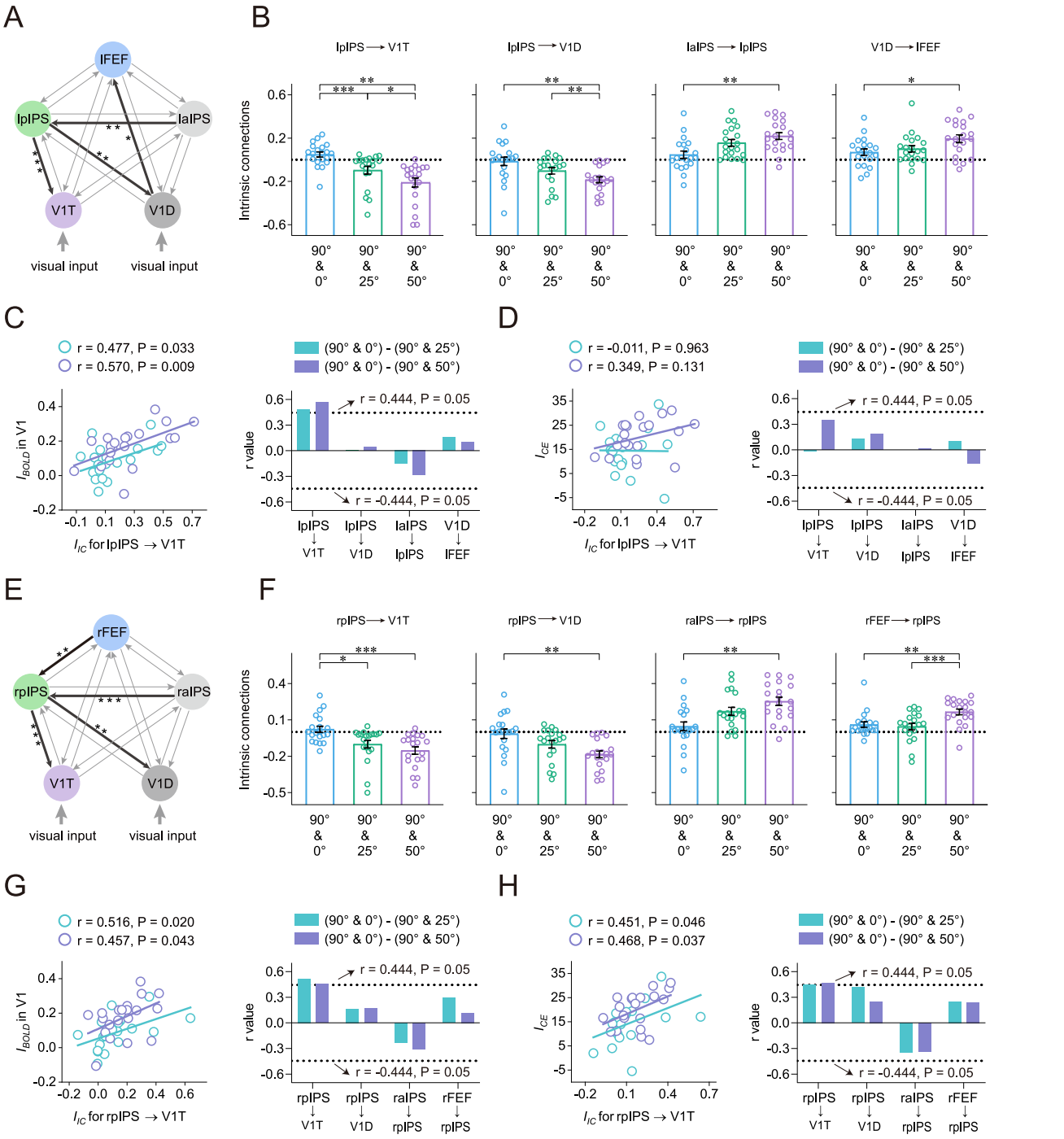

Fig. 5. Results of DCM analysis. ***A*** & ***E*** Given the extrinsic visual input into both V1T (i.e., the ROI in V1 that was evoked by the target: 90°-foreground) and V1D (i.e., the ROI in V1 that was evoked by the distractor), bidirectional connections were hypothesized to exist among the pIPS, aIPS, FEF, V1T, and V1D. Thick lines indicate intrinsic connections that are significantly modulated by the three types of texture stimuli (90° & 0°, 90° & 25°, and 90° & 50°) and their significance levels during the Visible condition (*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$) for the left (***A***) and right (***E***) hemisphere. ***B*** & ***F*** Intrinsic connections among the pIPS, aIPS, V1T, and V1D for the left (***B***) and right (***F***) hemisphere. Error bars denote 1 SEM calculated across subjects and colored dots denote the data from each subject. ***C*** & ***G*** Left: correlations between the I_{BOLD} in V1 and the I_{IC} of feedback connectivity from lpIPS (***C***) and rpIPS (***G***) to V1T, across individual subjects; Right: correlation coefficients (r values) between the I_{BOLD} in V1 and the I_{IC} of each significant intrinsic connection across individual subjects for the left (***C***) and right (***G***) hemisphere. ***D*** & ***H*** Left: correlations between the I_{CE} and the I_{IC} of feedback connectivity from lpIPS (***D***) and rpIPS (***H***) to V1T, across individual subjects; Right: correlation coefficients (r values) between the I_{CE} and the I_{IC} of each significant intrinsic connection across individual subjects for the left (***D***) and right (***H***) hemisphere.

4.1. Graded manner of the saliency map in pIPS

In accordance with previous studies demonstrating a direct representation for pIPS of the saliency map (Bisley and Goldberg, 2010; Bogler et al., 2011; Buschman and Miller, 2007; Constantinidis and Steinmetz, 2005; Gottlieb, 2007; Gottlieb et al., 1998; Serences et al., 2005), our study further revealed that pIPS could control the graded manner of the saliency map in V1. One should note that the graded manner in our study was indicated by a decreasing response to the 90°-foreground associated with increasing saliency of the distractor (i.e.,

reflecting increasing competition between the 90°-foreground and distractor). Previous studies have implicated IPS in the competition between the target and distractor (Mevorach et al., 2009, 2010; Payne and Allen, 2011; Ruff and Driver, 2006), and our findings are consistent with such an influence. Our findings suggested that the competition between the target (i.e., 90°-foreground) and distractor in V1 could be associated with feedback from pIPS. First, the bilateral pIPS responses increased with the competition between the 90°-foreground and distractor and showed a reverse pattern (Fig. 4 **A**) with that in V1, where the BOLD signal of 90°-foreground decreased with the competition. Second, across subjects, these decreased responses in V1 were significantly predicted by the enhanced response in the right pIPS (Fig. 4 **B**) rather than in other frontoparietal cortical areas (Fig. 4 **C**). Finally, the DCM analysis indicated that feedback from bilateral pIPS to V1 gained in suppressive influence with enhancement of competition (Fig. 5 **B** and 5 **F**); across subjects, this significantly predicted both the reduced responses in V1 (Fig. 5 **C** and 5 **G**) and, for right pIPS, the cueing effect in psychophysical experiments (Fig. 5 **H**). In addition, the more significant influences from the right rather than the left pIPS evident in our study also demonstrated a slight right pIPS advantage in controlling the graded manner of saliency map in V1, supporting the well-known hemispheric asymmetries of the attention networks in human visual system (Bartolomeo and Malkinson, 2019). Moreover, several transcranial magnetic stimulation (TMS, Hodsoll et al., 2009; Mevorach et al., 2006, 2009, 2010; Ruff et al., 2009; van Koningsbruggen, et al., 2010) and transcranial direct current stimulation (tDCS, Lo et al., 2019; Moos et al., 2012; Roy et al., 2015) studies have revealed a unique role of the right but not the left parietal cortex in the attentional control. Thus, further work is needed to use TMS, tDCS, or other techniques to parse the relative contributions of the left and right pIPS to the graded manner of saliency map in V1.

Our findings can be viewed as identifying the human parietal cortex as a source of the graded manner of saliency map. Note that, this conclusion is based mainly on our DCM analyses, which depended on time-series models of fMRI data for an interpretation of causality (Friston et al., 2003). The interpretation of causality in our study finds support in previous lesion (Corbetta and Shulman, 2011; Shomstein, 2012) and TMS (Hodsoll et al., 2009; Mevorach et al., 2006) studies showing a causal effect of parietal cortical disruption on bottom-up attention driven by the salient stimulus. The prominent role of the parietal cortex in the graded manner of saliency map evident here not only is consistent with recent neurophysiological findings that have revealed how parietal areas directly realize the saliency map (Bisley and Goldberg, 2010; Buschman and Miller, 2007; Constantinidis and Steinmetz, 2005; Gottlieb, 2007; Gottlieb et al., 1998), but also support the dominant model of the saliency map developed by Itti and Koch (2001), which proposes that higher cortical areas, particularly the parietal and frontal cortex, whose neurons are less selective to specific visual features (i.e., color, orientation, or other features, Koch and Ullman, 1985; Wolfe, 1994), are more likely to be possible candidates that realize the saliency map.

Although we emphasize the importance of pIPS in the graded manner of saliency map, we cannot deny a potential contribution from other frontoparietal cortical areas, such as aIPS and FEF. Actually, previous neurophysiological and brain imaging studies (Bogler et al., 2011; Geng and Mangun, 2009; Serences and Yantis, 2007; Thompson and Bichot, 2005) have reported that both aIPS and FEF could also represent the saliency map. Our study supported these findings by showing that the BOLD signal in aIPS and FEF (Fig. 4 **A**) and the feedforward from V1 (in particular, V1D) to FEF (Fig. 5 **B**) were significantly modulated by the texture stimuli. However, more importantly, our results showed that the feedback from either aIPS or FEF to V1 was not modulated by the texture stimuli (Fig. 5 **A** and 5 **E**), indicating that both aIPS and FEF is more likely to inherit or read out saliency signals from V1 or other earlier areas. Interestingly, we also found significant backward connections from both aIPS and FEF to pIPS in our study (Fig. 5 **B** and 5 **F**). These backward connections, however, could not significantly predict either the BOLD signal change in V1 (Fig. 5 **C** and 5 **G**) or the cueing effect in psychophysical experiments (Fig. 5 **D** and 5 **H**). These results suggest that both the involvement of aIPS and FEF in the graded manner of the saliency map in our study may occur via pIPS. Indeed, using multivariate pattern analysis of fMRI, previous studies have indicated that the pIPS and these two areas may be involved in different stages of saliency mapping (Bogler et al., 2011; Itti and Koch, 2001). Given the low temporal resolution of fMRI, further work is needed using neurophysiological techniques to parse how these two areas are involved in the mapping of saliency.

4.2. Non-graded manner of the saliency map in V1

Compared to a graded distribution of salience with awareness in pIPS, our results suggest that the salience without awareness is mapped as a non-graded manner and this manner is created in V1. Our claim was based on the following findings obtained with the Invisible condition. First, both the cueing effects (Fig. 1 **E**) and BOLD signal changes (Fig. 3 **B**) of 90°-foreground were not significantly interfered by the distractor (25°- and 50°-foregrounds) as evidenced by the difference among 90° & 0°, 90° & 25°, and 90° & 50° texture stimuli. Second, the psychophysical cueing effect was significantly correlated with the BOLD signal change in V1 (Fig. 3 **D**, left), but not elsewhere (Fig. 3 **D**, right). More importantly, the correlation coefficient for V1 was significantly larger than those for other cortical areas (Fig. S6). Finally, a voxel-wise analysis failed to detect any cortical or subcortical areas that were modulated by the distractor in the Invisible condition, indicating that the observed BOLD signals in V1 were unlikely to arise from interactions with other brain areas. Our study thus directly linked V1 activity with a non-graded manner of saliency mapping in the absence of awareness, consistent with existing psychophysical (Koene and Zhaoping, 2007; Zhaoping, 2008; Zhaoping and May 2007; Zhaoping and Zhe, 2015), neurophysiological (Kastner et al., 1997; Nothdurft et al., 1999; Yan et al., 2018), and brain imaging (Chen et al., 2016; Zhang et al., 2012) studies, as well as the computational model (Li, 1999, 2002), implying that V1 can be self-sufficient in generating a salience map. Notably, our study aims to identify the brain area that creates rather than represents the saliency map. According to our results, we believe that neural activities in V1 could determine an early attentional selection even if subjects are unaware of texture stimuli. Meanwhile, we are not claiming that other areas could not represent a saliency map. The intermediate or higher brain areas are more likely to inherit or read out saliency signals from V1 rather than create a saliency map within themselves.

In addition, our results suggested that the non-graded mapping of salience without awareness evident here might be derived by a WTA-like mechanism, i.e., a WTA-like competition between the high-salient target and low-salient distractor, as showing by a null difference in bottom-up attention of the 90°-foreground among our three types of texture stimuli (Figs. 1 **E** and 3 **B**). Moreover, this WTA-like mechanism was further supported through analyzing the reaction time in the valid and invalid cue conditions separately (Fig. S4**D**), and is compatible with the V1 saliency theory (Li, 1999, 2002), which proposes that saliency of a visual location is determined by its highest evoked V1 response relative to those evoked by other locations. However, we cannot deny a potential contribution from the priming mechanism in our non-graded mapping of salience without awareness (see “Results”) and further work is needed to clarify the relative contributions of these two mechanisms in this processing.

4.3. Awareness-dependent mapping of the saliency

Previous studies have reported ostensibly conflicting results with regard to the neural substrate of saliency map. Three of our previous studies (Chen et al., 2016; Huang et al., 2020; Zhang et al., 2012) suggest that discrepancies in the literature findings may have resulted from the

possible contamination by top-down signals, specifically the conscious access to salient stimuli, which have not been systematically controlled or manipulated. Here our study addresses it using a backward masking paradigm in which the low- and high-luminance mask renders the salient foreground (and indeed the whole texture stimuli) visible and invisible to subjects, respectively. We assume that relative to the visible foreground, the invisible foreground can maximally reduce various top-down contaminations, such as feature perception, object recognition, and subjects' intentions (Zhang et al., 2012). It has been proposed and widely accepted that subjective awareness is determined by top-down signaling (Del Cul et al. 2007; Mashour et al., 2020). Thus, rendering a stimulus invisible could maximally (although not completely, Boly et al., 2017) reduce top-down influences, particularly as indexed by temporally sluggish fMRI signals that typically reflect neural activities resulting from both bottom-up and top-down processes (Fang et al., 2008). We thus speculate that the higher cortical areas, particularly the parietal and frontal cortex, are more likely to be the neural substrate of saliency map with the visible stimulus, since these areas are able to integrate top-down and bottom-up attention (Bisley and Goldberg, 2010; Bogler et al., 2011; Geng and Mangun, 2009; Gottlieb et al., 1998; Katsuki and Constantinidis, 2012; Squire et al., 2013). Conversely, early visual areas are more likely to be possible candidates that realize a pure salience map of the invisible stimulus, specifically the area V1 that is nearly independent of top-down modulation. Indeed, previous studies have demonstrated that top-down control modulates extrastriate areas but not V1 (Kastner et al., 1998; Luck et al., 1997; Melloni et al., 2012) and decreases gradually from extrastriate visual areas to V1 (Buffalo et al., 2010; Liu et al., 2005). Accordingly, our findings support these speculations and identify the parietal cortex and V1 as the neural substrates of saliency maps with and without awareness, respectively. Meanwhile, it should be noted that some p values of the significant results in our study were between 0.01 and 0.05, which (although significant) often has a small effect size. We suggested that these small effect sizes might be due to the low sample size, low temporal resolution of fMRI, or both. Thus, further works are needed to address whether our conclusion can be replicated with a large sample size or using neurophysiological techniques.

In addition, at least four intriguing questions need to be addressed in further research. First, a number of studies have indicated a crucial role of the SC (Fecteau and Munoz, 2006; Krauzlis et al., 2013; Kustov and Robinson, 1996) in realizing the saliency map. Two recent papers even demonstrated that neurons in the superficial layers of SC encoded saliency earlier and more robustly than V1 neurons (White et al., 2017a; 2017b). However, ROIs in the SC defined in our study were across superficial and intermediate laminae. The SC superficial layer (SCs) is interconnected with multiple visual areas and can be in an ideal location to pool diverse visual inputs to form a feature-agnostic saliency representation; the SC intermediate layer (SCi), conversely, linked to the control of goal-directed attention and gaze behavior, best representing the priority rather than the saliency map (White et al., 2017a; 2017b). Besides, one could argue that compared with the invisible condition that maximally reduces top-down influences, the visible condition in our study, presumably, examined the mapping of priority map, in which activity is driven by a combination of low-level salience and various top-down influences (see Bisley and Mirpour, 2019 for a review). In other words, our study may not demonstrate an awareness-dependent mapping of saliency, but instead, it reveals the distinct neural substrates for the priority and saliency maps. Given the distinct involvements of SCi and SCs in these two maps, further work is thus needed using neurophysiological techniques or ultrahigh field fMRI to parse the exact role of awareness in the mapping of saliency. Second, several studies have indicated that the bilateral competition between the target and distractor is more marked in the superior hemifield than that in the inferior hemifield (Previc, 1996; Previc and Blume, 1993; Shipp, 2011). In other words, the awareness-dependent mapping of salience evident here could be modulated by whether the target and distractor are presented in the superior or inferior hemifield. Third, it might be argued that identifying V1 as the neural substrate of salience map without awareness in the current study owes much to our choice to use a backward masking paradigm to manipulate subjects' awareness. Indeed, the backward masking is well known to interrupt the recurrent processing and in blocking the processing of masked information in higher visual areas (Enns and DiLollo, 2000) and thus it might automatically infer a prominent role upon early visual cortex (particularly V1) in the mapping of saliency without awareness. Accordingly, further work is worthwhile to use the subliminal stimulation paradigm (Mulckhuyse and Theeuwes, 2010; Zhang and Fang, 2012) to manipulate subjects' awareness since in this paradigm, conversely, the weak stimuli were able to activate higher visual areas, although the activation magnitude could be much reduced compared to visible stimuli (Dehaene et al., 2001). Finally, compared with current stimuli that consist of simple oriented bars (Fig. 1 **A**), complex natural scenes that contain richer naturalistic low-level features (e.g., luminance, contrast, orientation, spatial frequency, and curvature, all of which are highly tuned by the visual system), that are thought to be optimal for automatic bottom-up attention (Bogler et al., 2011; Chen et al., 2016; White et al., 2017a). Further work is thus needed to address whether our conclusion can be generalized to the complex natural scene.

5. Conclusions

We conclude that, when multiple salient stimuli are presented simultaneously, the salience is mapped in either a graded or non-graded manner, depending on the conscious access to salient stimuli. Our study provides, to the best of our knowledge, the first evidence for awareness-dependent mapping of saliency and its distinct neural loci. Identifying the parietal cortex and V1 as the neural substrates of saliency maps with and without awareness, respectively, reconciles previous, seemingly contradictory findings regarding the nature of salience mapping.

Data and code availability

The code and MRI dataset generated during this study is available on Open Science Framework: https://osf.io/u7sbf/.

Declaration of Competing Interest

The authors declare no competing financial interests.

Credit authorship contribution statement

Lijuan Wang: Writing – original draft, Writing – review & editing, Methodology, Software, Validation, Investigation, Visualization. **Ling Huang:** Methodology, Investigation, Validation, Visualization. **Mengsha Li:** Methodology, Investigation, Validation, Visualization. **Xiaotong Wang:** Methodology, Investigation, Validation, Visualization. **Shiyu Wang:** Methodology, Investigation, Validation, Visualization. **Yuefa Lin:** Methodology, Investigation, Validation, Visualization. **Xilin Zhang:** Writing – original draft, Writing – review & editing, Conceptualization, Funding acquisition, Supervision.

Acknowledgments

We thank Li Zhaoping for valuable comments. This work was supported by the National Natural Science Foundation of China (Projects 32022032 and 31871135) and the Key Realm R&D Program of Guangzhou (202007030005).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2021.118864.

References

Allman, J., Miezin, F., McGuinness, E., 1985. Stimulus specific responses from beyond the classical receptive field: neurophysiological mechanisms for local-global comparisons in visual neurons. Annu. Rev. Neurosci. 8, 407–430.

Asplund, C.L., Todd, J.J., Snyder, A.P., Marois, R., 2010. A central role for the lateral prefrontal cortex in goal-directed and stimulus-driven attention. Nat. Neurosci. 13, 507–512.

Baluch, F., Itti, L., 2011. Mechanisms of top-down attention. Trends Neurosci. 34, 210–224.

Bartolomeo, P., Malkinson, T.S., 2019. Hemispheric lateralization of attention processes in the human brain. Curr. Opin. Psychol. 29, 90–96.

Basso, M.A., Wurtz, R.H., 2002. Neuronal activity in substantia nigra pars reticulata during target selection. J. Neurosci. 22, 1883–1894.

Bisley, J.W., Goldberg, M.E., 2010. Attention, intention, and priority in the parietal lobe. Annu. Rev. Neurosci. 33, 1–21.

Bisley, J.W., Mirpour, K., 2019. The neural instantiation of a priority map. Curr. Opin. Psychol. 29, 108–112.

Bogler, C., Bode, S., Haynes, J.D., 2011. Decoding successive computational stages of saliency processing. Curr. Biol. 21, 1667–1671.

Boly, M., Massimini, M., Tsuchiya, N., Postle, B.R., Koch, C., Tononi, G., 2017. Are the neural correlates of consciousness in the front or in the back of the cerebral cortex? Clinical and neuroimaging evidence. J. Neurosci. 37, 9603–9613.

Buffalo, E.A., Fries, P., Landman, R., Liang, H., Desimone, R., 2010. A backward progression of attentional effects in the ventral stream. Proc. Natl. Acad. Sci. USA. 107, 361–365.

Burrows, B.E., Moore, T., 2009. Influence and limitations of pop out in the selection of salient visual stimuli by area V4 neurons. J. Neurosci. 29, 15169–15177.

Buschman, T.J., Miller, E.K., 2007. Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. Science 315, 1860–1862.

Cavanagh, P., Alvarez, G.A., 2005. Tracking multiple targets with multifocal attention. Trends Cogn. Sci. 9, 349–354.

Cheal, M.L., Lyon, D.R., Gottlob, L.R., 1994. A framework for understanding the allocation of attention in location-cued discrimination. Q. J. Exp. Psychol. 47, 699–739.

Chen, C., Zhang, X., Wang, Y., Zhou, T., Fang, F., 2016. Neural activities in V1 create the bottom-up saliency map of natural scenes. Exp. Brain Res. 234, 1769–1780.

Constantinidis, C., Steinmetz, M.A., 2005. Posterior parietal cortex automatically encodes the location of salient stimuli. J. Neurosci. 25, 233–238.

Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus driven attention in the brain. Nat. Rev. Neurosci. 3, 201–215.

Corbetta, M., Shulman, G.L., 2011. Spatial neglect and attention networks. Annu. Rev. Neurosci. 34, 569–599.

Dehaene, S., Naccache, L., Cohen, L., Le Bihan, D., Mangin, J.F., Poline, J.B., Rivière, D., 2001. Cerebral mechanisms of word masking and unconscious repetition priming. Nat. Neurosci. 4, 752–758.

Del Cul, A., Baillet, S., Dehaene, S., 2007. Brain dynamics underlying the nonlinear threshold for access to consciousness. PLoS Biol 5, e260.

Engel, S.A., Glover, G.H., Wandell, B.A., 1997. Retinotopic organization in human visual cortex and the spatial precision of functional MRI. Cereb. Cortex 7, 181–192.

Enns, J.T., Di Lollo, V., 2000. What's new in visual masking? Trends Cogn. Sci. 4, 345–352.

Eriksen, C.W., Yeh, Y.Y., 1985. Allocation of attention in the visual field. J. Exp. Psychol. Hum. Percept. Perform. 11, 583.

Fang, F., Boyaci, H., Kersten, D., Murray, S.O., 2008. Attention-dependent representation of a size illusion in human V1. Curr. Biol. 18, 1707–1712.

Faul, F., Erdfelder, E., Buchner, A., Lang, A.G., 2009. Statistical power analyses using G* Power 3.1: tests for correlation and regression analyses. Behav. Res. Methods 41, 1149–1160.

Fecteau, J.H., Munoz, D.P., 2006. Salience, relevance, and firing: a priority map for target selection. Trends. Cogn. Sci. 10, 382–390.

Fisher, R.A., 1921. On the probable error of a coefficient of correlation deduced from a small sample. Metron 1, 1–32.

Folk, C.L., Hoyer, W.J., 1992. Aging and shifts of visual spatial attention. Psychol. Aging 7, 453–465.

Franconeri, S.L., Alvarez, G.A., Cavanagh, P., 2013. Flexible cognitive resources: competitive content maps for attention and memory. Trends. Cogn. Sci. 17, 134–141.

Friesen, C.K., Kingstone, A., 1998. The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. Psychon. Bull. Rev. 5, 490–495.

Friston, K.J., Harrison, L., Penny, W., 2003. Dynamic causal modelling. Neuroimage 19, 1273–1302.

Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J.P., Frith, C.D., Frackowiak, R.S., 1994. Statistical parametric maps in functional imaging: a general linear approach. Hum. Brain Mapp. 2, 189–210.

Geng, J.J., Mangun, G.R., 2009. Anterior intraparietal sulcus is sensitive to bottom-up attention driven by stimulus salience. J. Cogn. Neurosci. 21, 1584–1601.

Genovese, C.R., Lazar, N.A., Nichols, T., 2002. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. Neuroimage 15, 870–878.

Gottlieb, J., 2007. From thought to action: the parietal cortex as a bridge between perception, action, and cognition. Neuron 53, 9–16.

Gottlieb, J.P., Kusunoki, M., Goldberg, M.E., 1998. The representation of visual salience in monkey parietal cortex. Nature 391, 481–484.

Hawkins, H.L., Hillyard, S.A., Luck, S.J., Mouloua, M., Downing, C.J., Woodward, D.P., 1990. Visual attention modulates signal detectability. J. Exp. Psychol. Hum. Percept. Perform. 16, 802–811.

Hegdé, J., Felleman, D.J., 2003. How selective are V1 cells for pop-out stimuli? J. Neurosci. 23, 9968–9980.

Hodsoll, J., Mevorach, C., Humphreys, G.W., 2009. Driven to less distraction: rTMS of the right parietal cortex reduces attentional capture in visual search. Cereb. Cortex 19, 106–114.

Huang, L., Wang, L., Shen, W., Li, M., Wang, S., Wang, X., Ungerleider, L.G., Zhang, X., 2020. A source for awareness-dependent figure–ground segregation in human prefrontal cortex. Proc. Natl. Acad. Sci. USA. 117, 30836–30847.

Itti, L., Koch, C., 2001. Computational modelling of visual attention. Nat. Rev. Neurosci. 2, 194–203.

Jonides, J., 1981. Voluntary vs. automatic control over the mind's eye's movement. In: XI, M.I.Posner, Marin, O. (Eds.), Attention and Performance. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 187–205.

Kanwisher, N., Wojciulik, E., 2000. Visual attention: insights from brain imaging. Nat. Rev. Neurosci. 1, 91–100.

Kastner, S., De Weerd, P., Desimone, R., Ungerleider, L.G, 1998. Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. Science 282, 108–111.

Kastner, S., Nothdurft, H.C., Pigarev, I.N., 1997. Neuronal correlates of pop-out in cat striate cortex. Vis. Res. 37, 371–376.

Kastner, S., Ungerleider, L.G., 2000. Mechanisms of visual attention in the human cortex. Annu. Rev. Neurosci. 23, 315–341.

Katsuki, F., Constantinidis, C., 2012. Early involvement of prefrontal cortex in visual bottom-up attention. Nat. Neurosci. 15, 1160–1166.

Klein, R.M., 2000. Inhibition of return. Trends Cogn. Sci. 4, 138–147.

Knudsen, E.I., 2011. Control from below: the role of a midbrain network in spatial attention. Eur. J. Neurosci. 33, 1961–1972.

Knudsen, E.I., 2018. Neural circuits that mediate selective attention: a comparative perspective. Trends Neurosci. 41, 789–805.

Koch, C., Ullman, S., 1985. Shifts in selective visual attention: towards the underlying neural circuitry. Hum. Neurobiol. 4, 219–227.

Koene, A.R., Zhaoping, L., 2007. Feature-specific interactions in salience from combined feature contrasts: evidence for a bottom-up saliency map in V1. J. Vis. 7, 1–14.

Krauzlis, R.J., Lovejoy, L.P., Zénon, A., 2013. Superior colliculus and visual spatial attention. Annu. Rev. Neurosci. 36, 165–182.

Kustov, A.A., Robinson, D.L., 1996. Shared neural control of attentional shifts and eye movements. Nature 384, 74–77.

Liu, T., Pestilli, F., Carrasco, M., 2005. Transient attention enhances perceptual performance and fMRI response in human visual cortex. Neuron 45, 469–477.

Li, Z., 1999. Contextual influences in V1 as a basis for pop out and asymmetry in visual search. Proc. Natl. Acad. Sci. USA. 96, 10530–10535.

Li, Z., 2002. A saliency map in primary visual cortex. Trends Cogn. Sci. 6, 9–16.

Lo, O.Y., van Donkelaar, P., Chou, L.S., 2019. Effects of transcranial direct current stimulation over right posterior parietal cortex on attention function in healthy young adults. Europ. J. Neurosci. 49, 1623–1631.

Luck, S.J., Chelazzi, L., Hillyard, S.A., Desimone, R., 1997. Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. J. Neurophysiol. 77, 24–42.

Mashour, G.A., Roelfsema, P., Changeux, J.P., Dehaene, S., 2020. Conscious processing and the global neuronal workspace hypothesis. Neuron 105, 776–798.

Mazer, J.A., Gallant, J.L., 2003. Goal-related activity in V4 during free viewing visual search. Evidence for a ventral stream visual salience map. Neuron 40, 1241–1250.

Melloni, L., van Leeuwen, S., Alink, A., Müller, N.G., 2012. Interaction between bottom-up saliency and top-down control: how saliency maps are created in the human brain. Cereb. Cortex 22, 2943–2952.

Mevorach, C., Hodsoll, J., Allen, H., Shalev, L., Humphreys, G., 2010. Ignoring the elephant in the room: a neural circuit to downregulate salience. J. Neurosci. 30, 6072–6079.

Mevorach, C., Humphreys, G.W., Shalev, L., 2006. Opposite biases in salience-based selection for the left and right posterior parietal cortex. Nat. Neurosci. 9, 740–742.

Mevorach, C., Humphreys, G.W., Shalev, L., 2009. Reflexive and preparatory selection and suppression of salient information in the right and left posterior parietal cortex. J. Cogn. Neurosci. 21, 1204–1214.

Moos, K., Vossel, S., Weidner, R., Sparing, R., Fink, G.R., 2012. Modulation of top–down control of visual attention by cathodal tDCS over right IPS. J. Neurosci. 32, 16360–16368.

Mulckhuyse, M., Theeuwes, J., 2010. Unconscious attentional orienting to exogenous cues: a review of the literature. Acta Psychol 134, 299–309.

Nakayama, K., Mackeben, M., 1989. Sustained and transient components of focal visual attention. Vis. Res 29, 1631–1647.

Nothdurft, H.C., Gallant, J.L., Van Essen, D.C., 1999. Response modulation by texture surround in primate area V1: correlates of "pop out" under anesthesia. Vis. Neurosci. 16, 15–34.

Payne, H.E., Allen, H.A., 2011. Active ignoring in early visual cortex. J. Cogn. Neurosci. 23, 2046–2058.

Penny, W.D., Stephan, K.E., Mechelli, A., Friston, K.J., 2004. Modelling functional integration: a comparison of structural equation and dynamic causal models. NeuroImage 23, S264–S274.

Posner, M.I., 1980. Orienting of attention. Q. J. Exp. Psychol. 32, 3–25.

Posner, M.I., Snyder, C.R.R., 1975. Facilitation and inhibition in the processing of signals. Am. J. Int. Law 41, 509–530.

Posner, M.I., Snyder, C.R.R., Davidson, B.J., 1980. Attention and the detection of signals. J. Exp. Psychol. 109, 160–174.

Previc, F.H., 1996. Attentional and oculomotor influences on visual field anisotropies in visual search performance. Visual Cogn. 3, 277–302.

Previc, F.H., Blume, J.L., 1993. Visual search asymmetries in three-dimensional space. Vision Res 33, 2697–2704.

Roy, L.B., Sparing, R., Fink, G.R., Hesse, M.D., 2015. Modulation of attention functions by anodal tDCS on right PPC. Neuropsychologia 74, 96–107.
Ruff, C.C., Blankenburg, F., Bjoertomt, O., Bestmann, S., Weiskopf, N., Driver, J., 2009. Hemispheric differences in frontal and parietal influences on human occipital cortex: direct confirmation with concurrent TMS–fMRI. J. Cogn. Neurosci. 21, 1146–1161.
Ruff, C.C., Driver, J., 2006. Attentional preparation for a lateralized visual distractor: behavioral and fMRI evidence. J. Cogn. Neurosci. 18, 522–538.
Serences, J.T., Shomstein, S., Leber, A.B., Golay, X., Egeth, H.E., Yantis, S., 2005. Coordination of voluntary and stimulus-driven attentional control in human cortex. Psychol. Sci. 16, 114–122.
Serences, J.T., Yantis, S., 2006. Selective visual attention and perceptual coherence. Trends Cogn. Sci. 10, 38–45.
Serences, J.T., Yantis, S., 2007. Spatially selective representations of voluntary and stimulus-driven attentional priority in human occipital, parietal, and frontal cortex. Cereb. Cortex 17, 284–293.
Sereno, M.I., Dale, A.M., Reppas, J.B., Kwong, K.K., Belliveau, J.W., Brady, T.J., Rosen, B.R., Tootell, R.B.H., 1995. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. Science 268, 889–893.
Shipp, S., 2004. The brain circuitry of attention. Trends Cogn. Sci. 8, 223–230.
Shipp, S., 2011. Interhemispheric integration in visual search. Neuropsychologia 49, 2630–2647.
Shiu, L.P., Pashler, H., 1994. Negligible effect of spatial precuing on identification of single digits. J. Exp. Psychol. Hum. Percept. Perform. 20, 1037–1054.
Shiu, L.P., Pashler, H., 1995. Spatial attention and vernier acuity. Vision Res. 35, 337–343.
Shomstein, S., 2012. Cognitive functions of the posterior parietal cortex: top-down and bottom-up attentional control. Front. Integr. Neurosci. 6, 38.
Smith, A.M., Lewis, B.K., Ruttimann, U.E., Ye, F.Q., Sinnwell, T.M., Yang, Y., Duyn, J.H., Frank, J.A., 1999. Investigation of low frequency drift in fMRI signal. Neuroimage 9, 526–533.
Squire, R.F., Noudoost, B., Schafer, R.J., Moore, T., 2013. Prefrontal contributions to visual selective attention. Annu. Rev. Neurosci. 36, 451–466.
Talairach, J., Tournoux, P., 1988. Co-Planar Stereotaxic Atlas of the Human Brain: 3-Dimensional Proportional System: an Approach to Cerebral Imaging. Thieme, New York.
Thompson, K.G., Bichot, N.P., 2005. A visual salience map in the primate frontal eye field. Prog. Brain Res. 147, 251–262.
van Koningsbruggen, M.G., Gabay, S., Sapir, A., Henik, A., Rafal, R.D., 2010. Hemispheric asymmetry in the remapping and maintenance of visual saliency maps: a TMS study. J. Cogn. Neurosci. 22, 1730–1738.
Wandell, B.A., Dumoulin, S.O., Brewer, A.A., 2007. Visual field maps in human cortex. Neuron 56, 366–383.
Wang, S., Huang, L., Chen, Q., Wang, J., Xu, S., Zhang, X., 2021. Awareness-dependent normalization framework of visual bottom-up attention. J. Neurosci. 41, 9593–9607.
White, B.J., Berg, D.J., Kan, J.Y., Marino, R.A., Itti, L., Munoz, D.P. , 2017a. Superior colliculus neurons encode a visual saliency map during free viewing of natural dynamic video. Nat. Commun. 8, 1–9.
White, B.J., Kan, J.Y., Levy, R., Itti, L., Munoz, D.P. , 2017b. Superior colliculus encodes visual saliency before the primary visual cortex. Proc. Natl. Acad. Sci. USA. 114, 9451–9456.
Wolfe, J.M., 1994. Visual search in continuous, naturalistic stimuli. Vision Res. 34, 1187–1195.
Yan, Y., Zhaoping, L., Li, W., 2018. Bottom-up saliency and top-down learning in the primary visual cortex of monkeys. Proc. Natl. Acad. Sci. USA. 115, 10499–10504.
Zhaoping, L., 2008. Attention capture by eye of origin singletons even without awareness—A hallmark of a bottom-up saliency map in the primary visual cortex. J. Vis. 8, 1–18.
Zhaoping, L., May, K.A., 2007. Psychophysical tests of the hypothesis of a bottom-up saliency map in primary visual cortex. PLoS Comput. Biol. 3, e62.
Zhaoping, L., Zhe, L., 2015. Primary visual cortex as a saliency map: a parameter-free prediction its test by behavioral data. PLoS Comput. Biol. 11, e1004375.
Zhang, X., Fang, F., 2012. Object-based attention guided by an invisible object. Exp. Brain Res. 223, 397–404.
Zhang, X., Japee, S., Safiullah, Z., Mlynaryk, N., Ungerleider, L.G., 2016. A normalization framework for emotional attention. PLoS Biol 14, e1002578.
Zhang, X., Mlynaryk, N., Ahmed, S., Japee, S., Ungerleider, L.G., 2018. The role of inferior frontal junction in controlling the spatially global effect of feature-based attention in human visual areas. PLoS Biol. 16, e2005399.
Zhang, X., Qiu, J., Zhang, Y., Han, S., Fang, F., 2014. Misbinding of color and motion in human visual cortex. Curr. Biol. 24, 1354–1360.
Zhang, X., Zhaoping, L., Zhou, T., Fang, F., 2012. Neural activities in V1 create a bottom-up saliency map. Neuron 73, 183–192.