nature communications

Article

https://doi.org/10.1038/s41467-024-53968-x

Detecting biological motion signals in human and monkey superior colliculus: a subcortical-cortical pathway for biological motion perception

Received: 15 November 2023

Accepted: 25 October 2024

Published online: 07 November 2024

Check for updates

Xiqian Lu[1,2,6], Zhaoqi Hu[1,2,6], Yumeng Xin[3,4,6], Tianshu Yang[3,4,5], Ying Wang[1,2], Peng Zhang[3,4], Ning Liu[3,4,7] ✉ & Yi Jiang[1,2,7] ✉

Most vertebrates, including humans, are highly adept at detecting and encoding biological motion, even when it is portrayed by just a few point lights attached to the head and major joints. However, the function of subcortical regions in biological motion perception has been scarcely explored. Here, we investigate the role of the superior colliculus in local biological motion processing. Using high-field (3 T) and ultra-high-field (7 T) functional magnetic resonance imaging, we record the neural responses of the superior colliculus to scrambled point-light walkers (with local kinematics retained) in both humans and male macaque monkeys. Results show that the superior colliculus, especially the superficial layers, selectively responds to local biological motion. Furthermore, dynamic causal modeling analysis reveals a subcortical-cortical functional pathway that transmits local biological motion signals from the superior colliculus via the middle temporal visual complex to the posterior superior temporal sulcus in the human brain. These findings suggest the existence of a cross-species mechanism in the superior colliculus that facilitates the detection of local biological motion at the early stage of the visual processing stream.

Detecting the presence of living entities in the environment from their active movements is a fundamental capability of the human and other animal visual systems. One striking phenomenon is that humans can effectively perceive biological motion (BM) merely from a few point lights describing the moving trajectories of the major joints[1]. Many animal species can also discriminate a point-light BM from other motion patterns (e.g., chicks[2], cats[3], dogs[4], marmosets[5], chimpanzees[6], zebrafish[7], medaka fish[8]). In the last two decades, accumulated evidence points to the vital role of the kinematic cues carried by individual joints, also known as the local motion cues, in visual BM perception, independent of the processing of the global configuration cues[9–11]. Local BM processing is achieved within a fraction of a second in the human brain, even when the observers have no idea of what kind of creatures they see[12–14]. Moreover, local BM processing is supposed to be an inborn ability, given that the sensitivity to and preference for local BM cues appear in 2-day-old human infants[15,16], and individual variations in local BM perception abilities can be largely accounted for by genetic factors[17]. Several animal species also exhibit a preference for

[1]State Key Laboratory of Brain and Cognitive Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China. [2]Department of Psychology, University of Chinese Academy of Sciences, Beijing, China. [3]State Key Laboratory of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China. [4]College of Life Sciences, University of Chinese Academy of Sciences, Beijing, China. [5]Department of Radiology, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China. [6]These authors contributed equally: Xiqian Lu, Zhaoqi Hu, Yumeng Xin. [7]These authors jointly supervised this work: Ning Liu, Yi Jiang. ✉e-mail: liuning@ibp.ac.cn; yijiang@psych.ac.cn

local BM rather than the non-BM patterns (e.g., dogs[18], fish[19]), even from the very beginning of life (e.g., chicks[20]). Besides, this local BM processing is subject to a significant inversion effect, meaning that participants perform significantly better with upright BM stimuli compared to vertically flipped ones, despite the entire disruption of configural information in both[21]. Given the rapid and potentially inborn nature of local BM processing, researchers have hypothesized the existence of a cross-species mechanism for BM perception, potentially rooted in the primitive and homologous brain architecture[2,21–24]. However, the role of subcortical regions in BM perception remains largely unexplored.

In the present study, we focus on the superior colliculus (SC), an evolutionarily conserved noncortical structure lying on the roof of the mammalian midbrain, with its homologous structure known as the optic tectum in non-mammalian vertebrates[25]. In recent years, researchers have highlighted the assumption that the SC is involved in computing visual inputs of unlearned biologically relevant stimuli[26]; thus, it could be a candidate for the early detection of BM through its local motion cues[10,27]. However, to the best of our knowledge, no direct evidence has heretofore linked the function of the SC to visual BM perception in humans or animals. On the other hand, plenty of studies have located a broad cortical network responding to visual BM information in primates[11,28–31]. Within this network, the posterior superior temporal sulcus (pSTS) is a critical area for the orientation-dependent representation of BM in humans[32,33] and macaques[34,35]. Besides, the human middle temporal visual complex (MT +) is a vital region for analyzing the kinematic cues of BM[36]. If the SC is a subcortical neural substrate underlying local BM perception, how it functionally links with these cortical regions to form BM representations is also an important question.

Hence, the aims of our study are twofold. Firstly, we investigate whether the SC would selectively respond to local BM cues in both humans and animals (e.g., macaque monkeys) using high-field (3 T, for humans and monkeys) and ultra-high-field (7 T, for humans) functional magnetic resonance imaging (fMRI). Secondly, we explore the effective connectives in humans among the SC and two cortical regions crucial to local BM processing – the pSTS and the MT + , using dynamic causal modeling (DCM) to delineate the subcortical-cortical pathway for BM perception.

Results

SC selectively responds to local BM cues

To isolate the local BM cues, we derived natural scrambled BM stimuli by randomly relocating each dot of the intact point-light BM stimuli within the region they occupied (Fig. 1a)[17,21,22,37]. We focused on whether blood-oxygen-level-dependent (BOLD) signals in the SC were higher for the upright natural scrambled BM stimuli than their inverted counterparts that lacked critical life motion signals, i.e., the inversion effect. We also looked into the inversion effect for a type of unnatural BM stimuli, which were generated by forcing each dot of the natural stimuli to move along its original trajectory with constant speed[22,38], to further confirm the selectivity and specificity of the SC for local BM signals.

Figure 1c showed the BOLD signal time courses of the SC responses for upright and inverted natural scrambled BM for all 33 human observers and Fig. 1d showed the SC time courses for a subgroup of 17 observers who were tested with all four stimuli conditions. Note that because the patterns of the SC responses from 7 T and 3 T MRI scanning were similar (see Fig. S2 and Table S7 for details), they were combined to achieve a better signal-to-noise ratio. Paired t-tests or Wilcoxon signed-rank tests were employed according to the normality of the combined data (see Fig. S2 for details). For the natural scrambled BM stimuli, the SC exhibited a significantly stronger neural response for the upright condition than for the inverted condition in all observers (Fig. 1c; $t_{TR2} = 2.80$, $p_{uncorr} = 0.010$, $p_{FDR} < 0.05$, $BF_{10} = 8.38$; $Z_{TR3} = 2.83$, $p_{uncorr} = 0.005$, $p_{FDR} < 0.05$, $p_{Bonferroni} < 0.05$, $BF_{10} = 37.57$; $t_{TR4} = 1.93$, $p_{uncorr} = 0.067$, $BF_{10} = 1.76$; $t_{TR5} = 1.28$, $p_{uncorr} = 0.215$, $BF_{10} = 0.68$; $t_{TR6} = 2.15$, $p_{uncorr} = 0.043$, $BF_{10} = 2.54$) and in the subgroup (Fig. 1d; $t_{TR2} = 2.55$, $p_{uncorr} = 0.018$, $BF_{10} = 6.50$; $Z_{TR3} = 1.63$, $p_{uncorr} = 0.102$, $BF_{10} = 3.95$; $t_{TR4} = 2.48$, $p_{uncorr} = 0.021$, $BF_{10} = 5.65$; $t_{TR5} = 1.45$, $p_{uncorr} = 0.162$, $BF_{10} = 1.11$; $t_{TR6} = 1.65$, $p_{uncorr} = 0.113$, $BF_{10} = 1.49$). By contrast, for the unnatural scrambled BM, SC responses between the upright and inverted conditions showed no significant difference (Fig. 1d: $ps_{uncorr} > 0.3$, $BF_{10} < 0.48$). Analysis on the beta coefficients of SC activation yielded similar results. There was a significant or marginally significant inversion effect for the natural scrambled BM stimuli (Fig. 1c: $Z = 1.93$, $p = 0.054$, $BF_{10} = 2.85$; Fig. 1d: $Z = 1.97$, $p = 0.049$, $BF_{10} = 2.60$) but not for the unnatural scrambled BM stimuli (Fig. 1d: $Z = 1.07$, $p = 0.287$, $BF_{10} = 0.27$). Additionally, the SC responded significantly stronger to the upright natural scrambled BM than to the unnatural stimuli. Significant or marginally significant differences were revealed between the upright natural condition and the average of the upright and inverted unnatural conditions (Fig. 1d: $t_{TR2} = 2.65$, $p_{uncorr} = 0.015$, $BF_{10} = 7.83$; $t_{TR3} = 2.01$, $p_{uncorr} = 0.057$, $BF_{10} = 2.56$; $t_{TR4} = 2.62$, $p_{uncorr} = 0.016$, $BF_{10} = 7.34$; $t_{TR5} = 2.61$, $p_{uncorr} = 0.016$, $p_{FDR} < 0.05$, $BF_{10} = 7.27$; $t_{TR6} = 2.17$, $p_{uncorr} = 0.041$, $BF_{10} = 3.32$; for beta coefficients, $t = 1.90$, $p_{uncorr} = 0.075$, $BF_{10} = 2.05$). On the contrary, no significant differences were found between the inverted natural condition and the average of the unnatural conditions ($t_{TR2} = 0.43$, $p_{uncorr} = 0.67$, $BF_{10} = 0.36$; $t_{TR3} = 0.70$, $p_{uncorr} = 0.49$, $BF_{10} = 0.46$; $t_{TR4} = 1.33$, $p_{uncorr} = 0.20$, $BF_{10} = 0.94$; $t_{TR5} = 1.91$, $p_{uncorr} = 0.07$, $BF_{10} = 2.09$; $t_{TR6} = 1.51$, $p_{uncorr} = 0.15$, $BF_{10} = 1.18$; for beta coefficients, $t = 0.29$, $p_{uncorr} = 0.77$, $BF_{10} = 0.20$). Furthermore, an additional eye-tracking experiment in humans demonstrated that eye movement metrics did not significantly differ between conditions (see Fig. S5).

Consistent with the results obtained in humans, the SC of all three monkeys exhibited the strongest responses for the upright natural scrambled BM (Fig. 1e). More importantly, SC responses for the upright natural scrambled BM were significantly stronger than those for the inverted stimuli in all monkeys (Monkey Q: $t = 2.84$, $p = 0.005$; Monkey M: $t = 1.99$, $p = 0.047$; Monkey X: $t = 2.56$, $p = 0.011$). SC responses between the upright and inverted unnatural scrambled BM showed no significant difference in two monkeys (Monkey Q: $t = 0.31$, $p = 0.760$; Monkey M: $t = 0.34$, $p = 0.730$) and a weak difference in the other one (Monkey X: $t = 2.06$, $p = 0.039$). Additionally, SC responses for the upright natural scrambled BM were significantly stronger than the average responses for the upright and inverted unnatural conditions (Monkey Q: $t = 1.98$, $p = 0.047$; Monkey M: $t = 2.00$, $p = 0.045$; Monkey X: $t = 2.55$, $p = 0.011$), whereas there were no significant differences between the inverted natural condition and the average of the unnatural conditions (Monkey Q: $t = 0.03$, $p = 0.974$; Monkey M: $t = 0.59$, $p = 0.557$; Monkey X: $t = 0.73$, $p = 0.463$). Moreover, eye positions of the monkeys were monitored during the experiment, revealing no significant differences between conditions (see Table S4–S6).

Functional connectivities transmitting BM signals from SC to cortical regions

We further investigated the function of the SC in humans. We used dynamic causal modeling (DCM) analysis[39] to investigate the functional pathway between the SC and the cortical regions MT+ and pSTS. The regions of interest (ROIs) of MT+ and pSTS were separately defined using the data of the functional localizers.

We focused on the difference in modulatory influences on intrinsic connections between the upright scrambled BM and the inverted scrambled BM. For this purpose, we extracted the first eigenvariate of the BOLD time series from each ROI based on the individuals' peak activations induced by the upright natural scrambled BM > the inverted natural scrambled BM. Then, we systematically varied all intrinsic connections and modulations between regions to create reduced models, resulting in a total of 64 models for each

https://doi.org/10.1038/s41467-024-53968-x

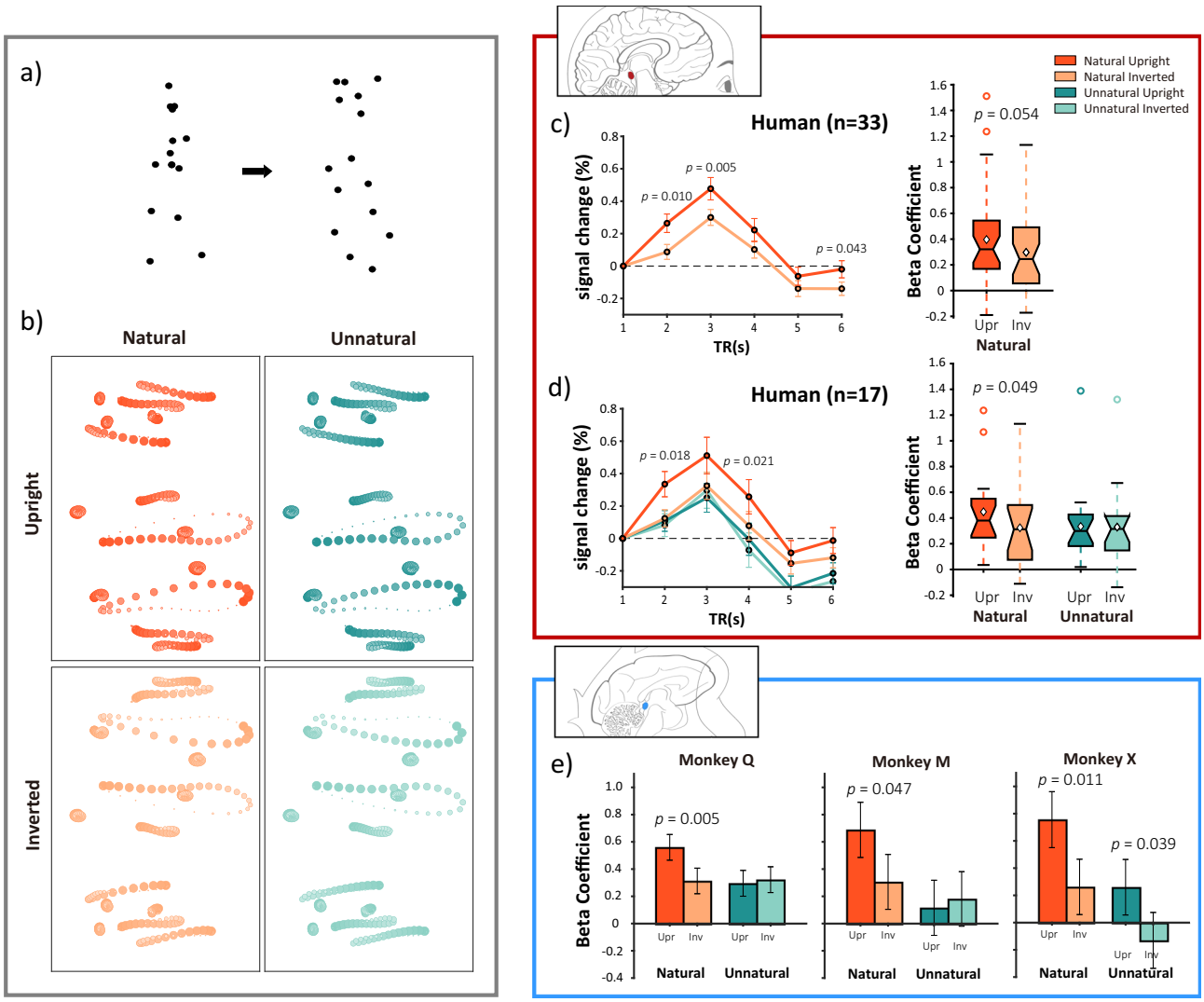

Fig. 1 | Schematic representations of the stimuli and responses of the whole superior colliculus (SC) in humans and monkeys. a Original point-light walker and an example of a spatially scrambled walker. **b** Schematic representations of the dots' trajectories from an example. The size and contrast of each dot decrease as they refer to more remote frames. For the natural scrambled biological motion (BM), the trajectory of each dot was extracted from a point-light walker and randomly relocated within the region occupied by the original walker. These trajectories contained a typical gravitational acceleration of BM. For the unnatural scrambled BM, each dot of natural scrambled BM was forced to move along its original trajectory with constant speed. Inverted walkers were conducted by vertically flipping the upright walkers. **c** The time courses and the beta coefficient of the SC for upright/inverted natural scrambled BM of all human observers. **d** The time courses and the beta coefficients of the SC for upright/inverted natural scrambled BM and upright/inverted unnatural scrambled BM of 17 human observers. **e** Beta coefficients of the SC of monkeys. The time courses for human and the beta coefficients for each monkey are presented as mean values ± SEM. For human beta coefficients, the center of the box indicates the median and the bottom/top edges of the box indicate the 25th/75th percentiles. The whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually. The diamond markers indicated the mean values. Two-sided paired t-tests or Wilcoxon signed-rank tests were used (see Results section for details). Source data are provided as a Source Data file.

observer. Bayesian model selection (BMS) analysis revealed that the full model with all intrinsic connections and modulations (Fig. 2a) had the highest posterior probability (close to 1; Fig. 2b). The difference in log evidence between the most likely model and the nearest alternative was larger than 232. We also conducted another DCM analysis with a model space, in which all the intrinsic connections were included in each model and only modulations were varied, and the full model still had the highest posterior probability (close to 1, $\Delta F > 32$). Paired t-tests over all 9 pairs of modulatory parameters revealed a functional pathway transmitting local BM signals from SC, via MT + , into pSTS. Significantly stronger modulatory strengths of the upright natural scrambled BM than those of the inverted natural scrambled BM were found in connectivity from SC to MT+ ($t = 2.73$, $p_{uncorr} = 0.010$, $p_{FDR} < 0.05$, $BF_{10} = 4.32$) as well as from MT+ to pSTS ($t = 2.55$, $p_{uncorr} = 0.016$, $p_{FDR} < 0.05$, $BF_{10} = 3.02$). Although the modulatory parameters in MT+ to SC connectivity showed differences between the upright and inverted conditions, the effect was quite weak ($t = 2.00$, $p_{uncorr} = 0.054$, $BF_{10} = 1.09$). The modulatory strengths of the upright and inverted natural scrambled BM also exhibited significant differences in the self-connections of SC and pSTS (SC: $t = 3.14$, $p_{uncorr} = 0.004$, $p_{FDR} < 0.05$, $BF_{10} = 10.71$; pSTS: $t = 2.91$, $p_{uncorr} = 0.006$, $p_{FDR} < 0.05$, $BF_{10} = 6.47$). Such differences were weaker on the self-connection of MT + , i.e., significant at an uncorrected level ($t = 2.32$, $p_{uncorr} = 0.027$, $BF_{10} = 1.93$). Other pairs showed no significant differences ($ps_{uncorr} > 0.20$, $BF_{10} < 0.40$). The driving inputs exhibited significant inversion effects for both SC ($t = 5.09$, $p < 0.001$, $BF_{10} = 1722$) and MT+ ($t = 6.06$, $p < 0.001$, $BF_{10} = 17776$). Additionally, one sample t-tests over all 18 modulatory parameters revealed significant positive modulatory strengths in connectivity from SC to MT+ ($t = 4.10$, $p_{uncorr} < 0.001$, $p_{FDR} < 0.01$, $BF_{10} = 115.19$) and from MT+ to SC ($t = 3.20$,

https://doi.org/10.1038/s41467-024-53968-x

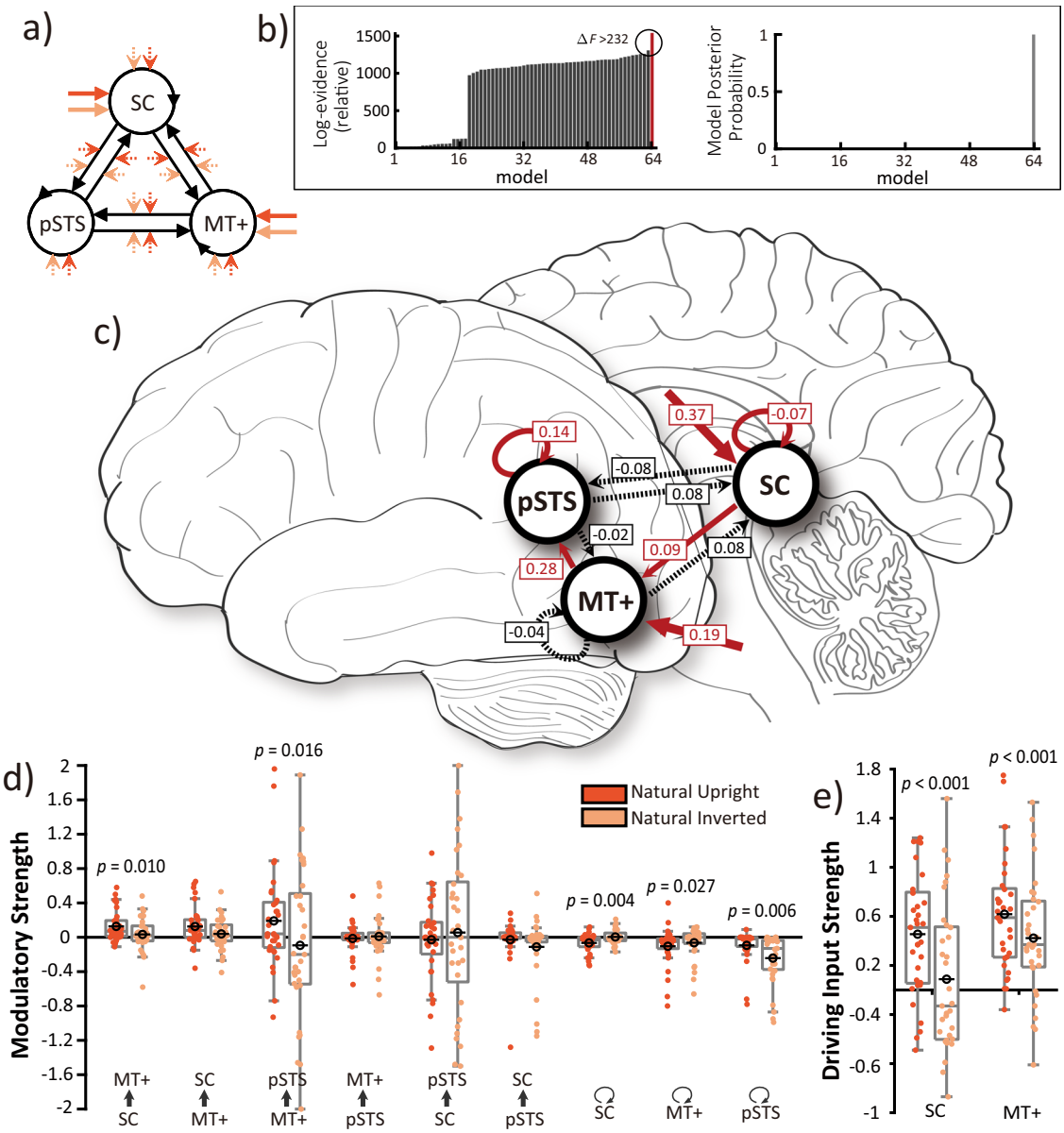

Fig. 2 | Dynamic causal modeling (DCM) results. a Depiction of the winning full model (Model 64). The solid arrows indicate driving inputs, and the dotted arrows indicate the modulatory effect. The red-orange arrows indicate the upright natural scrambled biological motion (BM) condition, and the light-orange arrows indicate the inverted scrambled BM condition. **b** The log-evidence and the posterior probability for all models. Model 64 (full model) had the highest probability of close to 1 and had $\Delta F > 232$ compared with the closest alternative model, representing robust evidence. **c** Depiction of the model and the results of paired t-tests over modulatory and driving input parameters. The values illustrate modulatory and driving input parameters for upright natural scrambled BM subtraction parameters for inverted natural scrambled BM. Red arrows indicate the significant (after FDR correction) inversion effect on modulatory and driving input parameters; that is, the modulatory parameters for upright natural scrambled BM are larger than inverted condition. A functional pathway for local BM is found from the superior colliculus (SC), via the middle temporal visual complex (MT +), into the posterior superior temporal sulcus (pSTS). Dashed black arrows indicate no significance was found after correction. **d** Mean modulatory parameters across observers (N = 33), indicating the strength of each connection or self-connection modulated by upright/inverted natural scrambled BM stimuli. **e** Mean driving input parameters across observers (N = 33). For each box plot, the center of the box indicates the median, and the bottom/top edges of the box indicate the 25th/75th percentiles. The whiskers extend to the most extreme data points not considered outliers. Individual data are plotted. The Saturn markers indicated the mean values. Two-sided paired t-tests were used. Source data are provided as a Source Data file.

$p_{uncorr} = 0.003$, $p_{FDR} < 0.05$, $BF_{10} = 12.33$) on the upright natural scrambled BM, significant negative modulatory strengths in self-connections of SC ($t = 3.31$, $p_{uncorr} = 0.002$, $p_{FDR} < 0.01$, $BF_{10} = 15.95$), MT+ ($t = 2.68$, $p_{uncorr} = 0.011$, $p_{FDR} < 0.05$, $BF_{10} = 3.95$) and pSTS ($t = 2.93$, $p_{uncorr} = 0.006$, $p_{FDR} < 0.05$, $BF_{10} = 6.65$) on the upright natural scrambled BM, and significant negative modulatory strengths in self-connections of pSTS on the inverted natural scrambled BM ($t = 4.74$, $p_{uncorr} < 0.001$, $p_{FDR} < 0.001$, $BF_{10} = 633.43$). No significant differences were found in other connections ($BF_{10} < 1.07$).

We also employed a DCM analysis focusing on the effect of the upright natural scrambled BM > the upright unnatural scrambled BM to verify whether the observed functional pathway was reliable. These two types of stimuli were strictly matched in all aspects except that the former conveyed critical local motion cues (i.e., accelerations). Results were consistent with those obtained from the main DCM analysis, suggesting the involvement of a functional pathway from SC, via MT + , into pSTS in local BM processing (see Fig. S3 for details).

https://doi.org/10.1038/s41467-024-53968-x

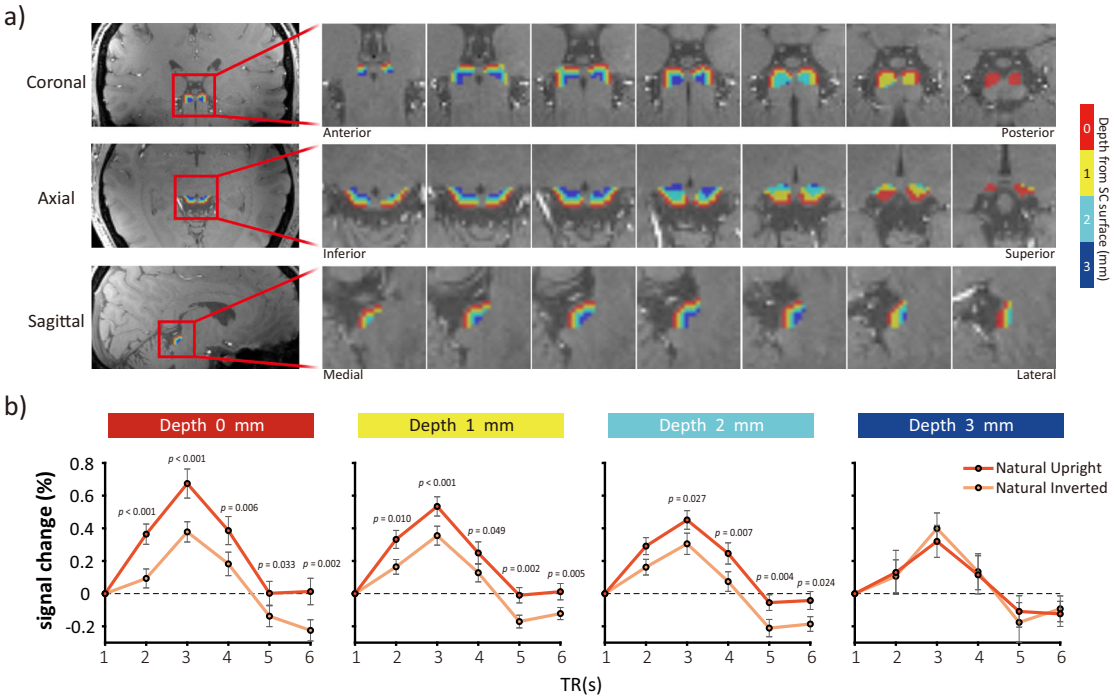

Fig. 3 | Results of human SC responses at different depths. a Regions of interest (ROIs) of different depths from the surface of the superior colliculus (SC) were defined from T1 anatomical images. Here are the SC ROIs with T1 images of a representative observer from 7 T MRI scanning. **b** Blood-oxygen-level-dependent (BOLD) response for upright/inverted natural scrambled BM at the different depths from the surface of the human SC (N = 33). The time courses are presented as mean values ± SEM. Two-sided paired t-tests or Wilcoxon signed-rank tests were used (see Results section for details). Source data are provided as a Source Data file.

Human SC responses at different depths

The SC is a structurally and functionally layered structure. In the primate, the superficial layers receive direct input from the retina and strongly respond to a variety of salient visual stimuli, while the deeper remaining layers are closely related to motor-related functions and respond to stimuli in multiple modalities[40,41]. Considering its divided function, we analyzed the SC's responses to the visual local BM signals at its different depths, in order to examine whether its responses were strongest at superficies and decayed with depth increasing. Therefore, we defined the surface and derived voxels at different depths of each human observer's SC (Fig. 3a)[42,43]. Figure 3b shows the time courses of BOLD signals in the SC at different depths from the surface for the upright/inverted natural scrambled BM in all human observers. Paired t-tests or Wilcoxon signed-rank tests were employed according to the normality (see Fig. S2 for details). As shown by Fig. 3b, significantly stronger responses for the upright than for the inverted local BM were established in voxels at the SC surface (Depth = 0 mm: $t_{TR2} = 3.861$, $p_{uncorr} < 0.001$, $p_{FDR} < 0.01$, $p_{Bonferroni} < 0.05$, $BF_{10} = 75.21$; $Z_{TR3} = 3.252$, $p_{uncorr} < 0.001$, $p_{FDR} < 0.01$, $p_{Bonferroni} < 0.05$, $BF_{10} = 75.17$; $t_{TR4} = 3.02$, $p_{uncorr} = 0.006$, $BF_{10} = 13.00$; $t_{TR5} = 2.28$, $p_{uncorr} = 0.033$, $p_{FDR} < 0.05$, $BF_{10} = 3.19$; $Z_{TR6} = 3.15$, $p_{uncorr} = 0.002$, $p_{FDR} < 0.01$, $p_{Bonferroni} < 0.05$, $BF_{10} = 43.55$) and those close to the surface (Depth = 1 mm: $t_{TR2} = 2.81$, $p_{uncorr} = 0.010$, $p_{FDR} < 0.05$, $BF_{10} = 8.52$; $t_{TR3} = 3.84$, $p_{uncorr} < 0.001$, $p_{FDR} < 0.01$, $p_{Bonferroni} < 0.05$, $BF_{10} = 71.07$; $t_{TR4} = 2.09$, $p_{uncorr} = 0.049$, $BF_{10} = 2.29$; $t_{TR5} = 3.48$, $p_{uncorr} = 0.002$, $p_{FDR} < 0.01$, $p_{Bonferroni} < 0.05$, $BF_{10} = 32.89$; $t_{TR6} = 3.10$, $p_{uncorr} = 0.005$, $p_{FDR} < 0.05$, $BF_{10} = 15.11$). Such difference in the deeper SC was weaker (Depth = 2 mm: $t_{TR2} = 1.87$, $p_{uncorr} = 0.074$, $BF_{10} = 1.62$; $t_{TR3} = 2.37$, $p_{uncorr} = 0.027$, $p_{FDR} < 0.05$, $BF_{10} = 3.74$; $t_{TR4} = 2.95$, $p_{uncorr} = 0.007$, $p_{FDR} < 0.05$, $BF_{10} = 11.14$; $t_{TR5} = 3.19$, $p_{uncorr} = 0.004$, $p_{FDR} < 0.05$, $BF_{10} = 18.36$; $Z_{TR6} = 2.26$, $p_{uncorr} = 0.024$, $p_{FDR} < 0.05$, $BF_{10} = 3.62$) and finally vanished (Depth = 3 mm: $ps_{uncorr} > 0.39$, $BF_{10} < 0.29$). Moreover, a 2 (Stimulus orientation) × 4 (Depth) repeated measures ANOVA on the peak points (TR3) revealed a significant interaction, $F(1.8,58.4) = 5.26$, $p = 0.010$, after Greenhouse-Geisser correction, $BF_{10} = 6.48$. The main effect of Stimulus orientation [$F(1,32) = 4.80$, $p = 0.036$, $BF_{10} = 37.05$] as well as Depth [$F(2.1,66.8) = 4.07$, $p = 0.020$, after Greenhouse-Geisser correction, $BF_{10} = 3.50$] also reached significance. For the interaction, simple effect analysis revealed that the neural responses exhibited significant differences at different depths only in the upright BM condition [$F(3, 96) = 8.18$, $p < 0.001$, $BF_{10} = 337.52$] but not in the inverted BM condition [$F(3, 96) = 0.66$, $p = 0.576$, $BF_{10} = 0.09$], indicating that only the SC's responses to upright local BM signals had a decreased tendency with increasing depths of voxels from the surface.

Discussion

The current study provides substantial evidence that the SC plays a vital role in local BM processing through high-field (3 T) and ultra-high-field (7 T) fMRI in both humans and non-human primates. Firstly, we recorded the neural responses of the SC in both humans and macaque monkeys and demonstrated that the SC, especially the superficial layers, selectively responds to local BM cues. In particular, the SC exhibited a stronger response to the upright scrambled BM stimuli conveying natural local BM cues than to their inverted counterparts, while showing no such orientation-dependent activation when local BM cues were removed from the visual stimuli. Secondly, by using DCM, we found a functional pathway in the human brain that transmitted local BM information from SC, via MT + , to pSTS, with stronger modulatory strengths for the upright scrambled BM stimuli than the inverted controlled stimuli in the feedforward connections.

Our findings suggest the existence of a cross-species mechanism in the SC that facilitates the detection of BM signals at the early stage of the visual processing stream. Researchers have previously proposed a link between the SC and life motion detection. Troje and Westhoff[21] found a significant inversion effect in scrambled BM perception and interpreted their findings in terms of a visual filter tuned to the characteristic local motion of animals. Such a visual filter possibly serves as an evolutionarily old life detection mechanism to support the innate

sensitivity to the local BM cues for neonates and newly hatched chicks at their first exposure to scrambled point-light displays. Troje and Chang[27] further suggested that the SC, a homologous component of the vertebrate midbrain (also known as the optic tectum in non-mammalians), is the neural structure for life motion detection in terrestrial, legged animals. The present study provides direct evidence for this hypothesis by demonstrating that both the human and non-human primate SC selectively responded to local BM cues, which encourages further exploration of the function of the SC on BM perception in different vertebrates.

Notably, the SC is also involved in a subcortical route for the detection of another important social stimulus, i.e., faces[44]. A recent study in patients with unilateral primary visual cortex damage showed that part of patients could detect and discriminate faces presenting inside their blind visual field, with the fMRI signals exhibiting a functional connection between the SC and the amygdala in their damaged hemisphere[45]. Neurophysiological studies in monkeys also suggest that the SC exhibits distinct responsive patterns between face-like and nonface stimuli as early as 50 ms after the stimulus onset[46,47]. The current finding regarding the role of the SC in detecting BM signals, combined with the evidence of the SC's function in face detection, supports the idea that the SC probably comprises an essential part of a general animacy-detection system[26].

Besides, considering that the SC is a structurally and functionally layered structure, we also investigated its response to local BM signals at different depths in the human SC. The inversion effect for natural scrambled BM was pronounced at the voxels in the superficial layer (which also respond to salient visual stimuli) of the SC, decreased in the intermediate layers, and diminished in the deep layers, suggesting a functional divergence for different SC layers in which the superficial layers are crucially involved in BM detection. This finding is consistent with the structure and function of the primate SC, where the superficial layer of the SC receives direct input from the retina and then transmits it to subsequent structures in the visual processing stream[41].

Plenty of human neuroimaging studies have identified a broad cortical network involved in BM perception, mainly consisting of the pSTS, MT + , posterior inferior temporal sulcus, fusiform gyrus, and extrastriate body areas[11,28,29,31]. However, only a few studies have reported noncortical areas responding to BM in humans and animals[48], including the amygdala[49], the left ventral lateral nucleus of the thalamus[22], and the cerebellum[50,51]. Among these studies, Chang et al.[22] explored the subcortical neural activity associated with local BM processing, demonstrating that the left ventral lateral nucleus could discriminate between natural and unnatural local BM signals. Our study extends this finding by revealing the crucial involvement of the SC in local BM processing.

More interestingly, we found functional connectivities transmitting local BM signals from the subcortical region SC to the cortical regions MT+ and pSTS in humans, as revealed by stronger modulatory strengths for the upright natural scrambled BM than the inverted or unnatural scrambled BM stimuli in the connection from SC to MT+ and that from MT+ to pSTS. The existence of the pathway from SC to MT or STS has been supported by previous studies in monkeys. The pathway through the pulvinar from the superficial visual layers of the SC to the MT was identified in macaque by microstimulation[52,53]. The floor of the STS has been identified as the primary cortical target of the SC activity in macaques[54]. The functional connectivity between the SC and the cortical regions may constitute a subcortical-cortical pathway for BM processing, transmitting the rapidly detected life motion signal to the cortical network responsible for more comprehensive analyzes of the BM signals. Our findings are also compatible with a recently proposed two-process theory of BM processing[10]. This theory divides BM processing into a rapid and pre-attentive step detector stage, which relies on a subcortical network, and a slower bodily action evaluator stage, which relies on cortical networks and receives information from the first stage. It should be noted that the functional pathway for BM perception is largely inferred from BOLD signals, and future studies using more direct methods to capture neurotransmission are needed to corroborate these findings.

The current study points to the pivotal role of the SC in local BM detection, while the mechanism of how the SC detects local BM cues remains to be elucidated. Here we propose some speculations alongside the typical characteristics of BM perception in prior research. Previous investigators highlighted the unique role of foot movements in local BM processing[14,21,38,55]. Specifically, for terrestrial animals, the ballistic trajectories created by their feet pushing off the ground and falling provide clues for identifying living organisms[21]. Based on this perspective, it is conceivable that the SC may also be more sensitive to the foot movements, but the present study cannot isolate the specific contribution of foot movements in the SC response since the scrambled BM stimuli retain all major joint trajectories of the body. On the other hand, a recent study reported that fish also exhibit a preference for upright BM over inverted or constant-speed stimuli despite lacking feet[7,19]. Another widely accepted interpretation for the inversion effect of BM perception is a gravity-dependent model, as the limb motion trajectories are constrained by gravity and the vertical acceleration contained in the local BM might be a cue for the visual system to identify life motion in the visual environment[21,38,56]. Further studies on the SC response to visual gravity cues may help illustrate whether the selective response in the SC for upright BM is related to the process of visual gravity cues.

Additionally, it should be noted that the behavioral results (Fig. S1) did not entirely correspond to the neural responses of the SC observed in this study. Specifically, the behavioral task was to discriminate the locomotion directions of the stimuli. Results indicated inversion effects not only under natural conditions but also under unnatural conditions, whereas only natural conditions exhibited significant inversion effects in the neural responses of the SC. The unnatural scrambled BM stimuli, characterized by a constant speed, were originally developed by Chang and Troje[38]. In their recent neuroimaging studies, behavioral results for such unnatural stimuli also showed a significant inversion effect[12,22]. Since our task gave the observers sufficient processing and response time (i.e., 2 s for stimulus presenting and 10 s for response), the difference between the behavioral results and the neural responses of the SC might reflect the fact that the behavioral outcomes captured an integrated processing involving multiple brain regions.

Methods

Human observers

A total of 35 adult observers (24.0 ± 2.8 years old; 23 females) were paid to participate in this study. All observers were right-handed, had normal or corrected-to-normal vision, and had no history of neurological or psychiatric disorders. All observers volunteered for this study and were paid. Written informed consent was obtained from all observers before the experiment according to procedures approved by the Institutional Review Board of the Institute of Psychology, Chinese Academy of Sciences. The gender of observers was determined based on self-report. We did not perform gender-based analyzes as gender was not relevant to the questions and hypotheses of the study.

Animals

Three male macaque monkeys (monkeys Q, M, and X; *Macaca mulatta*; 10–11 years old; 7.0 – 10.5 kg) were used. They were acquired from a primate breeding facility in China, where they had social group histories and a group-housing experience until their transfer to the Institute of Biophysics, Chinese Academy of Sciences (IBP) for quarantine at the age of approximately four years. Both animals used in this study had been housed at IBP for 4–7 years before this experiment. All experimental procedures complied with the US National Institutes of

Health Guide for the Care and Use of Laboratory Animals and were approved by the Institutional Animal Care and Use Committee of IBP. Each monkey was surgically implanted with a magnetic resonance (MR)-compatible head post under sterile conditions, using isoflurane anesthesia. After recovery, monkeys were trained to sit in a plastic restraint chair and fixate at a central target for long durations with their heads fixed, facing a screen on which visual stimuli were presented[57,58].

Stimuli

Point-light walker sequences were used as stimuli[59]. Each walker was represented by a set of 15 dots depicting the major joints of the body. We conducted eight views of walker for human observers, which changed from facing rightward to leftward by step of 25.71°. For monkeys, only the rightward and leftward walkers were used. The stimuli were presented as white dots against a gray background, lasting 1 s in each walking cycle. Each stimulus was located at the screen center with a jitter of 0.3° and started from a random frame. We derived scrambled BM by randomly relocating each dot within the region occupied by the original walker[37]. Four kinds of stimuli were used in formal experiments: Upright natural scrambled BM, inverted natural scrambled BM, upright unnatural scrambled BM, and inverted unnatural scrambled BM. Unnatural BM was generated by forcing each dot of natural BM to move along its original trajectory with constant speed[22,38]. This manipulation eliminates the vertical asymmetries caused by gravitational acceleration, destroying valid life motion signals embedded in BM. The inverted counterparts were conducted by vertically flipping the upright stimuli. Another two kinds of stimuli were added in the localizer session: Upright intact BM and inverted intact BM, which were the original walkers without relocating.

Human Experiments

Stimulus presentation and experimental manipulation were realized using MATLAB 2014b (MathWorks, Inc.) and the PsychToolbox-3 extensions[60]. Observers lay in a supine position inside the scanner and viewed the display through a mirror. In each trial, a BM stimulus was presented for 2 s, following a fixation displayed for 10, 12, 14, or 16 s. Stimuli spanned a visual angle of 11.7° (maximal horizontal and/or vertical extent) on a uniform gray background and were presented foveally with a fixation cross (0.6° in white) superimposed on each image. Observers were required to judge the locomotion direction (either toward left or right) of the undefinable creatures by pressing one of two keys with their right hand after the stimulus disappeared. Seventeen observers (12 in a 3 T scanner and 5 in a 7 T scanner, see below) were tested with two stimulus conditions: Upright natural scrambled BM and inverted natural scrambled BM. One of them (male) was discarded from further analysis due to his chance-level performance (0.5) on upright natural scrambled BM condition. The other 18 observers (12 in 3 T scanner and 6 in 7 T scanner) were tested with four stimulus conditions: Upright natural scrambled BM, inverted natural scrambled BM, upright unnatural scrambled BM, and inverted unnatural scrambled BM. There were 32 trials for each condition with each walking direction (left or right) 16 trials, dividing into several runs with 16 trials in each run. The first 3 volumes of each run were discarded to eliminate the effects of startup transients. Each run lasted 4 min and 8 s. Before the formal experiment, observers completed a practice session (at least 16 trials) outside the scanner to become familiar with the task.

Localizer runs were arranged in a block design. Each run included four main blocks comprising four stimulus conditions (i.e., upright natural scrambled BM, inverted natural scrambled BM, upright intact BM, and inverted intact BM). Blocks of each stimulus condition lasted 12 s, which consisted of 6 different examples of BM and were repeated 3 times within a run. Block order was randomized and was interleaved by 6 s fixation blocks. Each observer was tested with 2 runs, each lasting 3 min and 42 s.

Monkey Experiments

Stimulus presentation and experimental manipulation were realized in MATLAB 2014b with the PsychToolbox-3 extensions. In each trial, a BM stimulus was presented for 2 s, followed by a fixation presented for 12, 14, 16, or 18 s. The inter-trial interval (ITI) average duration was 15 s, considering the prolonged response function for MION in monkeys compared to BOLD in humans, where the average ITI was 13 s. Stimuli spanned a visual angle of 10.3° (maximal horizontal and/or vertical extent) on a uniform gray background and were presented foveally with a fixation square (0.2° in red) superimposed on each image. Each stimulus condition was presented 12 times in one run (i.e., 24 events in total), which lasted 6 min and 56 s. In one session, we firstly collected 6 runs for natural conditions in which upright and inverted natural scrambled BM were presented, then collected 6 runs for unnatural conditions in which upright and inverted unnatural scrambled BM were presented, or vice versa.

Eye position was monitored with an infrared pupil tracking system (ISCAN, Inc). The monkeys were required to maintain fixation on a square superimposed on the stimuli to receive a liquid reward. In the reward schedule, the reward frequency increased as the duration of fixation increased[57,58]. Each monkey was scanned in 2 sessions except for Monkey M (due to health concerns). Monkeys were required to fixate within a circular window (radius 2° of visual angle) centered over the fixation dot. Only data in which fixation was maintained at least 70% of the run time were included in the final analyzes, resulting in a total of 12–20 runs used in the final analyzes (Monkey Q: 20 runs, Monkey X: 16 runs, Monkey M: 12 runs), with an equal number of runs for natural and unnatural conditions. We did not find significant differences in the percentage of fixation, saccade size, and saccade frequency across conditions (Table S4–S6).

Human Data Acquisition

Imaging data for 24 observers were acquired at the Beijing MRI Center for Brain Research, using a 3 T Siemens Prisma MRI scanner (Siemens Healthcare, Erlangen, Germany) for both experimental and localizer runs. Blood oxygenated level-dependent signals were measured by echo-planar imaging sequences, with acquisition parameters as follows: repetition time (TR) = 2 s; echo time (TE) = 30 ms; flip angle = 90°. We acquired 26 axial slices (thickness = 1.5 mm, in-plane resolution = 128 × 128) covering the regions of SC, MT + , and pSTS but excluding most frontal and parietal regions as well as the cerebellum. We prioritized that the whole SC was perfectly covered for each observer, then we ensured the MT+ and pSTS were covered as intact as possible. In addition, T1-weighted sagittal images were collected as a high-resolution (1 mm^3) anatomical scan.

Imaging data for 11 observers were acquired at the Beijing MRI Center for Brain Research, using a whole-body human 7 T MR research system (Siemens Healthcare, Erlangen, Germany) for both experimental and localizer runs. Echo-planar imaging sequences were used. In 5 observers, 22 axial slices (thickness = 1.2 mm, in-plane resolution = 134 × 134) were acquired, with acquisition parameters as follows: TR = 2 s; TE = 24.4 ms; flip angle = 70°. In 6 observers, 44 axial slices (with multiband, factor = 2; thickness = 1.2 mm, in-plane resolution = 134 × 134) were acquired, with acquisition parameters as follows: TR = 2 s; TE = 25.6 ms; flip angle = 80°. The slices covered the regions of SC, MT + , and pSTS but excluded most frontal and parietal regions and the cerebellum. We prioritized that the whole SC was perfectly covered for each observer, then we ensured the MT+ and pSTS were covered as intact as possible. In addition, T1-weighted sagittal images were collected as a high-resolution (0.7 mm^3) anatomical scan.

Monkeys Data Acquisition

Imaging data for monkeys were acquired at the Beijing MRI Center for Brain Research, using a 3 T Siemens Prisma MRI scanner (Siemens Healthcare, Erlangen, Germany) with a surface coil array (eight

elements). Before each scanning session, an exogenous contrast agent [monocrystalline iron oxide nanocolloid (MION)] was injected into the femoral or external saphenous vein (8 mg/kg) to increase the contrast/noise ratio and to optimize the localization of fMRI signals[61]. We acquired 39 coronal slices (thickness = 1.5 mm, no gap, in-plane resolution = 129 × 129) using single-shot interleaved gradient-recalled echo-planar imaging. Acquisition parameters were as follows: voxel size: 1.5 mm isotropic; TR = 2 s; TE = 17 ms; flip angle = 90°; matrix size: 86 × 86.

A low-resolution T2 anatomical scan was also acquired in each session to serve as an anatomical reference (voxel size = 0.625 mm × 0.625 mm × 1.5 mm; TR = 11.2 s; TE = 101 ms; flip angle = 126°). To facilitate cortical surface alignment and the following local targeting, we also acquired high-resolution T1-weighted whole-brain anatomical scans in separate sessions under anesthesia (voxel size = 0.5 mm isotropic; TR = 2.2 s; TE = 2.84 ms; flip angle = 8°).

Human Data Analysis

fMRI data preprocessing and analysis were performed using MATLAB 2020a and SPM12 (Wellcome Center for Human Neuroimaging, London, United Kingdom). Functional images were corrected for head motion using a rigid body correction and corrected for slice acquisition time differences. Then, for extracting the SC signals, the functional images were resampled to 1 mm isotropic and co-registered with the anatomical images. The ROIs of SC were defined as their anatomical structure on individual T1 images with the drawing tool in MRIcron (https://people.cas.sc.edu/rorden/mricron). We first defined the surface voxels of the SC by drawing in the anatomical structure, whose depth was 0 mm. Then, the depth for each voxel in the SC was calculated as the shortest distance from the voxel to all surface voxels (Fig. 3a)[42,43]. Considering the individual difference in the size of the SC, we only analyzed the data from the first four layers of the SC for each observer because all the observers' SC could be divided into at least four layers of voxels. The average voxels' number (Mean ± SD) of each layer of the SC is as follows: 172.8 ± 44.9 for depth = 0 mm, 108.1 ± 22.9 for depth = 1 mm, 77.2 ± 23.5 for depth = 2 mm, 38.6 ± 21.3 for depth = 3 mm (see Table S1 for details). We used the MarsBar toolbox (http://marsbar.sourceforge.net) for SPM to extract time courses from the ROIs of SC. Time courses were extracted from the whole SC as well as each depth layer, respectively. These time courses were then converted to the percentage signal change by subtracting the mean of each run and then dividing by that value. We shifted the first point of each trial to zero and then averaged the signal intensity across trials for each condition at each of 6 time points (from 0 s to 12 s). One observer (male) was excluded due to the lack of reasonable BOLD responses in all four conditions.

For the other analyzes, anatomical images were co-registered with the functional images after motion and time correction, and iteratively segmented and normalized to the SPM MNI152 template. The resulting normalization parameters were applied to the functional images. Normalized images were smoothed using a Gaussian kernel of full-width half-maximum 4 mm. We defined the other two ROIs through the functional images in localizer runs. We created a first-level general linear model for localizer runs of each observer in SPM. Four regressors were defined with a duration of 12 s for upright intact BM, inverted intact BM, upright scrambled BM, and inverted scrambled BM, respectively. All regressors were convolved with the canonical hemodynamic response function. In addition, six realignment regressors (three translation parameters and three rotation parameters) were also included. Two t-contrasts were defined: motion, all the four conditions > baseline; BM-specific, upright intact BM > inverted scrambled BM. For each observer, the ROI of MT+ was defined as the suprathreshold ($p_{uncorr} < 0.001$) voxels in motion contrast, and the ROI of pSTS defined as the suprathreshold ($p_{uncorr} < 0.05$) voxels in BM-specific contrast, conforming to appropriate anatomy. The mean coordinates (MNI space) of the MT+ and pSTS ROI among observers are as follows: [−48.9 −64.3 6.9] for left MT + , [49.9 −64.8 6.3] for right MT + , [−56.1 −44.6 8.0] for left pSTS, [52.7 −43.6 8.3] for right pSTS. See Fig. S4 and Table S2 for details.

For the DCM analysis, we firstly created a general linear model for each observer in SPM. Events of each condition were modeled respectively (two or four conditions for different groups) with a duration of 0 s (assuming the events as impulse responses) and convolved with the canonical hemodynamic response function. In addition, six realignment regressors (three translation parameters and three rotation parameters) were also included. A contrast that upright natural scrambled BM > inverted natural scrambled BM was defined. Then, we extracted the first eigenvariate of the BOLD time series based on the individuals' peak activations induced by upright natural scrambled BM > inverted natural scrambled BM. For the SC, we found the observer-specific peak activation within individuals' ROI of SC and extracted the BOLD time series from the voxels both in a 3 mm radius sphere at the peak and within the ROI of SC. For the MT+ and pSTS, we found the observer-specific peak activations within the individual's ROI of MT+ or pSTS and extracted the BOLD time series from the voxels both in a 5 mm radius sphere at the peak and suprathreshold ($p_{uncorr} < 0.05$; see Table S3 for details). All time series were adjusted for confounding effects.

We focus on the difference of modulatory influences on intrinsic connections between upright scrambled BM and inverted scrambled BM. We included two conditions: Upright natural scrambled BM and inverted natural scrambled BM as both driving input and modulatory input. We applied driving inputs to SC and MT+ in all models. The full model included all the intrinsic connections between and within the three regions and the upright/inverted natural scrambled BM modulations of all connections. Then, we systematically varied all intrinsic connections and modulations between regions to create reduced models. To lessen the number of models in the model space, we made the intrinsic connections and the modulations change together all along. That is, if one model had an intrinsic connection between two regions, it must contain the modulations of this connection. We did not change the modulations of connections within regions. There was a total of 64 models for each observer. We used DCM in SPM12 to estimate these models. Then, the BMS with fixed effects was used to pool evidence at the group level and identify the best models[62]. The parameters under the winning models of each observer were extracted and imported into SPSS for further analyzes, which involved performing paired t-tests (with FDR correction for multiple comparisons) to all the modulatory effects and driving input effects.

We also employed another DCM analysis focusing on the effect of upright natural scrambled BM > upright unnatural scrambled BM. To do this, we extracted the BOLD time series based on peak activations induced by upright natural scrambled BM > upright unnatural scrambled BM from the observers who were tested by four conditions. The other aspects were consistent.

Statistical analyzes were conducted using SPSS Statistics 19.0 (IBM Corp., Armonk, United States of America). Bayes factors were added for all the statistical analyzes and conducted using jamovi (2.3.28) software (https://www.jamovi.org) with the Bayesian methods module (JASP Team, https://jasp-stats.org). We reported the BF_{10} for the t-test and the BF_{10}-inclusion value for each main effect and interaction effect of ANOVA. The hypothesis of Bayes paired t-test was group 1 > group 2 (i.e., upright > inverted). A $BF_{10} > 3$ indicates moderate evidence for the presence of an effect under consideration, and a $BF_{10} < 1/3$ indicates moderate evidence for the absence of the effect under consideration.

Monkeys Data Analysis

fMRI data preprocessing and analysis were performed using Analysis of Functional NeuroImages (AFNI) software (21.0.04)[63]. Images were

realigned to the base volume. Then, the data were smoothed with a 2-mm full-width half-maximum Gaussian kernel. Signal intensity was normalized to the mean signal value within each run. The anatomical ROIs of SCs were defined as their anatomical structure on individual high-resolution T1-weighted images with the drawing tool in AFNI. All visually responsive voxels that were more active during stimulus trial than during fixation for at least one category (Monkey Q: $p < 0.05$; Monkey M: $p < 0.15$; Monkey X: $p < 0.05$) within the anatomical ROI of SC were combined to yield the final SC ROI. Due to health issues, we were only able to collect one session from Monkey M, whereas two sessions were obtained from the other two monkeys. The relatively smaller size of the dataset reduced the statistical power of the analysis in Monkey M. Furthermore, in the present study, a surface coil was employed, which renders the signal-to-noise ratio (SNR) sensitive to the distance between the region of interest and the coil. It should be noted that Monkey M had the highest body weight among the three animals (Monkey Q: 7.0 kg, Monkey X: 9.0 kg, Monkey M: 10.5 kg). As a consequence, the muscle thickness in Monkey M was greater in comparison to the other two subjects, leading to weaker signals in the SC. Therefore, a relatively low threshold was applied in the analysis of Monkey M. We extracted the signal from the SC and then performed a single univariate linear model fit to estimate and compare the response amplitude for each condition. The model included a hemodynamic response predictor for each category and regressors of no interest (baseline, movement parameters from realignment corrections, and signal drifts). A general linear model and a MION kernel were used to model the hemodynamic response function[61]. All fMRI signals throughout the article have been inverted so that an increase in signal intensity indicates an increase in activation.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All data supporting the findings of this study are available within the paper and its Supplementary Information. The data generated in this study have been deposited in the Science Data Bank database under accession code https://doi.org/10.57760/sciencedb.13770. Source data are provided with this paper.

Code availability

Custom code can be found on the Science Data Bank: https://doi.org/10.57760/sciencedb.13770.

References

1. Johansson, G. Visual perception of biological motion and a model for its analysis. *Percept. Psychophys.* **14**, 201–211 (1973).
2. Vallortigara, G. & Regolin, L. Gravity bias in the interpretation of biological motion by inexperienced chicks. *Curr. Biol.* **16**, R279–R280 (2006).
3. Blake, R. Cats Perceive Biological Motion. *Psychol. Sci.* **4**, 54–57 (1993).
4. Kovács, K. et al. The effect of oxytocin on biological motion perception in dogs (Canis familiaris). *Anim. Cogn.* **19**, 513–522 (2016).
5. Brown, J., Kaplan, G., Rogers, L. J. & Vallortigara, G. Perception of biological motion in common marmosets (Callithrix jacchus): by females only. *Anim. Cogn.* **13**, 555–564 (2010).
6. Tomonaga, M. Visual search for biological motion patterns in chimpanzees (Pan troglodytes). *Psychol. Int. J. Psychol. Orient* **44**, 46–59 (2001).
7. Ma, X. et al. Gravity-dependent animacy perception in zebrafish. *Research* **2022**, 9829016 (2022).
8. Nakayasu, T. & Watanabe, E. Biological motion stimuli are attractive to medaka fish. *Anim. Cogn.* **17**, 559–575 (2014).
9. Giese, M. A. & Poggio, T. Neural mechanisms for the recognition of biological movements. *Nat. Rev. Neurosci.* **4**, 179–192 (2003).
10. Hirai, M. & Senju, A. The two-process theory of biological motion processing. *Neurosci. Biobehav. Rev.* **111**, 114–124 (2020).
11. Thompson, J. & Parasuraman, R. Attention, biological motion, and action recognition. *NeuroImage* **59**, 4–13 (2012).
12. Chang, D. H. F., Troje, N. F., Ikegaya, Y., Fujita, I. & Ban, H. Spatio-temporal dynamics of responses to biological motion in the human brain. *Cortex* **136**, 124–139 (2021).
13. Saunders, D. R., Suchan, J. & Troje, N. F. Off on the wrong foot: local features in biological motion. *Perception* **38**, 522–532 (2009).
14. Wang, L., Yang, X., Shi, J. & Jiang, Y. The feet have it: local biological motion cues trigger reflexive attentional orienting in the brain. *NeuroImage* **84**, 217–224 (2014).
15. Bardi, L., Regolin, L. & Simion, F. Biological motion preference in humans at birth: role of dynamic and configural properties: biological motion preference. *Dev. Sci.* **14**, 353–359 (2011).
16. Bardi, L., Regolin, L. & Simion, F. The first time ever i saw your feet: inversion effect in newborns' sensitivity to biological motion. *Dev. Psychol.* **50**, 986–993 (2014).
17. Wang, Y. et al. Heritable aspects of biological motion perception and its covariation with autistic traits. *Proc. Natl Acad. Sci.* **115**, 1937–1942 (2018).
18. Eatherington, C. J., Marinelli, L., Lõoke, M., Battaglini, L. & Mongillo, P. Local dot motion, not global configuration, determines dogs' preference for point-light displays. *Animals* **9**, 661 (2019).
19. Shibai, A. et al. Attraction of posture and motion-trajectory elements of conspecific biological motion in medaka fish. *Sci. Rep.* **8**, 8589 (2018).
20. Vallortigara, G., Regolin, L. & Marconato, F. Visually inexperienced chicks exhibit spontaneous preference for biological motion patterns. *PLoS Biol.* **3**, e208 (2005).
21. Troje, N. F. & Westhoff, C. The inversion effect in biological motion perception: evidence for a "life detector"? *Curr. Biol.* **16**, 821–824 (2006).
22. Chang, D. H. F., Ban, H., Ikegaya, Y., Fujita, I. & Troje, N. F. Cortical and subcortical responses to biological motion. *NeuroImage* **174**, 87–96 (2018).
23. Johnson, M. H. Biological motion: a perceptual life detector? *Curr. Biol.* **16**, R376–R377 (2006).
24. Troje, N. F. & Chang, D. H. F. Life detection from biological motion. *Curr. Dir. Psychol. Sci.* **32**, 26–32 (2023).
25. Stein, B. E., Stanford, T. R., Godwin, D. W. & McHaffie, J. G. *The Superior Colliculus and Visual Thalamus*. in *Neuroscience in the 21st Century* (ed. Pfaff, D. W.) 629–653 (Springer New York, New York, NY, 2013).
26. Rosa Salva, O., Mayer, U. & Vallortigara, G. Roots of a social brain: developmental models of emerging animacy-detection mechanisms. *Neurosci. Biobehav. Rev.* **50**, 150–168 (2015).
27. Troje, N. F. & Chang, D. H. F. *Shape-Independent Processing of Biological Motion*. in *People Watching* (eds Johnson, K. L. & Shiffrar, M.) 82–100 (Oxford University Press, 2013).
28. Blake, R. & Shiffrar, M. Perception of Human Motion. *Annu. Rev. Psychol.* **58**, 47–73 (2007).
29. Pitcher, D. & Ungerleider, L. G. Evidence for a third visual pathway specialized for social perception. *Trends Cogn. Sci.* **25**, 100–110 (2021).
30. Wang, R., Lu, X. & Jiang, Y. Distributed and hierarchical neural encoding of multidimensional biological motion attributes in the human brain. *Cereb. Cortex* **33**, 8510–8522 (2023).
31. Yovel, G. & O'Toole, A. J. Recognizing people in motion. *Trends Cogn. Sci.* **20**, 383–395 (2016).

32. Grossman, E. et al. Brain areas involved in perception of biological motion. *J. Cogn. Neurosci.* **12**, 711–720 (2000).
33. Grossman, E. & Blake, R. Brain activity evoked by inverted and imagined biological motion. *Vis. Res.* **41**, 1475–1482 (2001).
34. Jastorff, J., Popivanov, I. D., Vogels, R., Vanduffel, W. & Orban, G. A. Integration of shape and motion cues in biological motion processing in the monkey STS. *NeuroImage* **60**, 911–921 (2012).
35. Russ, B. E. & Leopold, D. A. Functional MRI mapping of dynamic visual features during natural viewing in the macaque. *NeuroImage* **109**, 84–94 (2015).
36. Jastorff, J. & Orban, G. A. Human functional magnetic resonance imaging reveals separation and integration of shape and motion cues in biological motion processing. *J. Neurosci.* **29**, 7315–7329 (2009).
37. Wang, L., Zhang, K., He, S. & Jiang, Y. Searching for life motion signals: visual search asymmetry in local but not global biological-motion processing. *Psychol. Sci.* **21**, 1083–1089 (2010).
38. Chang, D. H. F. & Troje, N. F. Acceleration carries the local inversion effect in biological motion perception. *J. Vis.* **9**, 19–19 (2009).
39. Friston, K. J., Harrison, L. & Penny, W. Dynamic causal modelling. *NeuroImage* **19**, 1273–1302 (2003).
40. Krauzlis, R. J., Lovejoy, L. P. & Zénon, A. Superior colliculus and visual spatial attention. *Annu. Rev. Neurosci.* **36**, 165–182 (2013).
41. May, P. J. et al. *The Mammalian Superior Colliculus: Laminar Structure And Connections*. in *Progress in Brain Research* 151 321–378 (Elsevier, 2006).
42. Zhang, P., Zhou, H., Wen, W. & He, S. Layer-specific response properties of the human lateral geniculate nucleus and superior colliculus. *NeuroImage* **111**, 159–166 (2015).
43. Zhang, P., Wen, W., Sun, X. & He, S. Selective reduction of fMRI responses to transient achromatic stimuli in the magnocellular layers of the LGN and the superficial layer of the SC of early glaucoma patients. *Hum. Brain Mapp.* **37**, 558–569 (2016).
44. Johnson, M. H. Subcortical face processing. *Nat. Rev. Neurosci.* **6**, 766–774 (2005).
45. Ajina, S., Pollard, M. & Bridge, H. The superior colliculus and amygdala support evaluation of face trait in blindsight. *Front. Neurol.* **11**, 769 (2020).
46. Le, Q. V. et al. A prototypical template for rapid face detection is embedded in the monkey superior colliculus. *Front. Syst. Neurosci.* **14**, 5 (2020).
47. Nguyen, M. N. et al. Neuronal responses to face-like and facial stimuli in the monkey superior colliculus. *Front. Behav. Neurosci.* **8**, 85 (2014).
48. Lorenzi, E., Nadalin, G., Morandi-Raikova, A., Mayer, U. & Vallortigara, G. Noncortical coding of biological motion in newborn chicks' brain. *Cereb. Cortex* **34**, bhae262 (2024).
49. Bonda, E., Ostry, D. & Evans, A. Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *J. Neurosci.* **16**, 3737–3744 (1996).
50. Ferrari, C., Ciricugno, A., Battelli, L., Grossman, E. D. & Cattaneo, Z. Distinct cerebellar regions for body motion discrimination. *Soc. Cogn. Affect. Neurosci.* **17**, 72–80 (2022).
51. Sokolov, A. A. et al. Biological motion processing: the left cerebellum communicates with the right superior temporal sulcus. *NeuroImage* **59**, 2824–2830 (2012).
52. Berman, R. A. & Wurtz, R. H. Functional identification of a pulvinar path from superior colliculus to cortical area MT. *J. Neurosci.* **30**, 6342–6354 (2010).
53. Berman, R. A. & Wurtz, R. H. Signals conveyed in the pulvinar pathway from superior colliculus to cortical area MT. *J. Neurosci.* **31**, 373–384 (2011).
54. Bogadhi, A. R., Katz, L. N., Bollimunta, A., Leopold, D. A. & Krauzlis, R. J. Midbrain activity shapes high-level visual properties in the primate temporal cortex. *Neuron* **109**, 690–699.e5 (2021).
55. Gurnsey, R. Limits of peripheral direction discrimination of point-light walkers. *J. Vis.* **10**, 1–17 (2010).
56. Wang, Y. et al. Modulation of biological motion perception in humans by gravity. *Nat. Commun.* **13**, 2765 (2022).
57. Liu, N. et al. Intrinsic structure of visual exemplar and category representations in macaque brain. *J. Neurosci.* **33**, 11346–11360 (2013).
58. Liu, N. et al. Oxytocin modulates fMRI responses to facial expression in macaques. *Proc. Natl Acad. Sci. USA* **112**, E3123–E3130 (2015).
59. Troje, N. F. et al. The little difference: fourier based synthesis of gender- specific biological motion. in *Dynamic Perception* 115–120 (2002).
60. Brainard, D. H. The psychophysics toolbox. *Spat. Vis.* **10**, 433–436 (1997).
61. Leite, F. P. et al. Repeated fMRI using iron oxide contrast agent in awake, behaving macaques at 3 tesla. *NeuroImage* **16**, 283–294 (2002).
62. Stephan, K. E. et al. Ten simple rules for dynamic causal modeling. *Neuroimage* **49**, 3099 (2010).
63. Cox, R. W. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res. Int. J.* **29**, 162–173 (1996).

Acknowledgements

We thank Dr. Zhentao Zuo for fMRI technical assistance. This research was supported by grants from the STI2030-Major Projects (2021ZD0203800 to Y. J., 2021ZD0200200 to N. L., 2021ZD0204200 to Y. X.), the National Natural Science Foundation of China (32430043 and 31830037 to Y. J.), the Interdisciplinary Innovation Team (JCTD-2021-06 to Y. J.), and the Fundamental Research Funds for the Central Universities to Y. J.

Author contributions

Y.J. conceptualized the study. Y.J., N.L., X.L., Z.H., and Y.X. designed the experiments. X. L., Z.H., Y.X., and T.Y. conducted data collection and analysis. P.Z. provided the analysis tool, and P.Z. and Y.W. contributed to the interpretation of the results. X.L., Z.H., and Y.X. wrote the original draft. All authors reviewed and approved the final version of the manuscript for submission.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-53968-x.

Correspondence and requests for materials should be addressed to Ning Liu or Yi Jiang.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.