Neuroimaging evidence of visual-vestibular interaction accounting for perceptual mislocalization induced by head rotation

Xin He [a,]* **and Min Bao** [a,b,c,]*
^aChinese Academy of Sciences, Institute of Psychology, CAS Key Laboratory of Behavioral Science, Beijing, China
^bUniversity of Chinese Academy of Sciences, Department of Psychology, Beijing, China
^cState Key Laboratory of Brain and Cognitive Science, Beijing, China

ABSTRACT. **Significance:** A fleeting flash aligned vertically with an object remaining stationary in the head-centered space would be perceived as lagging behind the object during the observer's horizontal head rotation. This perceptual mislocalization is an illusion named head-rotation-induced flash-lag effect (hFLE). While many studies have investigated the neural mechanism of the classical visual FLE, the hFLE has been hardly investigated.

Aim: We measured the cortical activity corresponding to the hFLE on participants experiencing passive head rotations using functional near-infrared spectroscopy.

Approach: Participants were asked to judge the relative position of a flash to a fixed reference while being horizontally rotated or staying static in a swivel chair. Meanwhile, functional near-infrared spectroscopy signals were recorded in temporal-parietal areas. The flash duration was manipulated to provide control conditions.

Results: Brain activity specific to the hFLE was found around the right middle/inferior temporal gyri, and bilateral supramarginal gyri and superior temporal gyri areas. The activation was positively correlated with the rotation velocity of the participant around the supramarginal gyrus and negatively related to the hFLE intensity around the middle temporal gyrus.

Conclusions: These results suggest that the mechanism underlying the hFLE involves multiple aspects of visual-vestibular interactions including the processing of multisensory conflicts mediated by the temporoparietal junction and the modulation of vestibular signals on object position perception in the human middle temporal complex.

1 Introduction

Motion can distort the position perception of objects. Many mislocalizations are accompanied by existence of a moving visual stimulus.[1–3] One of the best-known and most-studied motion-induced mislocalizations is the flash-lag effect (FLE),[4–6] where a flash would be perceived as lagging behind a moving object even if they are actually aligned. In addition to the classical FLE,

*Address all correspondence to Xin He, hex@psych.ac.cn; Min Bao, baom@psych.ac.cn

Schlag et al.[7] found an extension of the FLE without stimulation of visual motion: a flash would be perceived as lagging behind an object that remains in one spot relative to the head when the observer is making head rotations. To account for this novel head-rotation-induced FLE (herein abbreviated as hFLE), Schlag et al. followed the "motion extrapolation" hypothesis proposed by Nijhawan[5] that the neural system tends to predict the position of a moving object to compensate for the delay caused by the transmission and processing of neural signals. Although there was no retinal motion in this hFLE, Schlag et al. proposed that extraretinal motion could also be extrapolated.

As to the neural origin for the FLE, Maus et al.[8] and Wang et al.[9] proved the causal role of human middle temporal complex (hMT+) in FLE with TMS and tDCS, respectively. It is well known that hMT+ plays an essential role in forming the perception of visual motion and object position. This area not only represents the physical motion or position but also encodes implied motion, such as the motion aftereffect,[10–14] and the perceived position of the observer.[15,16] Thus, the hMT+ is involved in many mislocalizations induced by motion including the FLE.[17–19] By contrast, to our knowledge, the brain activity in the hFLE has not been investigated. As a multisensory illusion, it can be foreseen that the hFLE involves even more complicated processing than its classical visual version, but still surprisingly little is known about it. Elucidating the mechanism underlying the hFLE may shed light on the complex visual-vestibular interactions that occur every day in natural viewing conditions. Furthermore, it will add to our knowledge of the vestibular cognition, which, differing from other sensory modalities, is extensively distributed across the brain with distinct physiological characteristics and awaits deeper explorations.[20–23]

In previous work, we pointed out that the motion extrapolation hypothesis could not provide an adequate explanation of the forming mechanism of hFLE.[24] Instead, the account of visual-vestibular interaction was more preferred. We notice that vestibular stimulations[25–27] and visual-vestibular integration[28] can also activate hMT+. Moreover, hMT+ is functionally connected to several regions in the vestibular cortical network.[29,30] Considering the contribution of hMT+ to the classical FLE and the functions of hMT+ concerning both motion processing and visual-vestibular integration, we hypothesized that hMT+ would be activated when observer experienced the hFLE.

To test this hypothesis, one should seek an appropriate neuroimaging tool first. Unfortunately, this is considerably challenging owing to the technical limitations for common tools (e.g., fMRI and EEG). Compared with fMRI and EEG that are both susceptible to motion artifacts, functional near-infrared spectroscopy (fNIRS) is more tolerant to motion and has thus been broadly employed in research on exercises, under natural scenes, or with special participants like infants. In this study, we measured with fNIRS the cortical activity around hMT+ of the participants who performed horizontal head rotations to experience the hFLE.

2 Methods

2.1 Participants

Twenty-eight participants (11 males and 17 females) were recruited for the research, whose ages ranged from 19 to 32 years ($M \pm \text{SD} = 25 \pm 3$ years). The sample size was determined using G*Power[31] after a pilot experiment with 11 participants. All the participants had normal or corrected-to-normal visual acuity and reported no history of motion sickness [including simulator sickness with VR or three-dimensional (3D) display] nor dysfunction of visual or vestibular system.

This study protocol was approved by the Institutional Review Board of the Institute of Psychology, Chinese Academy of Sciences. All the participants gave their informed consent and were paid after the experiment.

2.2 Apparatus

The experiment program was written and run with MATLAB R2013a (MathWorks Inc.) and Psychtoolbox v3.0.12[32] on a Dell XPS 8700 PC (Dell Inc.), and the visual stimuli were presented to the participants on a Sony HMZ-T3 head-mounted goggle (Sony, Japan) with a field of view of 50 (horizontal) × 28 (vertical) degrees (deg) of visual angle, resolution of 1280 × 720 pixels, and
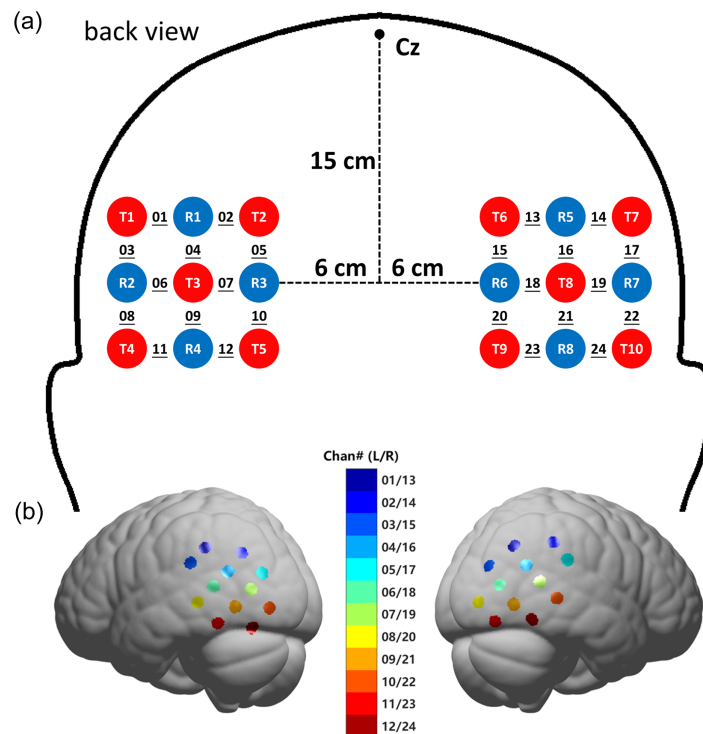
Fig. 1 (a) The optode layout used in the study. Emitters are marked with red (T) and detectors with blue circles (R). Each adjacent pair of emitter and detector were 3 cm apart from each other to form a channel (underlined). There were 12 channels on each hemisphere and 24 in total. The distances marked are geodesic along the scalp. All the distances and sizes diagramed in the figure are for illustration purpose only and not in real scale. (b) Average locations of the channels on the brain template.

refresh rate of 60 Hz. The goggles were worn outside the fNIRS headcap. Participants sat on a barber's chair to achieve rotation with their feet off the ground, and a TSS-WL 3-Space wireless motion sensor (Yost Labs Inc.) was fixed to one arm of the chair.

The fNIRS recording was implemented by a LABNIRS continuous-wave fNIRS system (Shimadzu Inc., Japan) with the near-infrared wavelengths of 780, 805, and 830 mm. Optode localization was carried out with a Fastrak 3D tracking and digitizing system (Polhemus Inc.).

2.3 fNIRS Acquisition

The optodes were arranged around T5/T6 of the international 10–20 system to cover bilateral temporo-parieto-occipital areas.[33–35] On each side was a 3×3 square layout with five emitters and four detectors yielding 12 channels [Fig. 1(a)]. The raw sampling rate of fNIRS data was 55.6 Hz. After each participant completed the main task, a localization was carried out for each optode as well as for anatomical landmarks of the vertex (Cz), nasion (Nz), and bilateral pre-auricular points (AL/AR). These localization data were then used to estimate the locations of the channels on the Montreal Neurological Institute (MNI) brain template.[36–38] The average locations of the channels on the brain template are shown in Fig. 1(b). The cortical subdivisions they fell on were then estimated based on the LONI Probabilistic Brain Atlas (LPBA40)[39] and Brodmann areas (BA).

2.4 Stimuli and Procedure

We assumed that the relevant brain activation during the hFLE was the result of highly independent components caused by three types of neural signals, namely: (i) response to visual stimulus (i.e., the flash bar), (ii) response to vestibular stimulus (i.e., the horizontal body rotation), and (iii) the signal of visual–vestibular interaction leading to the hFLE. Based on this assumption, we designed five experimental conditions to decompose the brain activation during the hFLE.

2.4.1 *Short flash + moving (sF+M)*

In light of previous findings where passive movement could also induce the hFLE,[40] in this study, participants were seated in a swivel barber's chair with their feet off the ground. The horizontal rotation was performed by the experimenter to reduce the influence of cervical movement on the fNIRS recording. The experimenter received a tone cue through earphones to start or stop the rotation. The display on the LCD monitor was identical to that on the goggles, allowing the experimenter to read the instruction.

Throughout the whole experimental session, the background of the display was pure black with a centered, red fixation point with 0.15 deg diameter. Before each block started, two lines of texts displayed the number of upcoming blocks and an instruction "Rotate!" to remind participants that they needed to rotate during the block. After participants were ready with their feet off the ground, the experimenter pressed a key to start the block. For the first 5 s, only a central reference bar and the fixation point were presented while a countdown with beep tones through earphones reminded the experimenter to prepare to rotate the chair. The reference bar was white with a size of 5 deg (height) × 0.5 deg (width).

The experimenter rotated the chair once per trial based on the auditory cues from the earphones. A high tone (1500 Hz) indicated starting the rotation, and a low tone (1000 Hz) indicated ceasing, each lasting for 50 ms. Each rotation lasted for 2 s with 1-s intertrial intervals (Fig. 2). The experimenter purposely controlled the ends of rotation trajectories on approximately the same positions to keep the rotation amplitude roughly constant across trials. Each block consisted of 15 trials with alternating rotating directions between leftward and rightward.
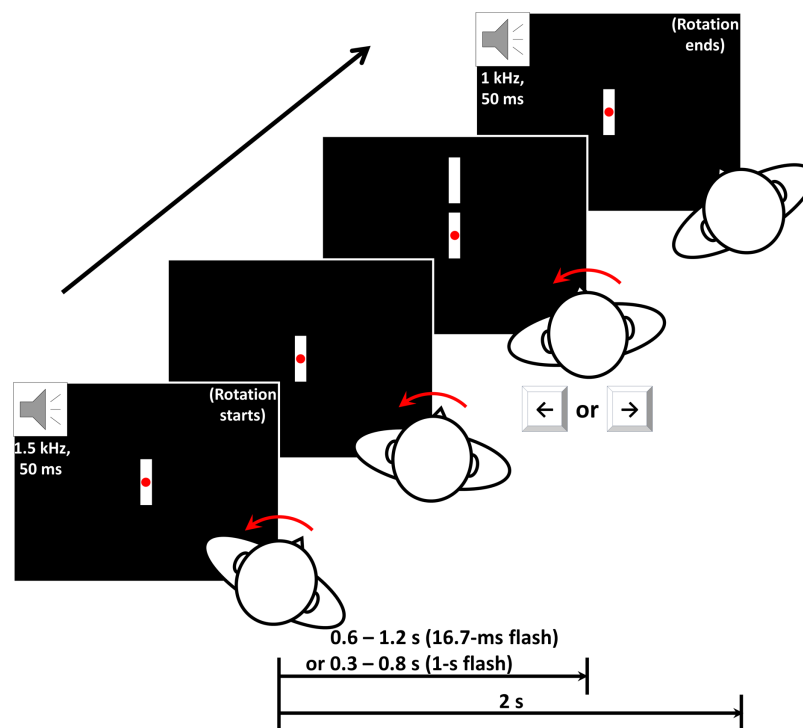
Fig. 2 Illustration of a single trial in condition A or C. The experimenter started rotating the chair on hearing a high-pitch tone. 0.6 to 1.2 s (condition A) or 0.3 to 0.8 s (condition C) later, the flash bar appeared above the reference bar for 16.7 ms (condition A) or 1 s (condition C). The participant was asked to judge whether the flash bar was to the left or the right of the reference bar although the two bars were actually vertically aligned. The rotation was stopped after the low-pitch tone, which was 2 s later than the high-pitch tone. These auditory cues were played to the experimenter through earphones, so the participant could not hear them. The figure only showed the case where the participant was rotated leftward, but the rotating direction alternated between trials actually. All the distances and sizes of the visual stimuli in the figure are for illustration purpose only and not in real scale.

Table 1 Summary of all conditions.

Signal	Condition				
	A (sF+M)	B (sF+S)	C (lF+M)	D (lF+S)	E (nF+M)
Visual	√	√	√	√	×
[duration (ms)]	[16.7]	[16.7]	[1000]	[1000]	—
Vestibular	√	×	√	×	√
[velocity (°/s)]	[42.11 ± 3.10]	—	[32.88 ± 2.77]	—	[41.81 ± 3.03]
hFLE (expected)	√	×	×	×	×

Note: Rotation velocities are reported in mean ± standard deviation. The velocity in condition C was lower than that in condition A because of their different flash onset delays relative to the beginning of trials (see Sec. 2.4).

After a random delay ranging from 0.6 to 1.2 s starting from the beginning of each rotation, a flash bar identical to the reference bar would appear 5 deg right above the reference bar for 1 frame (16.7 ms) and then vanish. Although the flash bar and target bar were vertically aligned physically, the participant was informed in advance that there was always a slight misalignment between the two bars and required to perform a 2AFC task discriminating whether the flash bar was to the left or the right of the reference bar by pressing the left or the right arrow key while maintaining fixation at the central fixation point. The reference bar remained centered on the display throughout each block.

The mean velocity across 0.5 s right before each flash onset was recorded as the rotation velocity. The average rotation velocities for conditions with rotation (i.e., conditions A, C, and E) are listed in Table 1.

2.4.2 Short flash + static (sF+S)

Condition B was designed to measure the brain activation caused by the visual stimulus alone in condition A. Therefore, participants in condition B were not rotated. The other procedures and tasks were the same as those in condition A, except that the instruction before the block read "Do NOT rotate!" instead.

2.4.3 Long flash + moving (lF+M)

Past studies have shown that the flash-lag effect may weaken with increasing duration of the flash stimulus.[41–44] Furthermore, the upper limit of the flash duration that could induce the FLE was 80 to 500 ms.[41,42,44,45] Thus, in condition C we lengthened the duration of the flash bar to 1 s. Hopefully, this modification might suppress the hFLE while maintaining similar visual stimuli to those in condition A. In terms of timing, the random delay before the flash was altered to 0.3 to 0.8 s (Fig. 2). Note that the rotation velocity recorded in condition C was lower than that in condition A (see Table 1) due to this shorter delay. The other procedures and tasks were identical to those in condition A, and participants were required to make responses after the flash bar disappeared.

2.4.4 Long flash + static (lF+S)

Condition D was almost the same as condition C except that participants in this condition were not rotated. The instruction before the block also read "Do NOT rotate!" instead of "Rotate!"
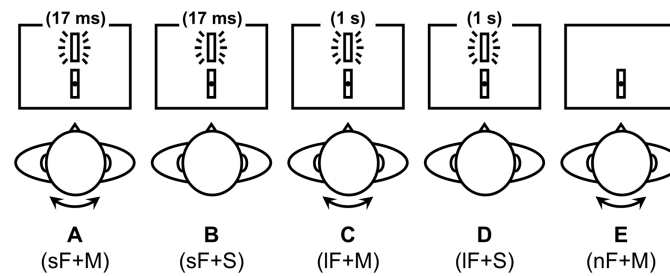
Fig. 3 Simplified illustration and comparison of the conditions. The visual stimuli are for illustration only. For individual GLM-based analysis, a β-value for each condition was obtained (β_a to β_e). Then, we examined the following three types of contrasts: (i) $\Delta\beta_1 = (\beta_a - \beta_b) - (\beta_c - \beta_d)$, (ii) $\Delta\beta_2 = (\beta_a - \beta_b) - \beta_e$, and (iii) $\Delta\beta_3 = (\beta_c - \beta_d) - \beta_e$.

2.4.5 No flash + moving (nF+M)

Condition E served as a control to measure the brain activation caused solely by the self-rotation. Therefore, the flash bar was not presented and participants did not have to make any response. All the participant had to do was to keep gazing at the fixation point. The other procedures were the same as those in condition A. In particular, the program sent a trigger to the LABNIRS system to mark a virtual "flash-onset" event at a random time point of each trial to maintain a consistent data structure with the other conditions for the fNIRS data analyses.

The conditions are summarized and compared in Fig. 3 and Table 1. Each condition consisted of five blocks, resulting in a total of 25 blocks for each participant, with the order randomized according to a Latin square design. Participants received a minimum of 15-s rest between blocks during which they could take a break without large head or body movements.

2.5 Behavioral Data Analysis

For conditions A and C, the proportion of responses that were in the opposite direction to self-rotation (e.g., reporting the flash bar to the left of the reference bar when rotating rightward and vice versa) was calculated to indicate the intensity of hFLE. If this proportion was significantly higher than 50%, it would indicate an hFLE. As for conditions B and D where the participants did not rotate, we simply calculated the proportion of responses reporting the flash bar to the left to measure response bias. A 2×2 repeated-measures ANOVA (short/long flash × moving/static) was then conducted on these response proportions to compare the intensity of hFLE across different conditions.

2.6 fNIRS Data Analysis

Both the absorbance (ABS) data of each wavelength and the relative changes in concentration of oxygenated, deoxygenated, and total hemoglobin (denoted as [HbO], [HbR], and [tHb], respectively) were obtained from the fNIRS system. These data were processed and analyzed using MATLAB 2017a (MathWorks Inc.) with the NIRS-KIT[46] and SPM12[47] packages.

2.6.1 Pre-processing

The pre-processing consisted of several steps. Figure 4 demonstrates the pre-processing pipeline step by step with an instance.

(I) Downsampling: We first downsampled the data from the sampling period of 18 to 54 ms (i.e., sampling rate from 55.56 to 18.52 Hz) using the resampling function built in the LABNIRS system. Then, both the ABS data and [Hb] data were exported as the raw data.

(II) Motion artifact correction: The head-motion-related artifacts were corrected with the temporal derivative distribution repair (TDDR) method[48] built in the NIRS-KIT toolbox. Based on iteratively reweighting the temporal derivatives of the fNIRS signal, this parameter-free algorithm uses a robust regression approach to remove large fluctuations such as spikes and baseline shifts attributed to motion artifacts while leaving smaller,
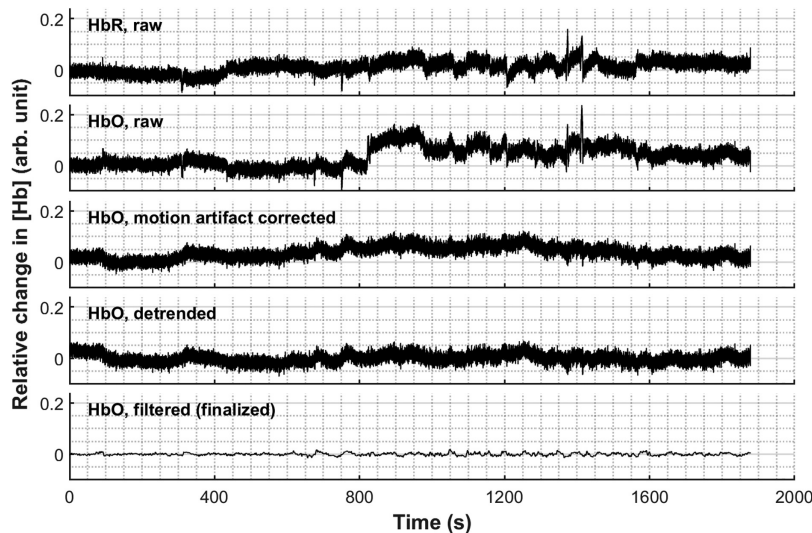
Fig. 4 Demonstration of a sample of raw [HbO] time series (second row; downsampled already) undergoing the pre-processing procedure step by step (third to bottom row). This signal was obtained from Ch. #3 of participant #5. The corresponding raw [HbR] time series (top row) is also plotted as a comparison. All the signals are in arbitrary unit.

hemodynamic fluctuations, which can be used on either optical density or hemoglobin concentration signals.[46,48–51]

To confirm the validity of the TDDR method, we conducted a repeated-measures ANOVA on the signal-to-noise ratio (SNR) of ABS data with three factors (wavelength, channel, and correction: raw versus TDDR-corrected). Focusing on the factor of correction, the results showed a significant main effect $[F(1,27) = 31.81, p < 0.001, \eta_p^2 = 0.54]$, with an SNR increment (ΔSNR) of 3.42 dB after the TDDR correction. Also, the wavelength $\times$ correction interaction was significant $[F(2,54) = 8.31, p = 0.001, \eta_p^2 = 0.24]$, showing that the correction effect varied across wavelengths but all valid (ΔSNR$_{780\ nm}$ = 3.61 dB, ΔSNR$_{805\ nm}$ = 3.29 dB, ΔSNR$_{830\ nm}$ = 3.34 dB, ps < 0.001). The channel $\times$ correction interaction was not significant $[F(23,621) = 1.17, p > 0.3, \eta_p^2 = 0.04]$, indicating a relatively homogeneous correction effect across channels (ΔSNR range = [2.25, 5.70] dB, ps < 0.014). The three-way interaction was not significant either $[F(46,1242) = 0.80, p > 0.6, \eta_p^2 = 0.03]$. These results proved that the TDDR method was suitable and effective for the current data.

(III) Detrending. The data were detrended using a second-order polynomial regression model.

(IV) Filtering. The data were filtered using a third-order IIR filter with a passband set at 0.01 to 0.39 Hz since the actual interstimulus intervals between flashes presented were all longer than 2.6 s.[52]

(V) Outlier rejection. The data at each channel of each participant were then segmented and analyzed block-wise, based on the time series from the first flash onset until the last flash offset of a block. The data were assessed with six indices: (1) SNR ($= 20 \log_{10} \frac{M}{SD}$); (2) number of data loss (NDL), the number of the local data sequences with a fixed value (≥ 2 consecutive data points); (3) maximum continuous data loss (MCDL), the maximum length of the local data sequences with a fixed value; (4) repetition rate (RR), the proportion of repetitive values (= 1-number of distinct values/data length); (5) the range (R); and (6) the standard deviation (SD). Among them, the indices (2) to (6) were specially defined to automatically detect a certain type of low-quality data: upon visual inspection of the data, we noticed that many abnormal data involved a common pattern of varying between several limited values or remaining fixed for a period of time, which might not be screened by SNR because their SNR often seemed normal due to abnormally small SD. SNR was calculated for the average ABS of the three wavelengths, and the others for [HbO].

First, SNR, NDL, and MCDL were calculated directly on the basis of raw data. Blocks with SNR < 20 dB, NDL > 6, or MCDL > 2 were first rejected. Then, R, SD, and RR of the remaining blocks were assessed by their deviation within the distribution. Considering the high skewness of these distributions, we adopted the adjusted boxplot method[53] with the LIBRA toolbox[54,55] instead of the ordinary normalization. Blocks with R or SD beyond $[Q_1 - 1.5e^{-4\text{MC}} \cdot \text{IQR}, Q_3 + 1.5e^{3\text{MC}} \cdot \text{IQR}]$ or RR beyond $[Q_1 - 30e^{-4\text{MC}} \cdot \text{IQR}, Q_3 + 30e^{3\text{MC}} \cdot \text{IQR}]$ were further rejected (Q_1, first quartile; Q_3, third quartile; IQR, interquartile range; MC, medcouple; see Ref. 53). All the criteria except for SNR were determined to distinguish as many abnormal data described above from the rest as possible by visual inspection.

Finally, only those channels that had all five conditions each containing at least one survival block could enter the further analyses. That is, if on a channel any condition had all its corresponding blocks rejected, then the whole channel would be excluded. Altogether, about 4.2% of all the blocks were rejected.

2.6.2 *First-level (individual) analysis*

Analyses below mainly focused on the change of [HbO] as it is considered a more sensitive measure compared to [HbR] or [tHb][56–58] and many past studies have selected [HbO] as the main analysis target.[34,59–61] However, the results for [HbR] will also be reported as a comparison.

The data were analyzed using a block design approach. The onset of the first flash and the offset of the last one in each block were defined as the beginning and end, respectively, and the duration between them was considered as the block duration. After pre-processed, the data of each participant were fitted to a GLM based on the canonical (two-gamma) hemodynamic response function provided by the SPM12 toolbox to obtain a β-value for each condition (β_a to β_e). Then, we focused on three types of contrasts below: (i) $\Delta\beta_1 = (\beta_a - \beta_b) - (\beta_c - \beta_d)$, (ii) $\Delta\beta_2 = (\beta_a - \beta_b) - \beta_e$, and (iii) $\Delta\beta_3 = (\beta_c - \beta_d) - \beta_e$.

According to the assumption mentioned above, β_a contained all three types of signals while β_b represented only the effect of the visual stimulus, so $(\beta_a - \beta_b)$ represented the vestibular component and the vestibular-visual interacting component (i.e., the hFLE). The case was similar for β_c and β_d, except that the long flash was supposed not to induce the hFLE, so $(\beta_c - \beta_d)$ represented only the vestibular component. Therefore, $\Delta\beta_1$ should represent a relatively pure effect of the hFLE. We also calculated $\Delta\beta_2$ by replacing $(\beta_c - \beta_d)$ with β_e. And to examine whether the long flash was capable of inducing the hFLE, $\Delta\beta_3$ was also calculated.

2.6.3 *Second-level (group) analysis*

We conducted one-sample t-tests to compare $\Delta\beta$ on each channel with a baseline of 0 in order to examine the differential brain activation across conditions. A significant deviation from 0 for $\Delta\beta_1$ or $\Delta\beta_2$ would indicate a potential locus that represented the hFLE on that channel. Besides, we analyzed the correlation between $\Delta\beta$ and rotation velocity or behavioral performance to relate neural activation to stimulus intensity or perception. As the analyses were conducted separately on each channel, multiple comparison corrections were implemented using the false discovery rate (FDR) method[62] within either hemisphere.

For all the statistics above and below, effect sizes are reported as Cohen's d (d for short) for t-tests or η_p^2 for ANOVA. All the p-values reported in the ANOVA have been adjusted where necessary using Greenhouse–Geisser correction, and post hoc tests were conducted with the Tukey–Kramer method.

3 Results

3.1 Behavioral Performance

According to the one-sample t-test results (Fig. 5), when the participants remained static (i.e., conditions B and D), their responses for the flash bar appearing on the left side did not deviate significantly from 50% [B: $M \pm SD = (47.36 \pm 21.32)\%$; D: $(52.31 \pm 23.93)\%$; $ps > 0.5$], indicating little response bias. However, when the participants were moving (conditions A and C), responses for the flash bar appearing in the opposite side of the rotating direction were
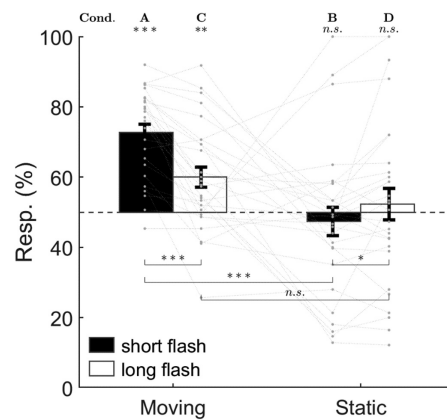
Fig. 5 The behavioral results. Gray dots and dashed lines represent the individual data. Error bars indicate 1 SEM. n.s., not significant; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

significantly greater than 50% of the total [A: $(72.69 \pm 12.64)\%$, $t(27) = 9.50$, $p < 0.001$, $d = 1.80$; C; $(59.98 \pm 15.19)\%$, $t(27) = 3.48$, $p = 0.002$, $d = 0.66$], which suggested an evident hFLE in both flash durations.

As for the ANOVA, the main effect of flash duration was significant [$F(1,27) = 9.68$, $p = 0.004$, $\eta_p^2 = 0.26$] with a higher response proportion for short flashes than for long ones, and so was the main effect of motion state [$F(1,27) = 12.01$, $p = 0.002$, $\eta_p^2 = 0.31$] with a higher response proportion when moving than when static. The interaction between flash duration and motion state was also significant [$F(1,27) = 22.33$, $p < 0.001$, $\eta_p^2 = 0.45$]: the response proportion was higher when moving than when static ($p < 0.001$) but only for short flashes. In addition, the response proportion for short flashes was much higher than that for long flashes when moving ($p < 0.001$) but slightly lower when static ($p = 0.043$). These results showed that although either short or long flashes could induce the hFLE, the effect was relatively weaker for long flashes than for short flashes.

3.2 fNIRS

The actual sample sizes in the fNIRS analyses were less than the number of participants since some channels of some participants were excluded in the pre-processing stage, which will be reported below.

First, we examined the deviations of $\Delta\beta$ from 0. For [HbO], as shown in Fig. 6 and Table 2, several channels showed significant deviation of $\Delta\beta_1$ from 0, but none of them survived the FDR
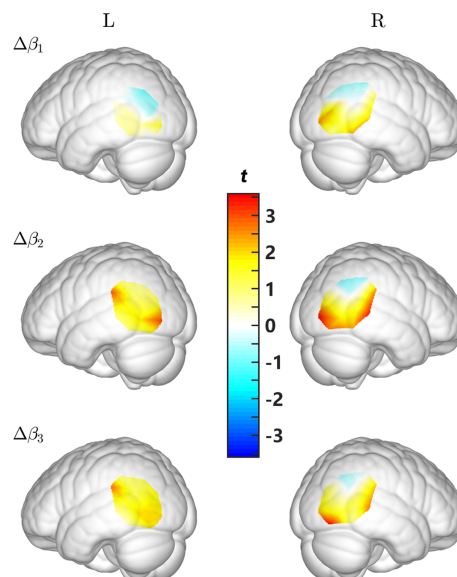
Fig. 6 Results of t-tests (from top to bottom for $\Delta\beta_1$ to $\Delta\beta_3$, respectively).

Table 2 Results of the examinations on $\Delta\beta$ with uncorrected ps < 0.05.

	Ch. #	N	t	d
$\Delta\beta_1$	—	—	—	—
	18	28	2.08*	0.39
	20	28	2.46*	0.47
	22	28	2.18*	0.41
	24	27	2.33*	0.48
$\Delta\beta_2$	—	—	—	—
	3	26	**2.98****	0.58
	10	25	**3.15*****	0.63
	17	28	2.41*	0.46
	20	28	**3.57*****	0.67
	22	28	**3.46*****	0.65
	23	27	**2.73***	0.52
	24	27	2.35*	0.45
$\Delta\beta_3$	—	—	—	—
	3	26	2.72*	0.53
	20	28	2.16*	0.41
	22	28	2.70*	0.51
	23	27	**3.59*****	0.69

Note: N means the size of the sample actually included in the analysis on each channel, and bold font indicates surviving the FDR correction.
*$p < 0.05$;
**$p < 0.01$;
***$p < 0.005$

correction. $\Delta\beta_2$ deviated significantly from 0 on more channels, among which Ch. #3, #10, #20, #22, and #23 remained significant after the FDR correction. Similarly, several channels showed significant $\Delta\beta_3$, and Ch. #23 passed the FDR correction. In contrast, for [HbR], Ch. #22 showed significant deviation of $\Delta\beta_2$ from 0 and #7, #14, and #22 showed significant deviation of $\Delta\beta_3$ from 0 (all ps < 0.05), but none of them survived the FDR correction.

We then decided to focus on Ch. #3, #10, #20, and #22, which passed the FDR correction, as the main results. Although Ch. #23 also passed the FDR correction for $\Delta\beta_2$, it was excluded with caution because it did so for the control $\Delta\beta_3$ as well, which was not specific enough to the hFLE. Due to a similar reason, Ch. #17 and #24 were leniently selected because they showed relatively large raw $\Delta\beta_2$ but no $\Delta\beta_3$, which was consistent with the anticipation of the experiment. The results of localization for them are listed in Table 3.
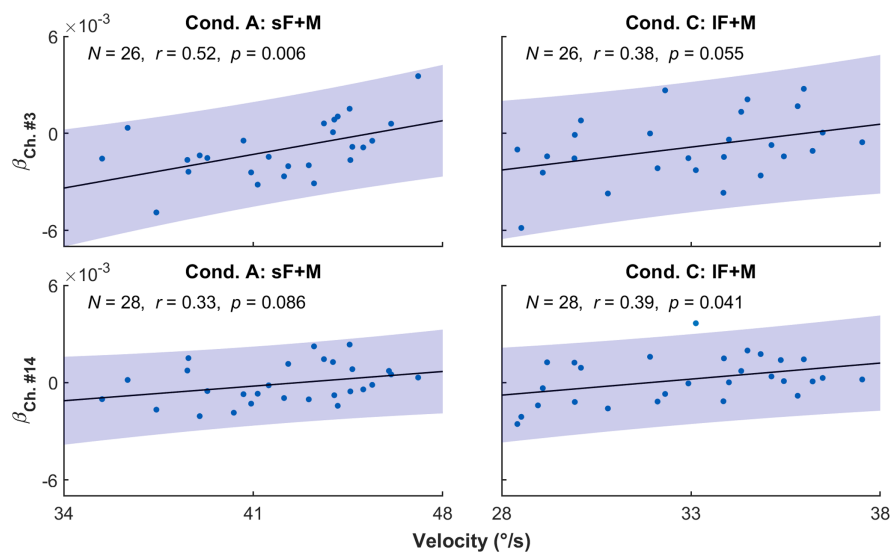
The correlation between β and rotation velocity in the two conditions with rotation (i.e., A and C) was then explored. As shown in Fig. 7 and Table 4, Ch. #3 showed significant positive correlation in condition A ($N = 26$, $r = 0.52$, $p = 0.006$) and Ch. #15 in condition C ($N = 28$, $r = 0.39$, $p = 0.041$). Besides, both channels showed the trend of positive correlation on the other condition (Ch. #3, condition C: $N = 26$, $r = 0.38$, $p = 0.055$; Ch. #15, condition A: $N = 28$, $r = 0.33$, $p = 0.086$). In fact, the posterior temporal-parietal area was found to have a slightly higher correlation than other tested regions (Fig. 8).

Finally, we tested the correlation between $\Delta\beta$ and the corresponding hFLE intensity in conditions A and C. None of the focused channels were found to have a significant correlation,

Table 3 Brain regions corresponding to the focused channels.

Ch. #	MNI coordinates			Regions (probabilities)	
	x	y	z	LPBA40	BA
3	−68	−44	22	SMG (0.52)	22 (0.72)
				STG (0.47)	48 (0.24)
10	−43	−91	−7	MOG (0.77)	19 (0.63)
				IOG (0.23)	18 (0.37)
17	69	−43	25	STG (0.44)	22 (0.58)
				AG (0.31)	48 (0.21)
20	38	−96	−3	MOG (0.74)	18 (0.96)
				IOG (0.26)	19 (0.04)
22	69	−51	−2	MTG (0.92)	37 (0.53)
				ITG (0.08)	21 (0.34)
24	59	−66	−16	ITG (0.70)	37 (0.86)
				IOG (0.14)	19 (0.14)

Note: SMG, supramarginal gyrus; STG, superior temporal gyrus; MOG, middle occipital gyrus; IOG, inferior occipital gyrus; AG, angular gyrus; MTG, middle temporal gyrus; ITG, inferior temporal gyrus.


Fig. 7 Correlation between β on Ch. #3 (upper row) and #14 (lower row) and rotation velocity in conditions A (left column) and C (right column).

but Ch. #6 ($N = 27$, $r = -0.43$, $p = 0.027$) and #8 ($N = 27$, $r = -0.44$, $p = 0.022$) showed significant negative correlation between $\Delta\beta_2$ and the response proportion in condition A, whereas Ch. #15 ($N = 28$, $r = 0.50$, $p = 0.007$) and #18 ($N = 28$, $r = 0.37$, $p = 0.050$) showed significant positive correlation between $\Delta\beta_3$ and the response proportion in condition C (Fig. 9). However, these results did not pass the FDR correction. The regions involved for these channels are listed in Table 5. As shown in Fig. 10, the negative correlation was mainly distributed across posterior part of the medium temporal gyrus in condition A, whereas the lateral middle occipital area showed positive correlation in condition C.
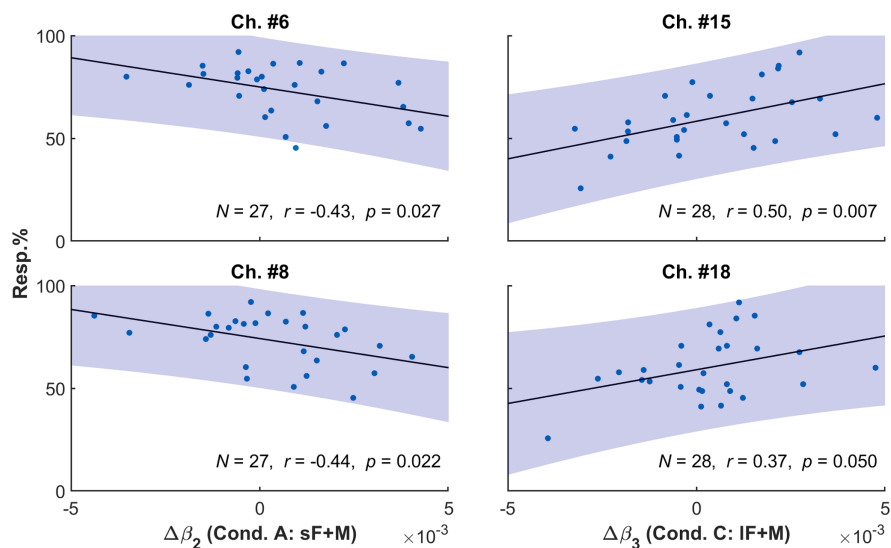
Table 4 Channels with correlation to the rotation velocity.

	MNI coordinates			Regions (probabilities)	
Ch. #	x	y	z	LPBA40	BA
3	−68	−44	22	SMG (0.52)	22 (0.72)
				STG (0.47)	48 (0.24)
14	64	−52	37	AG (0.91)	40 (0.63)
				SMG (0.07)	39 (0.22)

Note: MTG, middle temporal gyrus; ITG, inferior temporal gyrus; AG, angular gyrus; SOG, superior occipital gyrus; MOG, middle occipital gyrus.
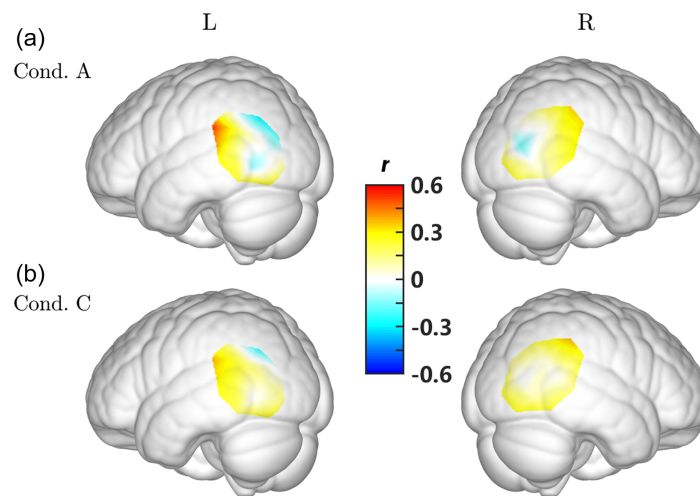
Fig. 8 Distribution of the correlation between β and rotation velocity in conditions (a) A and (b) C.
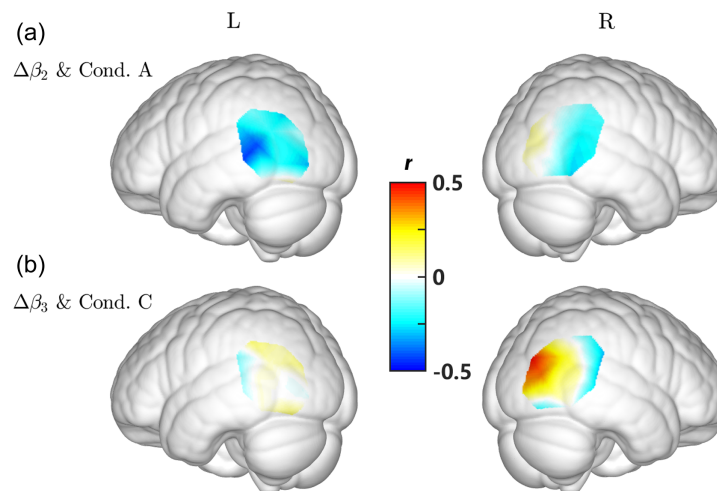

Fig. 9 Correlation between $\Delta\beta$ and the corresponding response proportion.

Table 5 Channels with correlation to the behavioral performance.

Ch. #	MNI coordinates			Regions (probabilities)	
	x	y	z	LPBA40	BA
6	−65	−59	6	MTG (0.95)	37 (0.77)
				AG (0.03)	21 (0.23)
8	−68	−49	−5	MTG (0.71)	37 (0.42)
				ITG (0.29)	21 (0.33)
15	40	−89	22	MOG (0.99)	19 (0.89)
				SOG (0.01)	18 (0.11)
18	49	−85	8	MOG (0.96)	19 (0.89)
				IOG (0.04)	18 (0.11)

Note: MTG, middle temporal gyrus; ITG, inferior temporal gyrus; AG, angular gyrus; SOG, superior occipital gyrus; MOG, middle occipital gyrus; IOG, inferior occipital gyrus.

Fig. 10 (a) Distribution of the correlation between $\Delta\beta_2$ and the behavioral performance in condition A and (b) between $\Delta\beta_3$ and the behavioral performance in condition C.

However, no significant correlation was observed directly between rotation velocity and behavioral performance (condition A: $r = -0.21$, $p > 0.2$; condition C: $r = 0.04$, $p > 0.8$).

In summary, the brain activity related to the hFLE was significant in right MTG/ITG (BA 37; Ch. #20 & #24), bilateral SMG/STG (BA 22; Ch. #3 & #17), and bilateral MOG/IOG (BA 18/19; Ch. #10 & #20). The activity in the SMG correlated with the rotation velocity, whereas that in the MTG correlated with the hFLE intensity.

4 Discussion

By using fNIRS, this study measured for the first time the brain activity corresponding to the hFLE. We provided the neuroimaging evidence for the engagement of left supramarginal gyrus (SMG)/superior temporal gyrus (STG) area (BA 22), the posterior part of bilateral middle temporal gyri (MTG)/inferior temporal gyri (ITG) areas (BA 37), and lateral occipital areas (BA 18/19).

To manipulate the hFLE while minimizing the alteration of the visual stimuli, we extended the duration of the flash bar based on previous literature on classical FLE.[41,42,44] However, it turned out that the hFLE was reduced but not fully eliminated when the flash duration was extended to 1 s, as shown by the behavioral results that the hFLE for the long flash was also significant though weaker than that for the short flash and not significantly different from the control condition when the participants were static. Although it is difficult to propose an explicit account for this finding with only current data, we believe that it might reflect a certain difference between the mechanisms for the hFLE and classical FLE, which requires further empirical evidence. For example, the formation of hFLE might involve higher-level sensory processing such as multisensory interaction compared to the classical FLE and thus took a longer time span, which could cover a longer flash stimulus. On the other hand, the current behavioral task forced participants to discriminate the mislocalization between two physically aligned bars, resulting in virtually no signal to detect. Therefore, intrinsic neural noise might have a significant impact on perception,[63,64] leading to the noticeable individual response biases in both static control conditions (B and D). This could also explain the observed hFLE for the long flash.

Consistent with our hypothesis that the hMT+ activates during the hFLE, Ch. #22 and #24 were identified by the analyses of $\Delta\beta$ contrasts of interest. Both channels are located in the middle and inferior temporal gyri (MTG and ITG) and both in BA 37. Despite that we did not functionally localize the exact divisions, we could infer from these estimated locations that they are likely situated in the middle temporal (MT) area, which is a main subregion of the hMT+, typically located around the occipital continuation and ascending branch of the inferior temporal sulcus.[65–68] The hMT+ and its homolog in the macaque have been well recognized as key areas for processing visual motion information, both in macaques and humans. Interestingly, recent studies have demonstrated that they also to respond to vestibular stimuli and participate in visual-vestibular integration to form heading and self-motion perception.[25–27,69–74] Della-Justina et al.[28] presented participants with visual stimulus of flickering checkboards, vestibular stimulus of galvanic vestibular stimulation (GVS), and combined stimulus, and observed activation of ITG to single visual stimulus and visual-vestibular combined stimulus. Furthermore, Indovina et al.[30] used graph-theoretical network analysis to find that BA 37 was structurally connected to the posterior insula cortex (PIC), a key hub in the vestibular cortical network, and that BA 37 itself also showed some hubness in the vestibular cross-modal network.

Based on these previous findings, we reckon that the activity of BA 37 is related to visual-vestibular integration processing that occurs in hMT+. The negative correlation between the activity in BA 37 and the behavioral performance may help clarify its role. Given that there was no visual motion stimulus in this study, it is likely that the activity of hMT+ represented the afferent retinal motion signals generated by eye movements and intrinsic neural noises of the participant. One possible explanation for this is that weaker activation of the hMT+ in a participant reflects fewer eye movements or neural intrinsic noises, resulting in weaker feedforward visual signals. This would lead to less weight of visual signals in the visual-vestibular integration and ultimately leave the estimation of object position more susceptible to the modulation of vestibular signals. Alternatively, the hMT+ could be subject to the inhibition by vestibular signals or by the feedback from visual-vestibular integration. Indeed, several past studies have observed mutual inhibition in visual-vestibular integration,[29,75,76] especially when the stimulus from the suppressed modality was weak. This suppression might aim to reduce neural noise from the weak modality and resolve sensory conflict.[75] In any case, the intensity of hFLE reflects the relative disadvantage of the visual modality compared to the vestibular modality in their integration during self-motion in our study.

On the other hand, Ch. #3 and #17 (and #14 in the correlation analysis) are located around SMG and STG (especially the posterior part, BA 22), two areas that also participate in visual-vestibular integration, especially the representation and processing of sensory conflicts.[27,33,77] Both areas are parts of the temporoparietal junction (TPJ), a functional complex between the parietal and temporal lobes and around the posterior end of the Sylvian fissure that plays a critical role in human vestibular system (see Ref. 78 for review) and found to be functionally connected to PIC and the parietal-insular vestibular cortex, two essential hubs in the vestibular network.[30,79–81] Previous neuroimaging research studies have found that the temporoparietal areas, including the SMG and STG, respond to caloric vestibular stimulation or GVS[22,82] and

activate in body-related tasks such as standing balance or inner verticality where the cues from other sensory modalities than vestibular were deprived,[83,84] and the STG response increased with the GVS intensity.[82] In an fNIRS study using horizontal rotation as the vestibular stimulation, Nguyen et al.[33] found that TPJ was sensitive to the incongruency between visual and vestibular stimuli, and that the activation of dorsal SMG negatively correlated with the intensity of subjective vertigo reported by the participant.

Although both hMT+ and TPJ engage in visual-vestibular integration, there are still some differences. Compared with hMT+, which is more of a visual related area, TPJ is closer to vestibular processing in visual-vestibular integration. This could also explain the different correlation results. The brain activation to the hFLE on Ch. #3 and #14 (TPJ) was positively correlated to the velocity probably because TPJ was more active in solving bigger visual-vestibular conflicts when the vestibular stimulation was stronger. In contrast, the brain activity on Ch. #6 and #8 identified in the correlation analysis on β and located in hMT+ influenced behavioral responses, which reflected the visual perception and subsequent sensory decision making affected by the visual-vestibular interaction.

As to Ch. #10 and #20 (and #15 and #18 in the correlation analysis), they fell in the lateral part of the medial occipital area (BA 18/19, extrastriate visual cortex V2/V3). The brain activity here might be reflective of the remaining effect from the visual stimulus of the flash, which was not fully eliminated in $\Delta\beta_2$ and $\Delta\beta_3$. This also explains the two distinct types of correlations between $\Delta\beta$ and the hFLE intensity in conditions A and C: in condition A with a weak visual input and strong hFLE, the performance was mainly determined by the cross-modal interaction in hMT+ as stated above, whereas in condition C where the visual input was relatively strong and the illusion was weak, the perception was likely influenced much by the response bias, which reflected the intrinsic neural noise of the extrastriate visual cortex V2/V3, as discussed at the beginning of the section.

Finally, despite that the brain activity was observed to correlate with both the rotation velocity and the hFLE intensity, no significant correlation was found directly between the latter two, which was consistent with our previous findings.[24] In that study, an hFLE was significantly induced both when the participants actively rotated the head (experiment 1, HM condition) and when they were passively rotated in the chair by the experimenter (experiment 1, BM condition), but the hFLE intensity significantly correlated to the rotation velocity only in the HM and not in the BM condition, which is interesting especially because these two conditions did not show significant difference in the hFLE intensity. And that finding in the BM condition was replicated in this study. We suspect that the decisive process to cause hFLE depends solely on the vestibular stimulation and does not distinguish its source, i.e., whether it comes from an active or a passive self-motion, but the intensity of the hFLE might be modulated by some mechanisms regarding this source, such as corollary discharge.[85–87]

Note that our findings were not symmetric across both hemispheres, with activation in the right hemisphere being generally stronger. This might be attributed to the fluctuations and noise in neural activity and data acquisition, but another possibility could be lateralization. Many studies have shown that at least part of the vestibular network was lateralized both anatomically and functionally toward a right-sided dominance,[23,29,30,88–90] including when processing horizontal rotating directions.[91] This lateral dominance was proved dependent partly on the handedness and often showed in the non-dominant hemisphere,[81,92] which might account for the current findings, but other factors and mechanisms are also involved.[22,30,88,93]

An important concern awaiting future investigations is to detail the distinct division of labor in the contribution of these brain regions to the hFLE as well as other visual-vestibular interaction instances, considering the widespread distribution of the vestibular network across the cortex.[21,23,94–96] These different roles may even relate to unique physiological characteristics such as elongated hemodynamics response.[22] And the current results should be interpreted cautiously and further evidence is still needed due to several limitations. First, the spatial resolution of fNIRS is much lower than fMRI, and fNIRS can only detect cortical surface activation due to its imaging principle. Therefore, the results may not be sufficiently accurate or complete. Besides, the brain atlas used here is mainly anatomical rather than functional, making it challenging to match the current findings to past fMRI research about visual-vestibular integration. Future studies can use neuronavigation techniques to help access more precise ROIs and

improve data reliability[97–99] and neuromodulation tools such as TMS to confirm whether these areas play a causal role. Also, a motion platform[100] can be used to implement finer and stricter manipulations of the motion profiles.

5 Conclusion

This study observed the activation of right SMG/STG area (BA 22), bilateral MTG/ITG areas (BA 37), and bilateral MOG/IOG areas (BA 18/19) in participants experiencing the hFLE. The activation of SMG was positively correlated with the rotation velocity while the activation of MTG was negatively correlated with the intensity of the hFLE. These findings demonstrate the role of visual-vestibular interaction in the formation of hFLE and provide indirect support to the visual-vestibular interaction account of the hFLE,[24] suggesting that multiple visual-vestibular interactions help form the hFLE including the processing of multisensory conflicts in TPJ and the biasing of position perception by vestibular information in hMT+. Another minor but intriguing observation was that the hFLE could tolerate a longer duration of the flash stimulus compared to the classical visual FLE, indicating distinct mechanisms of visual processing between these two phenomena. This study also lends more support to the application of fNIRS in multisensory neuroimaging investigations that allows vestibular stimulations of real motion.

References

1. V. S. Ramachandran and S. M. Anstis, "Illusory displacement of equiluminous kinetic edges," *Perception* **19**(5), 611–616 (1990).
2. J. Müsseler, S. Stork, and D. Kerzel, "Comparing mislocalizations with moving stimuli: the Fröhlich effect, the flash-lag, and representational momentum," *Vis. Cognit.* **9**(1–2), 120–138 (2002).
3. X. Dong, J. Bai, and M. Bao, "Robust size illusion produced by expanding and contracting flow fields," *Vis. Res.* **133**, 87–94 (2017).
4. D. M. Mackay, "Perceptual stability of a stroboscopically lit visual field containing self-luminous objects," *Nature* **181**(4607), 507–508 (1958).
5. R. Nijhawan, "Motion extrapolation in catching," *Nature* **370**(6487), 256–257 (1994).
6. D. Whitney, "The influence of visual motion on perceived position," *Trends Cogn. Sci.* **6**(5), 211–216 (2002).
7. J. Schlag et al., "Extrapolating movement without retinal motion," *Nature* **403**(6765), 38–39 (2000).
8. G. W. Maus et al., "The perceived position of moving objects: transcranial magnetic stimulation of area MT+ reduces the flash-lag effect," *Cereb. Cortex* **23**(1), 241–247 (2013).
9. W. Wang et al., "Facilitation and inhibition effects of anodal and cathodal tDCS over areas MT+ on the flash-lag effect," *J. Neurophysiol.* **128**(1), 239–248 (2022).
10. R. B. H. Tootell et al., "Visual motion aftereffect in human cortical area MT revealed by functional magnetic resonance imaging," *Nature* **375**(6527), 139–141 (1995).
11. H. Théoret et al., "Repetitive transcranial magnetic stimulation of human area MT/V5 disrupts perception and storage of the motion aftereffect," *Neuropsychologia* **40**(13), 2280–2287 (2002).
12. M. Castelo-Branco et al., "Type of featural attention differentially modulates hMT+ responses to illusory motion after effects," *J. Neurophysiol.* **102**(5), 3016–3025 (2009).

13. H. Hogendoorn and F. A. J. Verstraten, "Decoding the motion aftereffect in human visual cortex," *Neuroimage* **82**, 426–432 (2013).

14. K.-Y. Li et al., "Cognitive and neural mechanisms of implied motion," *Prog. Biochem. Biophys.* **42**(6), 519–532 (2015).

15. J. Fischer, N. Spotswood, and D. Whitney, "The emergence of perceived position in the visual system," *J. Cogn. Neurosci.* **23**(1), 119–136 (2011).

16. G. W. Maus, J. Fischer, and D. Whitney, "Motion-dependent representation of space in area MT+," *Neuron* **78**(3), 554–562 (2013).

17. C. Senior, J. Ward, and A. S. David, "Representational momentum and the brain: an investigation into the functional necessity of V5/MT," *Vis. Cogn.* **9**(1–2), 81–92 (2002).

18. P. V. McGraw, V. Walsh, and B. T. Barrett, "Motion-sensitive neurones in V5/MT modulate perceived spatial position," *Curr. Biol.* **14**(12), 1090–1093 (2004).

19. D. Whitney et al., "Visually guided reaching depends on motion area MT+," *Cereb. Cortex* **17**(11), 2644–2649 (2007).

20. E. R. Ferrè and P. Haggard, "Vestibular cognition: state-of-the-art and future directions," *Cogn. Neuropsychol.* **37**(7–8), 413–420 (2020).

21. Y. Gu, "Vestibular signals in primate cortex for self-motion perception," *Curr. Opin. Neurobiol.* **52**, 10–17 (2018).

22. H. T. Karim et al., "Neuroimaging to detect cortical projection of vestibular response to caloric stimulation in young and older adults using functional near-infrared spectroscopy (fNIRS)," *Neuroimage* **76**, 1–10 (2013).

23. M. Dieterich and T. Brandt, "Functional brain imaging of peripheral and central vestibular disorders," *Brain* **131**(10), 2538–2552 (2008).

24. X. He et al., "Beyond motion extrapolation: vestibular contribution to head-rotation-induced flash-lag effects," *Psychol. Res.* **86**(7), 2083–2098 (2022).

25. B. M. Seemungal et al., "Vestibular activation differentially modulates human early visual cortex and V5/MT excitability and response entropy," *Cereb. Cortex* **23**(1), 12–19 (2013).

26. F. Aedo-Jury et al., "Antero-posterior vs. lateral vestibular input processing in human visual cortex," *Front. Integr. Neurosci.* **14**, 43 (2020).

27. T. Stephan et al., "Functional MRI of galvanic vestibular stimulation with alternating currents at different frequencies," *Neuroimage* **26**(3), 721–732 (2005).

28. H. M. Della-Justina et al., "Interaction of brain areas of visual and vestibular simultaneous activity with fMRI," *Exp. Brain Res.* **233**(1), 237–252 (2015).

29. R. M. Rühl et al., "Functional correlate and delineated connectivity pattern of human motion aftereffect responses substantiate a subjacent visual-vestibular interaction," *Neuroimage* **174**, 22–34 (2018).

30. I. Indovina et al., "Structural connectome and connectivity lateralization of the multimodal vestibular cortical network," *Neuroimage* **222**, 117247 (2020).

31. F. Faul et al., "G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences," *Behav. Res. Methods* **39**(2), 175–191 (2007).

32. D. H. Brainard, "The psychophysics toolbox," *Spat. Vis.* **10**(4), 433–436 (1997).

33. N. T. Nguyen et al., "Cerebral hemodynamic responses to the sensory conflict between visual and rotary vestibular stimuli: an analysis with a multichannel near-infrared spectroscopy (NIRS) system," *Front. Hum. Neurosci.* **14**, 125 (2020).

34. D. Carius et al., "Hemodynamic response alteration as a function of task complexity and expertise: an fNIRS study in jugglers," *Front. Hum. Neurosci.* **10**, 126 (2016).

35. T. Hashimoto, Y. Minagawa-Kawai, and S. Kojima, "Motion illusion activates the visual motion area of the brain: a near-infrared spectroscopy (NIRS) study," *Brain Res.* **1077**(1), 116–122 (2006).

36. J. Mazziotta et al., "A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM)," *Philos. Trans. R. Soc. London. Ser. B Biol. Sci.* **356**(1412), 1293–1322 (2001).

37. D. L. Collins et al., "Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space," *J. Comput. Assist. Tomogr.* **18**(2), 192–205 (1994).

38. A. C. Evans et al., "3D statistical neuroanatomical models from 305 MRI volumes," in *IEEE Conf. Record Nucl. Sci. Symp. and Med. Imaging Conf.*, Vol. 3, pp. 1813–1817 (1993).

39. D. W. Shattuck et al., "Construction of a 3D probabilistic atlas of human cortical structures," *Neuroimage* **39**(3), 1064–1080 (2008).

40. R. H. Cai et al., "Vestibular signals can distort the perceived spatial relationship of retinal stimuli," *Exp. Brain Res.* **135**(2), 275–278 (2000).

41. C. R. L. Cantor and C. M. Schor, "Stimulus dependence of the flash-lag effect," *Vis. Res.* **47**(22), 2841–2854 (2007).

42. D. M. Eagleman and T. J. Sejnowski, "The position of moving objects," *Science* **289**(5482), 1107a (2000).

43. B. Krekelberg and M. Lappe, "Temporal recruitment along the trajectory of moving objects and the perception of position," *Vis. Res.* **39**(16), 2669–2679 (1999).
44. M. Lappe and B. Krekelberg, "The position of moving objects," *Perception* **27**(12), 1437–1449 (1998).
45. T. L. Hubbard, "The flash-lag effect and related mislocalizations: findings, properties, and theories," *Psychol. Bull.* **140**(1), 308–338 (2014).
46. X. Hou et al., "NIRS-KIT: a MATLAB toolbox for both resting-state and task fNIRS data analysis," *Neurophotonics* **8**(1), 010802 (2021).
47. K. J. Friston et al., *Statistical Parametric Mapping: The Analysis of Functional Brain Images*, Academic Press, Cambridge (2007).
48. F. A. Fishburn et al., "Temporal derivative distribution repair (TDDR): a motion correction method for fNIRS," *Neuroimage* **184**, 171–179 (2019).
49. R. Huang et al., "Motion artifacts removal and evaluation techniques for functional near-infrared spectroscopy signals: a review," *Front. Neurosci.* **16**, 878750 (2022).
50. C. P. Hoyniak et al., "Adversity is linked with decreased parent-child behavioral and neural synchrony," *Dev. Cogn. Neurosci.* **48**, 100937 (2021).
51. F. A. Fishburn et al., "Irritability uniquely predicts prefrontal cortex activation during preschool inhibitory control among all temperament domains: a LASSO approach," *Neuroimage* **184**, 68–77 (2019).
52. M. A. Yücel et al., "Best practices for fNIRS publications," *Neurophotonics* **8**(1), 012101 (2021).
53. M. Hubert and E. Vandervieren, "An adjusted boxplot for skewed distributions," *Comput. Stat. Data Anal.* **52**(12), 5186–5201 (2008).
54. S. Verboven and M. Hubert, "LIBRA: a MATLAB library for robust analysis," *Chemom. Intell. Lab. Syst.* **75**(2), 127–136 (2005).
55. S. Verboven and M. Hubert, "MATLAB library LIBRA," *Wiley Interdiscip. Rev. Comput. Stat.* **2**(4), 509–515 (2010).
56. Y. Hoshi, N. Kobayashi, and M. Tamura, "Interpretation of near-infrared spectroscopy signals: a study with a newly developed perfused rat brain model," *J. Appl. Physiol.* **90**(5), 1657–1662 (2001).
57. G. Strangman et al., "A quantitative comparison of simultaneous BOLD fMRI and NIRS recordings during functional brain activation," *Neuroimage* **17**(2), 719–731 (2002).
58. M. Wolf et al., "Different time evolution of oxyhemoglobin and deoxyhemoglobin concentration changes in the visual and motor cortices during functional stimulation: a near-infrared spectroscopy study," *Neuroimage* **16**(3), 704–712 (2002).
59. L. Chen et al., "Distinct contributions of genes and environment to visual size illusion and the underlying neural mechanism," *Cereb. Cortex* **32**(5), 1014–1023 (2022).
60. S. Bae, Y. Lee, and P. Chang, "There is no test–retest reliability of brain activation induced by robotic passive hand movement: a functional NIRS study," *Brain Behav.* **10**(10), e01788 (2020).
61. H. Takakura et al., "Cerebral hemodynamic responses during dynamic posturography: analysis with a multi-channel near-infrared spectroscopy system," *Front. Hum. Neurosci.* **9**, 620 (2015).
62. Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: a practical and powerful approach to multiple testing," *J. R. Stat. Soc. Ser. B* **57**(1), 289–300 (1995).
63. A. H. Wertheim and G. Reymond, "Neural noise distorts perceived motion: the special case of the freezing illusion and the Pavard and Berthoz effect," *Exp. Brain Res.* **180**(3), 569–576 (2007).
64. A. A. Faisal, L. P. J. Selen, and D. M. Wolpert, "Noise in the nervous system," *Nat. Rev. Neurosci.* **9**(4), 292–303 (2008).
65. A. C. Huk, R. F. Dougherty, and D. J. Heeger, "Retinotopy and functional subdivision of human areas MT and MST," *J. Neurosci.* **22**(16), 7195–7205 (2002).
66. A. Malikovic et al., "Cytoarchitectonic analysis of the human extrastriate cortex in the region of V5/MT+: a probabilistic, stereotaxic map of area hOc5," *Cereb. Cortex* **17**(3), 562–574 (2006).
67. H. Kolster, R. Peeters, and G. A. Orban, "The retinotopic organization of the human middle temporal area MT/V5 and its cortical neighbors," *J. Neurosci.* **30**(29), 9801–9820 (2010).
68. B. Wild and S. Treue, "Primate extrastriate cortical area MST: a gateway between sensation and cognition," *J. Neurophysiol.* **125**(5), 1851–1882 (2021).
69. W. K. Page and C. J. Duffy, "Heading representation in MST: sensory interactions and population encoding," *J. Neurophysiol.* **89**(4), 1994–2013 (2003).
70. C. J. Duffy, "MST neurons respond to optic flow and translational movement," *J. Neurophysiol.* **80**(4), 1816–1827 (1998).
71. Y. Gu et al., "Visual and nonvisual contributions to three-dimensional heading selectivity in the medial superior temporal area," *J. Neurosci.* **26**(1), 73–85 (2006).
72. Y. Gu, G. C. DeAngelis, and D. E. Angelaki, "A functional link between area MSTd and heading perception based on vestibular signals," *Nat. Neurosci.* **10**(8), 1038–1047 (2007).

73. Y. Gu, D. E. Angelaki, and G. C. DeAngelis, "Neural correlates of multisensory cue integration in macaque MSTd," *Nat. Neurosci.* **11**(10), 1201–1210 (2008).

74. K. H. Britten and R. J. A. Van Wezel, "Electrical microstimulation of cortical area MST biases heading perception in monkeys," *Nat. Neurosci.* **1**(1), 59–63 (1998).

75. T. Brandt et al., "Reciprocal inhibitory visual-vestibular interaction. Visual motion stimulation deactivates the parieto-insular vestibular cortex," *Brain* **121**(9), 1749–1758 (1998).

76. J. Conrad et al., "Recovery from spatial neglect with intra- and transhemispheric functional connectivity changes in vestibular and visual cortex areas: a case study," *Front. Neurol.* **9**, 112 (2018).

77. S. Bense et al., "Metabolic changes in vestibular and visual cortices in acute vestibular neuritis," *Ann. Neurol.* **56**(5), 624–630 (2004).

78. C. Lopez and O. Blanke, "The thalamocortical vestibular system in animals and humans," *Brain Res. Rev.* **67**(1–2), 119–146 (2011).

79. A. M. Wirth et al., "White matter connectivity of the visual–vestibular cortex examined by diffusion-weighted imaging," *Brain Connect.* **8**(4), 235–244 (2018).

80. F. Lacquaniti et al., "Visual gravitational motion and the vestibular system in humans," *Front. Integr. Neurosci.* **7**, 101 (2013).

81. V. Kirsch et al., "Handedness-dependent functional organizational patterns within the bilateral vestibular cortical network revealed by fMRI connectivity based parcellation," *Neuroimage* **178**, 224–237 (2018).

82. C. Helmchen et al., "Increased brain responsivity to galvanic vestibular stimulation in bilateral vestibular failure," *NeuroImage Clin.* **24**, 101942 (2019).

83. H. Karim et al., "Functional brain imaging of multi-sensory vestibular processing during computerized dynamic posturography using near-infrared spectroscopy," *Neuroimage* **74**, 318–325 (2013).

84. F. Fiori et al., "The right temporoparietal junction plays a causal role in maintaining the internal representation of verticality," *J. Neurophysiol.* **114**(5), 2983–2990 (2015).

85. R. W. Sperry, "Neural basis of the spontaneous optokinetic response produced by visual inversion," *J. Comp. Physiol. Psychol.* **43**(6), 482–489 (1950).

86. E. von Holst and H. Mittelstaedt, "Das Reafferenzprinzip," *Naturwissenschaften* **37**(20), 464–476 (1950).

87. D. Subramanian, A. Alers, and M. A. Sommer, "Corollary discharge for action and cognition," *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **4**(9), 782–790 (2019).

88. S. Becker-Bense et al., "Direct comparison of activation maps during galvanic vestibular stimulation: a hybrid H2[15 O] PET—BOLD MRI activation study," *PLoS One* **15**(5), e0233262 (2020).

89. M. Dieterich, V. Kirsch, and T. Brandt, "Right-sided dominance of the bilateral vestibular system in the upper brainstem and thalamus," *J. Neurol.* **264**(S1), 55–62 (2017).

90. C. Lopez, O. Blanke, and F. W. Mast, "The human vestibular cortex revealed by coordinate-based activation likelihood estimation meta-analysis," *Neuroscience* **212**, 159–179 (2012).

91. J. W. Philbeck et al., "Asymmetrical perception of body rotation after unilateral injury to human vestibular cortex," *Neuropsychologia* **44**(10), 1878–1890 (2006).

92. M. Dieterich, "Dominance for vestibular cortical function in the non-dominant hemisphere," *Cereb. Cortex* **13**(9), 994–1007 (2003).

93. Q. Arshad et al., "Left cathodal trans-cranial direct current stimulation of the parietal cortex leads to an asymmetrical modulation of the vestibular-ocular reflex," *Brain Stimul.* **7**(1), 85–91 (2014).

94. W. O. Guldin and O.-J. Grüsser, "Is there a vestibular cortex?" *Trends Neurosci.* **21**(6), 254–259 (1998).

95. Z. Cheng and Y. Gu, "Vestibular system and self-motion," *Front. Cell. Neurosci.* **12**, 456 (2018).

96. A. T. Smith et al., "Distributed visual–vestibular processing in the cerebral cortex of man and Macaque," *Multisens. Res.* **30**(2), 91–120 (2017).

97. S. L. Novi et al., "Integration of spatial information increases reproducibility in functional near-infrared spectroscopy," *Front. Neurosci.* **14**, 746 (2020).

98. A. Machado et al., "Optimal positioning of optodes on the scalp for personalized functional near-infrared spectroscopy investigations," *J. Neurosci. Methods* **309**, 91–108 (2018).

99. F. Herold et al., "Applications of functional near-infrared spectroscopy (fNIRS) neuroimaging in exercise–cognition science: a systematic, methodology-focused review," *J. Clin. Med.* **7**(12), 466 (2018).

100. K. N. de Winkel et al., "Neural correlates of decision making on whole body yaw rotation: an fNIRS study," *Neurosci. Lett.* **654**, 56–62 (2017).

Xin He received his PhD in cognitive neuroscience in 2021 from the Institute of Psychology, Chinese Academy of Sciences. His research interests include visual-vestibular interaction and the attentional effect on ocular dominance plasticity.

Min Bao received his BA degree in biology and his PhD in cognitive psychology from the School of Life Sciences, University of Science and technology of China. In 2012, he established the Visual Perception and Plasticity Laboratory in the Institute of Psychology, Chinese Academy of Sciences. The research focuses of the lab include ocular dominance plasticity, visual adaptation, and visual-vestibular interactions, especially using altered or virtual reality.