ORIGINAL ARTICLE

Beyond motion extrapolation: vestibular contribution to head-rotation-induced flash-lag effects

Xin He[1,2] · Jianying Bai[1,3,4] · Yi Jiang[2,5,6] · Tao Zhang[2,5] · Min Bao[1,2,5]

Abstract

The perceived position of a flash aligned with a moving object usually lags behind that object. This illusion is well known as the flash-lag effect. Interestingly, head rotation alone can also induce a flash-lag effect. To date, the underlying mechanism for the head-rotation-induced flash-lag effect remains unclear. Using a virtual reality approach, we examined the contribution of vestibular signal processing in producing the effect. We found that vestibular, rather than kinesthetic, signal processing is critical for this type of flash-lag effect to occur. When head rotation induced a stationary reference stimulus in space to move on the retina, we observed a flash-lead effect relative to the reference (or a flash-lag effect relative to the head). Moreover, after a short-term adaptation training on a novel association between head rotation and retinal motion, the direction of the flash-lag effect was consistent with the newly trained association. These findings disagree with a previous account extended from the influential motion extrapolation hypothesis. Rather, they support a cross-modal bias hypothesis that the visual-vestibular associations developed from multisensory experiences may generate biasing visual signals in the associated direction with the vestibular signals, which help produce the head-rotation-induced flash-lag effects. Our findings may provide new insight into other multisensory integration phenomena.

Introduction

For millions of years, human brains usually processed visual input signals when the head or body was moving. However, in modern vision researches, human observers often sit in front of a screen to complete a perceptual task with their heads stabilized. Does the brain process visual information in the same way when the observer's head or body is moving or still? Imagine a soldier is rotating the head when searching for enemies in a jungle. Suddenly, the soldier notices a flash which could be a reflection of light on an enemy's weapon. If the soldier is facing the flash with the head stationary, it would be easy for the soldier to accurately tell the position of the flash. However, since the soldier's head is rotating, whether the soldier can make the same accurate judgment as when the soldier is stationary largely depends on the answer to the above question.

Unfortunately, the answer to this question is probably no. As we know, sensory inputs are arguably full of uncertainties due to both the physical nature of stimuli and the noisy responses of neurons to them (Faisal et al., 2008). Because multisensory signals can interact, these uncertainties can cause perception in one modality to be distorted by signals in another (Cuturi & MacNeilage, 2014; Shams et al., 2000). For example, based on the classic illusion of flash-lag effect in which a moving target is perceived ahead of a non-moving flash stimulus presented just beside it (Mackay, 1958; Nijhawan, 1992), several cross-modal variants of the illusion have been found, including visuo-auditory (Alais & Burr, 2003; Hayashi & Murakami, 2019), visuo-haptic (Drewing

✉ Min Bao
baom@psych.ac.cn

[1] CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, 16 Lincui Road, Chaoyang District, Beijing 100101, People's Republic of China

[2] Department of Psychology, University of Chinese Academy of Sciences, Beijing, People's Republic of China

[3] Xinjiang Astronomical Observatory, Chinese Academy of Sciences, Urumqi, People's Republic of China

[4] University of Chinese Academy of Sciences, Beijing, People's Republic of China

[5] State Key Laboratory of Brain and Cognitive Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, People's Republic of China

[6] CAS Center for Excellence in Brain Science and Intelligence Technology, Shanghai, People's Republic of China

et al., 2018), visuo-motor (Nijhawan & Kirschfeld, 2003), or purely in non-visual modalities (Alais & Burr, 2003; Cellini et al., 2016). Among them, one fascinating observation is that a brief flash aligned with a vertical bar stationary in the visual field is also perceived as lagging behind the bar during the observer's horizontal head rotation with steady fixation of the bar (Schlag et al., 2000). Apparently, this head-rotation-induced flash-lag effect is unrelated to the movement across the retina (hereinafter referred to as "retinal motion") because the reference bar remained stationary in the observer's visual field. This makes Schlag et al.'s finding a distinct phenomenon from the classic flash-lag effect. The goal of the present study was to disclose the distinct mechanisms for the head-rotation-induced flash-lag effect that does not contribute to the classic flash-lag effect. Considering that vestibular signals represent movement of the head in space, vestibular processing presumably plays a role in generating the phenomenon.

Schlag and colleagues, who first report the head-rotation-induced flash-lag effect (Schlag et al., 2000), explained it with the motion extrapolation hypothesis that the neural system infers the movement of the bar, thus shifts its position forward to counteract the influence of neural delays in processing moving objects (Nijhawan, 1994). In fact, this classic motion extrapolation account for the flash-lag effect can apply to various kinds of reference frames in which the movement of the target can be perceived, regardless of the reference frame being the retina or the external space (Schlag et al., 2000). Therefore, the account provides no detailed information whether and how vestibular inputs and processing contribute to this type of flash-lag effect. Besides vestibular signals, kinesthetic signals might also play a role, which can represent movements of the body parts relative to one another. Furthermore, head rotation involves active motor planning and commands. To date, it remains unknown to what extent the head-rotation-induced flash-lag effect is related to these factors. Our Experiment 1 was designed to address this issue. We measured the flash-lag effect induced by self-movement (head or torso) but without retinal motion. For a more adequate control of stimulus presentation, we adopted our previously developed virtual reality approach (Bai et al., 2019) in which visual stimuli were presented on a head-mounted display (HMD) system with a 3-Space sensor (Yost Labs Inc.) attached.

The reference bar in Schlag et al.'s experiment was LEDs fixed 36 cm in front of the eyes, and participants were instructed to gaze on the bar during head rotation in darkness. Such stimuli are rarely found in real life. A more common visual experience occurring every day is that static objects in the world leave retinal motion signals in the opposite direction to our head rotation. Suppose we are rotating our heads to the left, and a flash is now physically aligned with a vertical bar that is static in space. It is easy to know that the bar would leave rightward motion signals on our retinas. But would the flash be perceived to lag behind or be beyond the bar? Or we may ask the question in another way: in which direction would the head-motion-induced retinal motion be extrapolated?

Our Experiment 2 particularly tested this condition. Since the stimulus presentation mimics this more natural visual experience, the result of Experiment 2 can to some extent reflect how our brains represent this kind of retinal motion signals that occur thousands of times in everyday life. This question can be made more concrete when we return to the soldier example. The soldier perceives a flash right above a flower when rotating the head in the jungle. Does that mean the flash is physically above the flower? If the flash is physically above the flower, the classic retinal motion extrapolation hypothesis would predict that the flash lags behind the flower which leaves motion signals on the retina in the opposite direction to the head rotation. Thus, the direction of the flash-lag effect would be consistent with the direction of head rotation, which is opposite to Schlag et al.'s finding. By contrast, Schlag et al.'s environmental motion extrapolation hypothesis predicts no flash-lag effect to occur, since the soldier's brain may infer that the head-rotation-induced moving flower, though moving on the retina, is actually static in the jungle.

To account for the contradiction between our observation in Experiment 2 (i.e., the flash perceptually appeared ahead of the moving reference bar when the observers rotated their heads) and the predictions of the motion extrapolation hypotheses, we turned to the "cross-modal bias hypothesis" which was proposed in our previous work to account for the vestibular modulation on the processing of motion aftereffect (Bai et al., 2020). In that work, participants were required to compare the velocity of motion aftereffects of two gratings. The illusory motion direction of one of the two gratings was consistent with the head rotation, while the other opposite. We found that the motion aftereffect with the opposite direction to head rotation was perceived as moving faster than that with the same direction as head rotation. The cross-modal bias hypothesis argues that living in a multisensory environment leads the brain to develop an association between different senses, e.g., signals from the visual and vestibular pathways. As a result, weak biasing visual signals in the associated direction can spontaneously emerge with the input of vestibular signals, thus modulating the velocity of the motion aftereffect (Bai et al., 2020). In the present study, we hypothesize that the biasing visual signals can also be responsible for the head-rotation-induced flash-lag effect. To further validate this hypothesis, in Experiment 3 we trained people to associate a head turn with retinal motion in an unusual direction. Consistent with our hypothesis, we found a head-rotation-induced flash-lag effect when retinal motion was rendered perpendicular to head rotation.

Experiment 1

It is known that vestibular signals carry information about movement of the head in space, and kinesthetic signals represent movement of the body parts relative to one another. Thus, during head rotation, there were both vestibular and kinesthetic inputs that might contribute to the head-rotation-induced flash-lag effect. To understand the role of the vestibular and kinesthetic signals in this type of flash-lag effect, we included three conditions in Experiment 1: a head-movement (HM) condition which retained both vestibular and kinesthetic inputs, a body-movement (BM) condition with unchanged vestibular inputs but substantially reduced kinesthetic signals (especially for those representing movement of the head relative to the torso), and a torso-movement (TM) condition which eliminated the vestibular signals but preserved the kinesthetic signals.
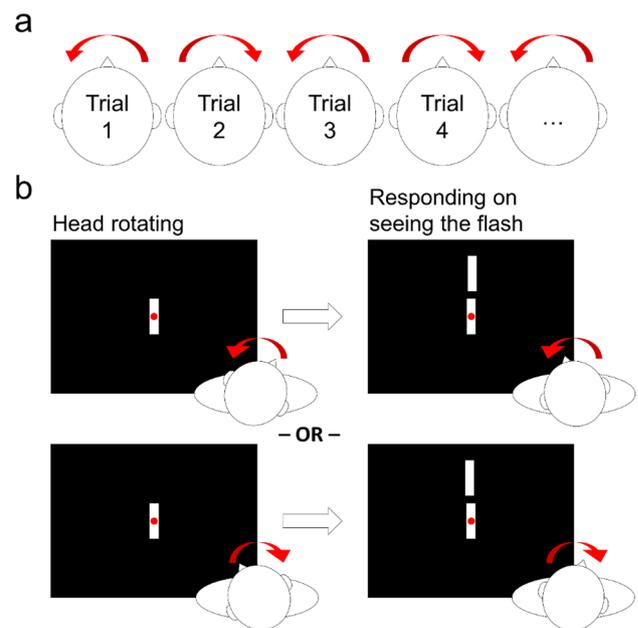
Method

Participants

Twenty participants (age range 19–29 years; 9 males and 11 females) participated in both Experiment 1 and Experiment 2. The number of participants exceeded the suggested sample size (11 participants) by G*Power based on the result of the pilot experiment in the pure-black-background (BB) condition.

All our experimental procedures in the present study were approved by the Institutional Review Board of the Institute of Psychology, Chinese Academy of Sciences, and informed consent was obtained from all participants. All participants had normal or corrected-to-normal vision.

Materials and apparatus

Before the formal experiments, we first conducted a pilot experiment to ensure that Schlag et al.'s (2000) main finding could be observed using our virtual reality experimental setup (see below), since Schlag et al. used LED for the reference and flash stimuli in otherwise complete darkness. It turned out that our pilot experiment replicated Schlag et al.'s finding that head rotation led to a brief flashed bar to be perceived as lagging behind a vertical bar fixed in front of the eyes when the two bars were physically aligned (see the Appendix for details).

Stimuli were presented on a Sony HMZ-T3 (Sony Corp., Tokyo, Japan) head-mounted goggles (50° × 28° visual angle, 1280 × 720 pixel resolution at 60 Hz) connected to a Dell XPS 8700 (Dell, Round Rock, TX, USA) computer, and programmed in MATLAB (The MathWorks, Natick,

Fig. 1 Procedures of the head-movement (HM) condition in Experiment 1. **a** The head-movement procedure in a session. Participants started each session by turning their heads from one side to the other, and then rotated back and forth. **b** The procedure of a single trial. A vertical white bar was fixed at the center of the screen from the beginning of a trial and during participant's head rotation (left half of the figure), serving as the reference. At a certain time in the middle of the head rotation (right half of the figure), an identical white bar flashed above the stationary bar. Participants were required to report, on seeing the flashed bar, whether it appeared to the left or right of the reference bar by pressing the corresponding arrow key. For illustration purpose, sizes of the stimuli in the graphs are not proportional to the actual. The depiction of head movement is from a top view. The ellipse represents the participant's shoulders

MA) and Psychtoolbox (Brainard, 1997). A 3-Space Sensor (TSS-WL Sensor, YEI technology, Portsmouth, OH, USA), which was used to record the participant's movement data in real time, was attached on top of the goggles or on one armrest of the swivel chair, depending on the experimental condition. The same set of equipment was used in all the experiments of the current study.

Design and procedure

All the three experimental conditions, HM, BM, and TM, were completed by each participant with identical displays and tasks but different states of self-motion. All the experimental conditions were presented in a random order for each participant. In the HM condition (see Fig. 1), participants pressed spacebar on a keyboard to start a trial, then immediately started turning their heads from the rightmost (leftmost) side to the left (right) in the horizontal plane. Each trial contained 1 head rotation. A white stationary vertical bar (also referred to as the reference bar, 0.25°

Fig. 2 The self-motion to be executed in the torso-movement (TM) condition of Experiment 1 and in the torso-movement/retinal-movement (TMRM) condition of Experiment 2. The participant in the figure put his head on a chinrest and remained the bite-bar in the mouth so as to keep the head static while rotating the torso using the feet. His hands were put in the lap holding the keyboard. The photos show different views of postures with the torso facing straight ahead (Photos 1 and 2) or the rightmost (Photos 3 and 4). Note that the head was static and always facing straight ahead

width × 2.5° height) and a red fixation point (0.15°) were first centrally presented on a black background. During a head turn and between 0.5 and 3 s after the start of the trial, another white vertical bar (also 0.25° × 2.5°) flashed above the stationary bar for 1 frame (~ 16.6 ms). There was a 0.5° gap between the two bars. Physically, the flash could be presented vertically aligned with the stationary bar, or horizontally shifted by 2, 4, or 6 pixels to the left or right of the stationary bar (a single pixel subtended a visual angle of approximately 2.34 arcmin), yielding 7 spatial offsets. The stationary bar disappeared 200 ms after the onset of the flash. Participants were told to press Left Arrow or Right Arrow key on the keyboard immediately once they saw the flash to indicate whether it appeared to the left or right of the stationary bar.

The BM condition had the same configurations as the HM condition except that the participant sat inside a swivel chair with the feet off the ground, and the experimenter executed the rotation by standing behind the swivel chair and manually turning it. And in the TM condition, participants sat in the swivel chair with their head stabilized in the upright position on a chinrest with a customized bite-bar. The bite-bar was a straight stainless-steel straw fastened to the chinrest, which guaranteed the heads were immobilized while the participants rotated the torso using their feet (see Fig. 2). This design helped eliminate the vestibular input as much as possible. In addition, their hands were in the laps holding the keyboard, so the arms would not have to reach against the direction of rotation and limit the rotation range. All other configurations were the same as in HM condition (see Fig. 3). The sensor was attached to the top of the goggles in the HM and BM conditions and to one arm of the swivel chair in the TM condition.
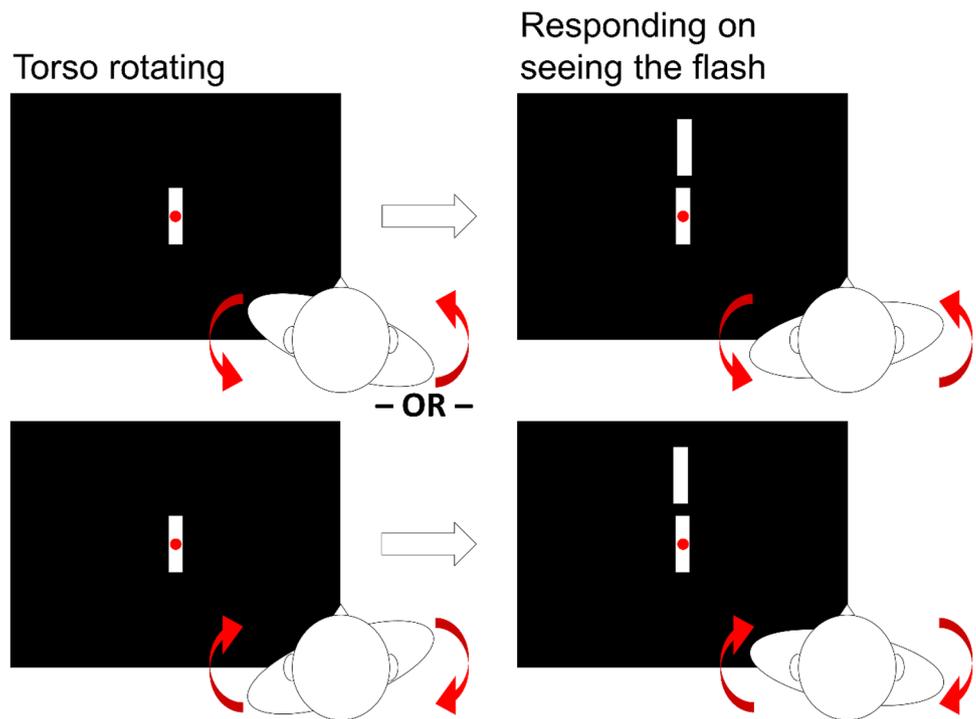
Additionally, there was a no-movement (nM) control condition, the data of which were used in both Experiments 1 and 2. Participants did not have to do any kinds of self-motion, and the reference bar also stayed static during the trials.

Each trial contained 1 rotation in the HM, BM, or TM condition. Twenty measurements were completed for each offset value, resulting in 140 trials in each condition. Each condition was completed in 2 consecutive sessions with a 5-min interval between. Participants were instructed to keep gazing on the central fixation point throughout each trial.

Statistical analysis

Points of subjective alignment (PSAs) were estimated from the fits with a logistic function ($P = \frac{1}{1+e^{-a(x-\text{PSA})}}$) on the probabilities ($P$) that participants perceived the flash as shifted opposite the head movement ("lagging") by the seven preset spatial offset values (x). Here, we defined positive values of x as physically shifted to the opposite direction to the head rotation and vice versa. PSAs were reported together with the 95% confidence intervals (CIs). One-sample t-tests (two-tailed) were used to compare the PSAs with zero. Paired t-tests (two-tailed) and repeated-measure ANOVAs (rANOVAs) were used to compare the

Fig. 3 Procedure of the torso-movement (TM) condition in Experiment 1. It was similar to that in HM condition (see Fig. 1) except that the participant had to turn the torso while keeping the head static

PSAs between the conditions. Effect size was reported as Cohen's d for one-sample t-test, d_z for paired t-test calculated with G*Power (Faul et al., 2007), and partial eta squared (η_p^2) for rANOVA via the formula (see Keppel, 1991):

$$\eta_p^2 = \frac{F \cdot df_1}{F \cdot df_1 + df_2}.$$

In addition, for conditions in which participants had to rotate their head or body, we calculated the rotation velocity of the participants and its correlation to the PSA.

The statistical analyses of Experiments 2 and 3 were almost the same as above. The minor difference for Experiment 3 was that the P and x were defined not based on the opposite direction to the head rotation, but on the associated direction with the head rotation in the adaptation training (see below for details).
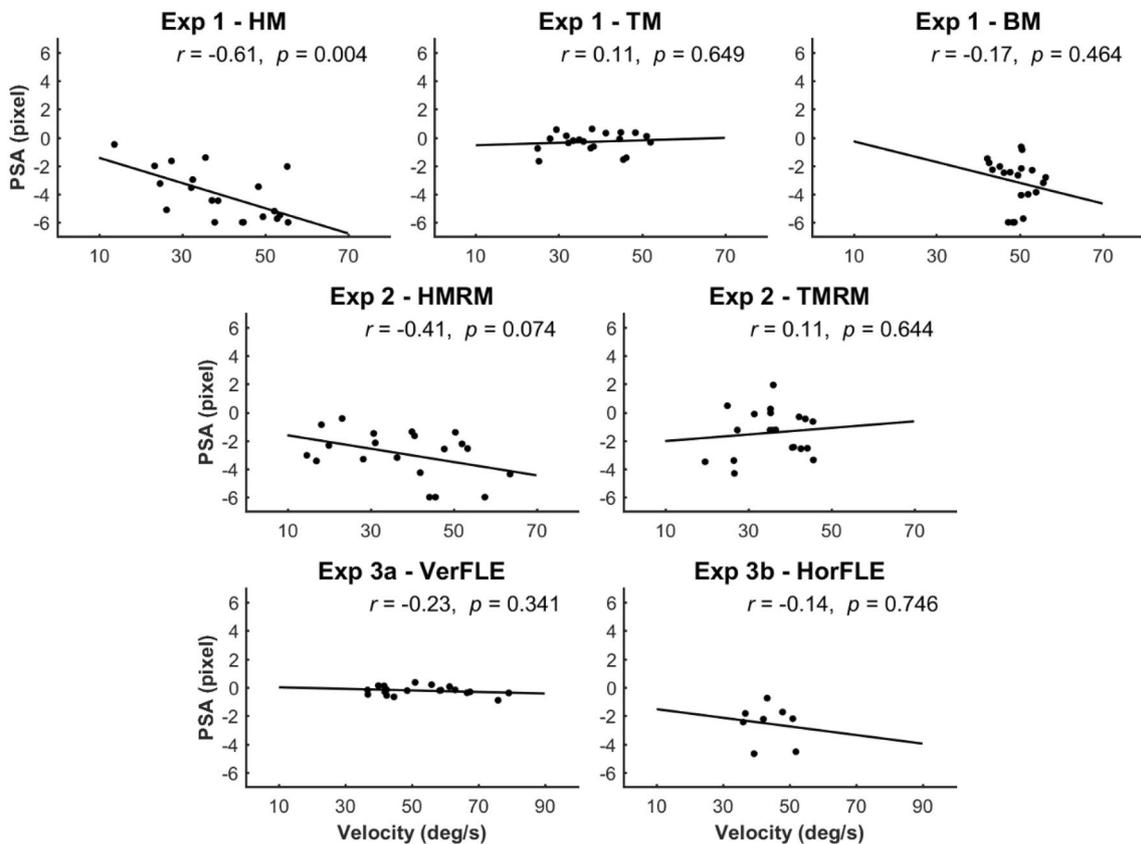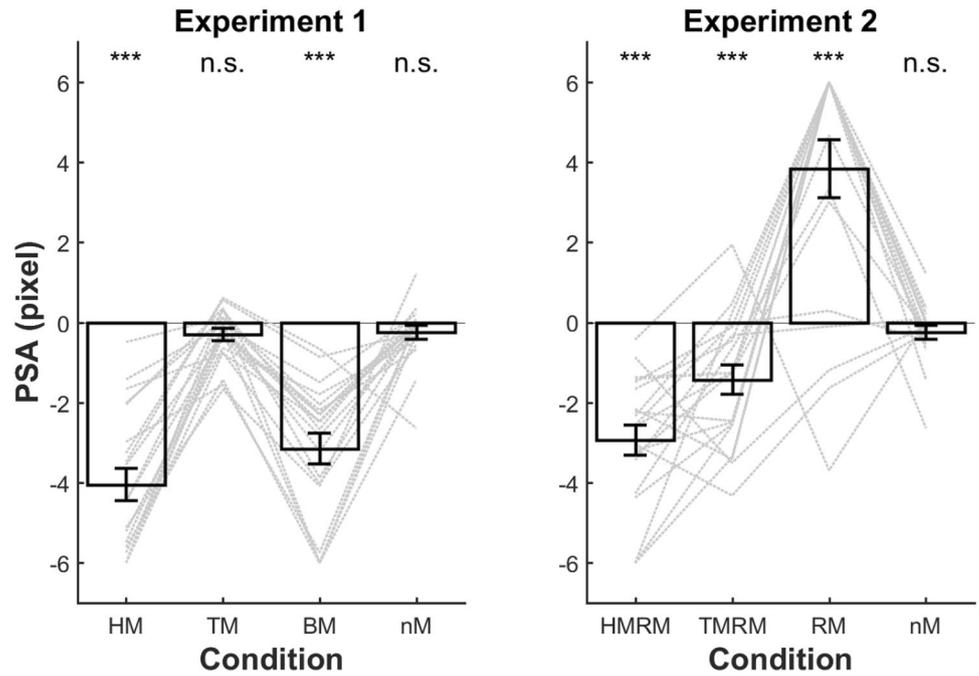
Results

As revealed by one-sample t-tests, the PSA in the nM control condition did not systematically deviate from zero (see Fig. 4; PSA $= -0.25 \pm 0.75$ pixels, $t(19) = 1.46$, $p = 0.159$, $d = 0.33$, 95% CI $= [-0.60, 0.11]$ pixels), suggesting no mislocalization bias when both the reference bar and the observer were static. Meanwhile, PSAs in both the HM (PSA $= -4.05 \pm 1.79$ pixels, $t(19) = 10.09$, $p < 10^{-8}$, $d = 2.26$, 95% CI $= [-4.89, -3.21]$ pixels) and

BM (PSA $= -3.15 \pm 1.68$ pixels, $t(19) = 8.35$, $p < 10^{-7}$, $d = 1.88$, 95% CI $= [-3.94, -2.36]$ pixels) conditions were significantly negative, indicating a head-rotation-induced flash-lag effect in both conditions (i.e., the flash physically aligned to the static reference bar was perceived to shift opposite the head movement). Though the PSA in the TM condition did not reach significance, it yielded a trend in the predicted direction (PSA $= -0.30 \pm 0.67$ pixels, $t(19) = 2.00$, $p = 0.060$, $d = 0.45$, 95% CI $= [-0.60, 0.11]$ pixels).

As for the comparison between conditions, rANOVA on PSA showed a significant main effect of Condition ($F(3, 57) = 52.52$, $p < 10^{-13}$, $\eta_p^2 = 0.73$). Specifically, the PSA of nM condition was significantly smaller than those of HM ($p < 10^{-6}$) and BM ($p < 10^{-4}$), but not TM ($p > 0.99$). Also, TM condition yielded a PSA smaller than HM and BM ($ps < 10^{-6}$). The PSAs of HM and BM did not differ significantly from each other ($p = 0.159$), suggesting that corollary discharge did not play a crucial role in the effect (Sperry, 1950; von Holst & Mittelstaedt, 1950).

Finally, we calculated the correlation between the rotation velocity and the PSA of the participants. The rotation velocity ($M \pm$ SD) was 39.36 ± 12.29 deg/s, 38.21 ± 8.35 deg/s, and 49.26 ± 3.99 deg/s in the HM, TM, and BM condition, respectively. As shown in Fig. 5, the rotation velocity was significantly negatively correlated with the PSA (i.e., positively correlated with the magnitude of the head-rotation-induced flash-lag effect) in the HM
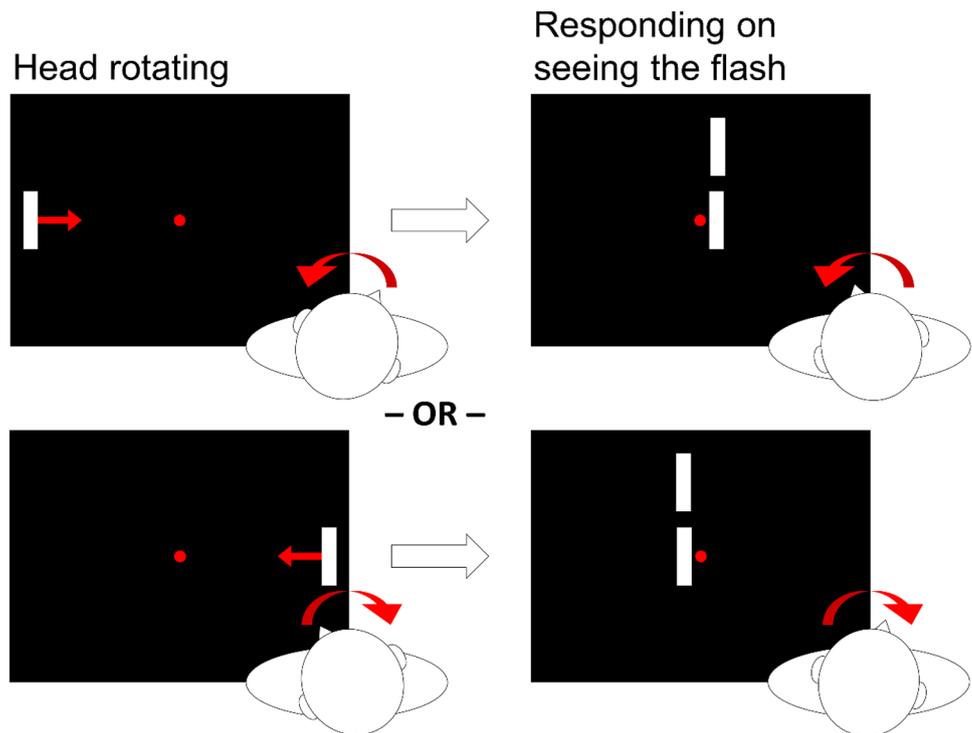
Fig. 4 Results of Experiment 1 (left) and 2 (right). Gray dashed lines represent individual data and error bars indicate 1 SEM. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. *HM* head movement, *TM* torso movement, *BM* body movement, *HMRM* head movement/retinal movement, *TMRM* torso movement/retinal movement, *RM* retinal movement, *nM* no movement

Fig. 5 Results of the correlation analysis between the rotation velocity and the PSA in conditions involving head or body rotation. Rows from upper to lower show the results for Experiments 1–3, respec-tively. *Exp* experiment, *VerFLE* vertical flash-lag effect, *HorFLE* horizontal flash-lag effect

Fig. 6 Procedure in the head-movement/retinal-movement (HMRM) condition of Experiment 2. At the beginning of a trial, participants started rotating their head from one side. Meanwhile, the reference bar appeared on the edge of the other side of the screen, and moved in the opposite direction to head rotation (left half of the figure). At the first frame that the reference bar had passed the midline (right half of the figure), an identical white bar would flash above. Participants were required to report, on seeing the flashed bar, whether it appeared to the left or right of the reference bar by pressing the corresponding arrow key. Arrows in the graphs are for demonstration purpose only, not physically presented in the experiment. Sizes of stimuli in the graphs are not proportional for illustration purpose

condition ($r = -0.61$, $p = 0.004$) while in the TM and BM conditions, there was no significant correlation ($ps > 0.4$).

The results of Experiment 1 indicated strong head-rotation-induced flash-lag effect in both the HM and BM conditions, but not the TM condition. This suggested that vestibular signals played a major role in driving this type of flash-lag effect. In contrast, the contribution from the kinesthetic signals was fairly limited.

Experiment 2

The findings of Experiment 1 indicated the significant role of vestibular signals in producing the head-rotation-induced flash-lag effect. However, it remains uncertain how to relate the contribution of vestibular inputs with Schlag et al.'s environmental motion extrapolation account (Schlag et al., 2000). Here, we tested their environmental motion extrapolation account in a condition in which the reference bar was static in the world but did have retinal motion due to head rotation, which differed from the HM condition in Experiment 1 in which there was no retinal motion. The classic retinal motion extrapolation hypothesis would predict that the flash lags behind the bar. The direction of this predicted effect is opposite to Schlag et al.'s finding. On the other hand, Schlag et al.'s environmental motion extrapolation hypothesis would predict no flash-lag effect to occur, since the bar, though moving across the retina, is considered static in the three-dimensional world.
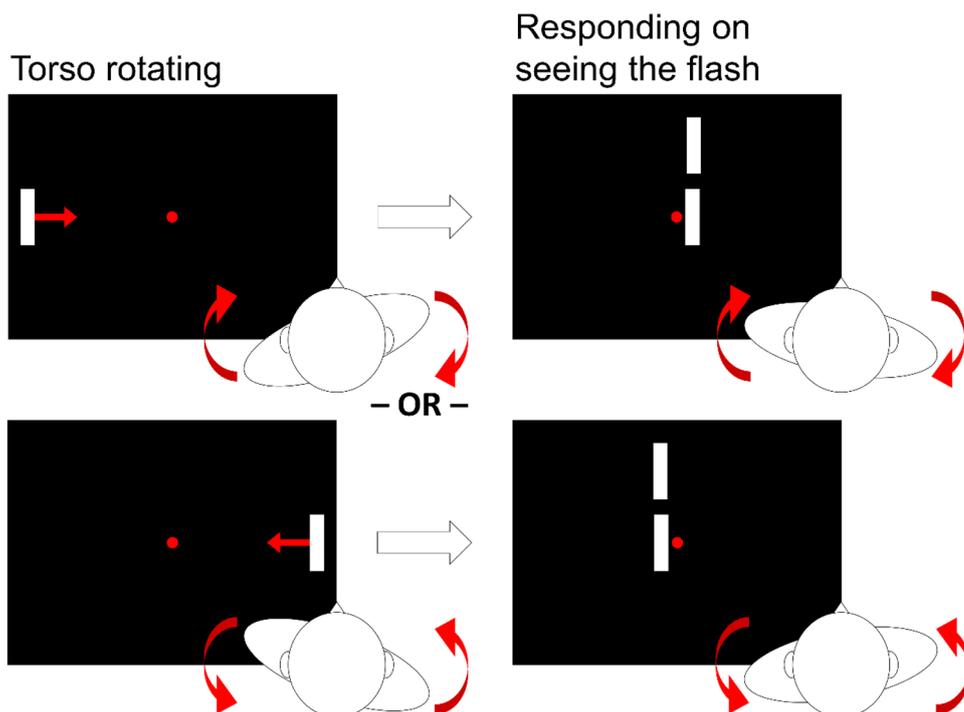
Method

Participants

The same participants as in Experiment 1 took part in Experiment 2. The order of Experiments 1 and 2 was counterbalanced across participants, and no difference in the pattern of the results was observed whether Experiment 1 or 2 was completed first.

Design and procedure

There were three within-participant experimental conditions. All the experimental conditions were presented in a random order for each participant. To realize the aforementioned situation in which the reference bar was static in the world but left motion signals across the retina due to head rotation, we designed the head-movement/retinal-movement (HMRM) condition. The procedure was similar to the HM condition in Experiment 1 except for one major difference (see Fig. 6). Participants pressed spacebar of a keyboard to start a trial and immediately started turning the head from the rightmost (leftmost) side to the left (right) while maintaining a central fixation. Each trial contained one head rotation. The reference bar was no longer stationary at the center of the screen. Instead, when participants started rotating their head from one side, the reference bar would appear on the edge of the other side of the screen and move in the opposite direction to head rotation, e.g., if a participant was turning the head from the leftmost side to the right, the reference bar would appear

Fig. 7 Procedure of the torso-movement/retinal-movement (TMRM) condition in Experiment 2. It was similar to that in HMRM condition (see Fig. 6) except that the participant had to turn their torso while keeping the head static, and that the reference bar moved in the same direction as the torso

on the right edge of the screen and move leftwards, and vice versa. In virtue of the recording by the sensor, the velocity of the reference bar was rendered the same as the head rotation in real time. At the first frame that the reference bar had already passed the vertical midline of the screen, an identical bar would flash for 16.6 ms (i.e., one frame) above the reference bar with a 0.5° gap. Then both bars would disappear simultaneously. Physically, the horizontal shift of the flashed bar relative to the reference bar could be 0 or 1, 3, or 5 pixels to the left or right, yielding 7 spatial offsets. Participants pressed Left Arrow or Right Arrow key on the keyboard immediately once they saw the flash to indicate whether it appeared to the left or right of the stationary bar.

The torso-movement/retinal-movement (TMRM) condition was approximately replacing head movements in the HMRM condition with torso movements, or in other words, adding the retinal motion on the basis of the TM condition in Experiment 1. To facilitate a direct comparison between the TMRM and HMRM conditions, we kept the same relative movement of head vs. torso between the TMRM and HMRM conditions. Therefore, the rightward rotation of the torso in the TMRM condition was seen as the leftward rotation of the head in the HMRM condition and vice versa; and since in the HMRM condition the reference bar moved oppositely to the head, here in the TMRM condition the reference bar moved in the same direction as the torso rotation (see Fig. 7).

The retinal-movement (RM) condition served as a control that examined the classical flash-lag effect without any self-motion. The reference bar automatically started moving at a random speed within 20–40 deg/s once the trial began,

which imitated the speed of head rotations. The directions were also randomized and counterbalanced across trials.

As in Experiment 1, twenty measurements were completed for each offset value, resulting in 140 trials in each condition. Each condition was completed in 2 consecutive sessions. The order of conditions was counterbalanced across participants. Participants were instructed to keep gazing on the central fixation point throughout each trial.

Results

PSAs in both the HMRM (PSA = $-$ 2.94 $\pm$ 1.67 pixels, $t(19)$ = 7.85, $p < 10^{-6}$, d = 1.76, 95% CI = [$-$ 3.72, $-$ 2.15] pixels) and TMRM (PSA = $-$ 1.43 $\pm$ 1.61 pixels, $t(19)$ = 3.97, $p < 0.001$, d = 0.88, 95% CI = [$-$ 2.18, $-$ 0.67] pixels) conditions were significantly different from zero, with most participants showing negative PSAs (see Fig. 4). That is, when physically aligned to the moving reference bar, the flash was perceived as shifted in the opposite direction to the head movement (HMRM) or the relative movement of head to torso (TMRM), i.e., ahead of the moving reference bar. In contrast, in the RM condition, the PSA was positive (PSA = 3.83 $\pm$ 3.23 pixels, $t(19)$ = 5.31, $p < 10^{-4}$, d = 1.19, 95% CI = [2.32, 5.34] pixels), which indicated a classical flash-lag effect that the flash physically aligned to the moving reference bar was perceived lagging behind it.[1]

---

[1] Readers may notice that the sign of PSA for the flash-lag effect (FLE) in the RM condition was opposite to those in the other conditions. Actually, it was challenging to devise a totally consistent sign of PSA for all the conditions in Experiments 1 and 2 because they did

Therefore, the head-movement-induced flash-lag effect in this experiment could also be deemed as a flash-lead effect (for a review of the flash-lead effect, see Hubbard, 2014) if we chose the location of the reference bar rather than the head as the reference. Note that the flash-lead effect has been observed in many studies through manipulating the luminance or contrast of the stimuli in a typical flash-lag paradigm (Arnold et al., 2009; Öğmen et al., 2004; Purushothaman et al., 1998; for reviews, see Hubbard, 2014, 2018). Unlike the current finding above, however, in these studies the flash-lead effect appeared more of a weakened and thus reversed flash-lag effect; that is, one effect could decrease and transform into the other in the same experiment merely by quantitatively altering one single dimension of the stimulus (contrast, latency, luminance, etc.) as if they were just on opposite sides of the origin on one axis.

The rANOVA again showed a significant main effect of Condition ($F(3, 57) = 38.59$, $p < 10^{-7}$, $\eta_p^2 = 0.67$). Specifically, the PSA in nM condition was significantly different than those in HMRM ($p < 10^{-6}$), TMRM ($p = 0.042$), and RM ($p < 0.001$) conditions. The amplitude of PSA in the TMRM condition was significantly smaller than that in the HMRM ($p = 0.017$) condition, and both (HMRM: $p < 10^{-5}$; TMRM: $p < 10^{-4}$) were significantly different from that in the RM.

The result in the HMRM condition was paradoxical. On the one hand, if the underlying mechanism was environmental motion extrapolation, as explained by Schlag et al. (2000), then no flash-lag effect should occur, because the moving reference was equivalent to, and likely inferred as, a stationary bar in the space due to the causal relationship between the head motion and retinal motion signals. On the other hand, if it was the retinal motion that was extrapolated, the direction of the flash-lag effect should remain consistent with the direction of head rotation. However, the results were inconsistent with the prediction of either motion extrapolation hypothesis.

The rotation velocity ($M \pm SD$) of the participants was 37.87 ± 14.45 deg/s and 35.88 ± 7.63 deg/s in the HMRM and TMRM condition, respectively. No significant correlation between the rotation velocity and the PSA was observed in either condition, but in the HMRM condition there was a trend ($r = -0.41$, $p = 0.074$; see Fig. 5).

Experiment 3

The findings of Experiment 2 cannot be explained by either Nijhawan's retinal or Schlag et al.'s environmental motion extrapolation hypothesis. To explain the findings, we advocate the cross-modal bias hypothesis proposed in our recent work (Bai et al., 2020): through long time living in a multisensory environment, the brain likely develops a strong association between signals from the visual and vestibular pathways; consequently, vestibular signals can produce a weak but effective biasing visual signal in the associated direction. Specifically, a leftward head turn may produce a weak rightward biasing visual signal in the brain. If a brief flash is now presented to the participant, the biasing visual signal would cause mislocalization of the flash in the opposite direction to the head rotation. This explains why a leftward head turn always induced a rightward mislocalization of the flash whether the reference bar was static on the retina or static in space yet moved on the retina. If this hypothesis is true, training people to associate retinal motion in unusual directions with head turns should produce a distinct flash-lag effect, the direction of which depends on the direction of the trained retinal motion.
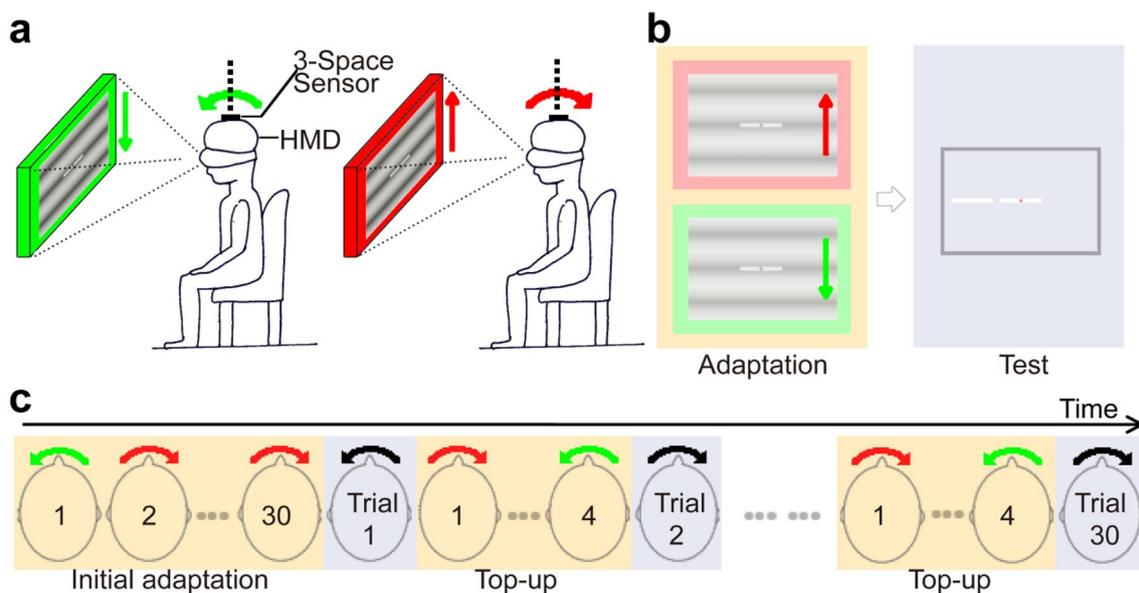
Method

Participants

Nineteen participants (age range 18–31 years; 11 males and 8 females) participated in Experiment 3a. Eight participants (age range 20–32 years; 5 males and 3 females) participated in Experiment 3b, four of which also participated in Experiment 3a. The author J.B. participated in both Experiment 3a and Experiment 3b.

Design and procedure

Experiment 3a Eleven of the participants first completed 7 head-movement (HM) sessions, and then completed 7 head-still (HS) sessions, while the other eight completed 14 sessions in each condition. A horizontal grating (0.13 cpd) was presented which occupied the whole screen. Participants started each session by turning their heads from the rightmost to the left, and then rotated back and forth on each trial. Meanwhile they were told to fixate at a central black

---

Footnote 1 (continued)

not have a common reference for FLE: in the HM condition of Experiment 1 there was only head movement but no retinal motion while in the RM condition of the Experiment 2 there was only retinal motion but no head movement, and in the HMRM condition existed both types of motion. Moreover, the retinal motion direction was opposite to the head-movement direction in the HMRM condition. So, we chose the head (or body) movement direction as the reference of FLE for conditions HM, BM, TM, HMRM, and TMRM, and exceptionally the retinal motion direction as reference for the RM condition. That was why the sign of PSA for FLE in the RM condition was opposite to those in the other 5 conditions.

Fig. 8 Stimuli in the head-movement condition of Experiment 3a. **a** The adapting stimuli. During the initial adaptation period, when the head was turning to the left (green), the grating drifted downward, and when the head turned to the right (red), the grating drifted upward. A white horizontal stationary bar was always presented during the adaptation. For demonstration purpose, the grating displays are drawn in front of the participant. In the experiment, the grating displays were actually presented on the screens of the HMD. **b** The stimuli in a top-up trial in the head-movement condition. The adapting stimuli during the top-up adaptation period (left half) were the same as in the initial adaptation period. During the test period (right half), an identical white bar flashed to the left (as shown in the example) or right of the stationary bar. Participants were required to report whether the flash was shifted above or below the stationary bar. Sizes of stimuli are not proportional for illustration purpose. **c** The procedure of a head-movement session. The initial adaptation lasted for about 1 min, followed by the first test period (i.e., Trial 1). In the example shown here, the participant finished 30 head turns. After the first trial, participants completed 27 top-up trials, with each top-up trial composed of 4 head turns and 1 test
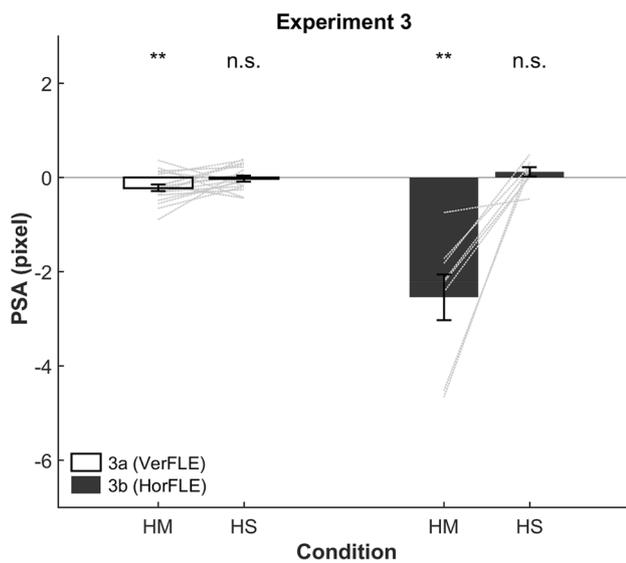
dot (0.15°). The fixation point was centered on a stationary horizontal white bar (2.5° width × 0.25° height). Participants were instructed to keep gazing at the central fixation point throughout each trial in Experiments 3a and 3b.

Each trial consisted of an adaptation period and a test phase. In the HM sessions, during the adaptation period (~ 1 min for every first trial of a session and ~ 8 s for the rest; see below for details), participants kept turning the head leftwards and rightwards alternately; when the head turned to the left, the grating drifted downward, and when the head to the right, the grating drifted upward (see Fig. 8a), through which participants adapted to the abnormal associations between the direction of head rotation and direction of retinal motion signals. The velocity of the grating was the same as that of the head rotation. The fixation point turned red during the last head rotation in this period. Immediately after the end of the adaptation period was a test phase during which the grating was removed and replaced with a black background. The head continued to turn, while at a random time (0.7–1 s after the start of the test phase) a white bar with the same size as the stationary bar flashed for 1 frame either to the left or right of the stationary bar (see Fig. 8b). Physically, the flash could be horizontally aligned with, or *vertically* shifted by 2, 4, or 6 pixels upward or downward

from, the stationary bar. There was a 0.5° gap between the flash and stationary bar. Participants continued to complete the head turn after the onset of the flash, and then judged whether the flash was above or below the stationary bar by pressing Up Arrow or Down Arrow key on the keyboard.

Once the participants made the choice, the current trial finished, the full screen grating reappeared and the fixation point disappeared. Then 500 ms later, the black fixation point reappeared, the next trial started and the participants needed to begin turning their heads while the grating drifted again. In each session, the adaptation period on the first trial, referred to as the initial adaptation, consisted of 30 head rotations (i.e., 15 back-and-forth periods) lasting for ~ 1 min in total, while in the remaining trials the adaptation period consisted of only 4 head rotations, and was referred to as the top-up adaptations (see Fig. 8c), allowing us to repeatedly measure the head-rotation-induced flash-lag effect while the participants remained adapting to the abnormal visuo-motor association.

There were 4 trials for each spatial offset, yielding 28 trials in each session. In HM sessions, whenever a head turn finished, the average drifting velocity of the adapting grating during the head turn and the time relative to the start of the session were saved. These data were used in the subsequent

Fig. 9 Result of Experiment 3. Experiment 3a tested the vertical flash-lag effect (VerFLE), whereas Experiment 3b tested the horizontal flash-lag effect (HorFLE). Gray dashed lines represent individual data and error bars indicate 1 *SEM*. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. *HM* head movement, *HS* head still

HS sessions to simulate the visual inputs in HM sessions. The procedure of HS sessions was almost the same as that of the HM sessions except that the participants did not rotate their heads but remained stationary.

Experiment 3b In this experiment, the fixation point was centered on a small white *vertical* stationary bar ($0.25°$ width × $2.5°$ Height). The flash was also a *vertical* bar of the same size as the stationary bar, displayed either above or below the stationary bar. Physically, the flash could be vertically aligned with, or *horizontally* shifted by 2, 4, or 6 pixels leftward or rightward from, the stationary bar. Other stimulus parameters and experimental procedures were identical to Experiment 3a.

Results

Experiment 3a

Interestingly, in the HM condition we observed a flash-lag effect along the vertical (i.e., perpendicular) direction that was congruent to the trained resultant retinal motion. As shown in Fig. 9, the PSA in the head-movement condition was significantly different from zero (PSA = − 0.22 ± 0.31 pixels, $t(18) = 3.13$, $p = 0.006$, $d = 0.72$, 95% CI = [− 0.37, − 0.07] pixels), showing a clear vertical flash-lag effect. By contrast, the PSA in the HS control condition was not significantly different from zero (PSA = − 0.03 ± 0.27 pixels, $t(18) = 0.52$, $p = 0.613$, $d = 0.12$, 95% CI = [− 0.16, 0.10]

pixels). When comparing the PSA between the head-movement and head-still conditions, we observed a marginal trend ($t(18) = 1.95$, $p = 0.067$, $d_z = 0.45$, 95% CI = [− 0.40, 0.01] pixels).

Experiment 3b

In Experiment 3b, we examined the typical horizontal flash-lag effect by using *vertical* flash and stationary bars. As shown in Fig. 9, the PSA in the HM condition was significantly different from 0 (PSA = -2.55 ± 1.37 pixels, $t(7) = 5.27$, $p = 0.001$, $d = 1.86$, 95% CI = [− 3.69, − 1.40] pixels), indicating a typical horizontal flash-lag effect. However, the effect was absent in the HS condition (PSA = 0.12 ± 0.28 pixels, $t(7) = 1.16$, $p = 0.282$, $d = 0.41$, 95% CI = [− 0.12, 0.35] pixels). A paired *t*-test also showed a significant difference between the two conditions ($t(7) = 4.96$, $p = 0.002$, $d_z = 1.75$, 95% CI = [− 3.93, − 1.39] pixels).

Therefore, short-term adaptation to perpendicular retinal motion triggered by head rotation could produce a vertical flash-lag effect. Although the effect was small (established by only 1-min adaptation) as compared to the horizontal flash-lag effect (established by long-term everyday experience), it cannot be explained by either version of the motion extrapolation hypothesis (Nijhawan, 1994; Schlag et al., 2000), because the retinal motion extrapolation hypothesis predicts no flash-lag effect, and the environmental motion extrapolation hypothesis predicts a horizontal rather than vertical flash-lag effect.

The rotation velocity ($M \pm SD$) of the participants was 53.28 ± 13.17 and 43.51 ± 6.18 deg/s in Experiment 3a and 3b, respectively. In both conditions the correlation between the rotation velocity and the PSA was non-significant (ps > 0.3; see Fig. 5).

General discussion

The present study tried to answer two questions regarding the mechanisms of head-rotation-induced flash-lag effect (Schlag et al., 2000) which have never been addressed in previous work.

The first question is on the role of vestibular and kinesthetic signal processing in producing the head-rotation-induced flash-lag effect. In Experiment 1, we found that the flash-lag effects between the HM and BM conditions did not statistically differ; no significant flash-lag effect was observed in the TM condition in which the vestibular inputs were eliminated but the kinesthetic signals representing movement of the head relative to the torso were preserved. These results suggest that vestibular (rather than the kinesthetic) signal processing is critical for this type of flash-lag

effect to occur. However, it is still hard to exclude with full confidence the contributions of the kinesthetic signals to the flash-lag effect, considering the marginal trend ($p = 0.062$, $d = 0.44$) in the TM condition. A larger sample size might turn this weak effect more prominent. Moreover, the significant flash-lag effect in the TMRM condition of Experiment 2 also endorsed the role of kinesthesia in the head-rotation-induced flash-lag effect.

The second question is how vestibular signal processing helps generate the head-rotation-induced flash-lag effect. To answer this question, we presented observers with a reference bar that moved at the same velocity as, but in the opposite direction to, their head rotation. Meanwhile, we measured the perceived spatial offset between the reference bar and a physically aligned flash that was presented when the reference bar had just passed the central fixation. The presentation of the reference bar simulated a natural circumstance in which a retinally moving stimulus (i.e., the reference bar) was actually stationary in the three-dimensional environment during the head rotation. We found that the flash was perceived ahead of the retinally moving reference, even though they were physically aligned. The direction of this effect is opposite to the prediction of the original retinal motion extrapolation hypothesis (Nijhawan, 1994) that the flash stimulus would be perceived as 'lagged' relative to the moving one. Moreover, the finding is inconsistent with the environmental motion extrapolation hypothesis proposed by Schlag et al. (2000). According to that account, no flash-lag effect would have been found in Experiment 1, since the reference bar would be inferred to be static in the environment. Therefore, the present findings are at odds with the predictions from both the retinal and the environmental motion extrapolation hypotheses.

Nevertheless, the present findings are in line with our recently proposed cross-modal bias hypothesis (Bai et al., 2020). The cross-modal bias hypothesis also stems from an everyday experience that a head turn naturally results in retinal motion in the opposite direction. This long-term experience may lead the neural system to develop a natural association between the vestibular self-motion signals and the self-motion-induced retinal signals. The underlying mechanisms for establishing and expressing the association may be based on Hebbian synaptic learning (Hebb, 1949) that has successfully accounted for the presence of mirror neurons (Gallese et al., 1996). Over one's own lifetime, the association becomes so strong that the signals from one modality could produce a bias signal in a congruent direction for the other modality. Given the intrinsic noises in neural responses, when the visual input signal is sufficiently weak and (or) uncertain, we hypothesize such a bias signal may easily manifest its perceptual outcome. In our previous work (Bai et al., 2020), the motion aftereffect (or slow real motion in Experiment 1b of that work) was predominantly

perceived as moving faster when its direction was opposite to the direction of the head rotation than when its direction was the same as the head rotation. Similar asymmetry in velocity perception was not observed for faster thus stronger visual motion stimuli (Bai et al., 2020). The unique findings for the weak visual motion signals are presumed to reflect a modulation by the visual motion bias signal resulted from head rotation.

The flash-lag effect has been found to diminish quickly when the flash was presented for 500 ms or even longer (Cantor & Schor, 2007; Eagleman & Sejnowski, 2000b; Lappe & Krekelberg, 1998; and see Hubbard, 2014 for a review). It is reasonable to infer that the position information for a brief flash was fairly uncertain as compared to when the flash is presented for long. Since motion and position perception may derive from common mechanisms (Dong et al., 2017; Kwon et al., 2015; Whitney & Cavanagh, 2000), it is possible that Schlag et al.'s head-rotation-induced flash-lag effect also reflects an effect of the cross-modal bias signal that is in the opposite direction to the head rotation. According to the cross-modal bias hypothesis, one would predict that the perceived position of the flash always lags behind the head position, regardless of whether the reference bar at the central fixation is retinally static (e.g., in Schlag et al.'s work) or retinally moving against the head rotation (though remaining stationary in the world) as in the present study. In other words, the cross-modal bias hypothesis can offer a unified explanation for the head-rotation-induced flash-lag effect in both Schlag et al.'s and our work.

Consistent with our previous work (Bai et al., 2020) in which the biasing effect was positively correlated to the head rotation velocity for and only for the motion aftereffect, in the current study we also found a moderate positive correlation between the strength of head-rotation-induced flash-lag effect in the HM condition and a trend of that in the HMRM condition. These results highlighted the contribution of the vestibular self-motion signals to the head-rotation-induced flash-lag effect and, thus, offered support for the cross-modal bias account. Yet the correlation did not reach significance in all the conditions containing head rotation, probably due to the concentration of the rotation velocity (BM condition), the weakness of the flash-lag effect (Experiment 3a), or the small sample size (Experiment 3b).

The cross-modal bias account was further supported by Experiment 3a. We observed a novel flash-lag effect produced following 1 min of adaptation to an abnormal association between head rotation and retinal motion. The finding can be well explained by the hypothesis that vestibular self-motion signals bias the perceived location of the flash towards the direction in which visual and vestibular inputs are recently associated. During the adaptation in Experiment 3a, participants saw a downward-drifting grating whenever making a leftward head turn. This re-afferent sensory signal

triggers activity in visual neurons encoding downward motion. Because the activities of these visual neurons consistently overlap in time with those of the vestibular neurons responding to leftward head turns, Hebbian learning would predict that the synapses connecting these two types of neurons should be potentiated. After repeated experience of this re-afference, the synapses would be strong enough so that those visual neurons would start firing to the associated head turns even without any retinal motion inputs. This, in turn, produces the flash-lag effect consistent in direction with the trained retinal motion.

It should be noted that though under the abnormal association (head-movement condition of Experiment 3a) a significant flash-lag effect was observed, the effect failed to be significantly different from that in the head-still condition. Even if there existed the possibility that this marginal trend ($p = 0.067$, $d = 0.45$) might reach significance with a larger sample size, we should admit that it was a rather weak effect, due partly to the relatively large individual differences. Also, it was obviously weaker than the effect under the natural associations such as the pilot experiment, Experiment 1 or Experiment 3b. The small magnitude was probably because the trained association based on short-term adaptation (for about 1 min) is temporary, and much weaker than the natural association based on the daily experiences. Specifically, during the early development of the neural system, multisensory stimulations may help form a relatively hardwired information integration network, especially the bottom-up integration network. By contrast, the outcome of the current short-term adaptation training, which might be regulated in a top-down manner, is believed to be much weaker in magnitude (Carriere et al., 2007; Wallace & Stein, 2007; Wallace et al., 2004). Indeed, motion extrapolation might still be playing an essential role in generating the flash-lag effect in Experiment 3b which had almost the same experiment settings as Experiment 3a. However, given presumably similar vestibular signals across the experiments, it is difficult to explain this magnitude difference by the motion extrapolation account alone. Therefore, one may parsimoniously use the cross-modal bias hypothesis to explain all types of head-rotation-induced flash-lag effect we explored in the present study.

On the other hand, the findings in Experiment 3a might also relate to the cross-modal contingent aftereffect. The cross-modal contingent aftereffect is the illusory perception in one modality that is associative to the stimulus from another modality after adaptation to these stimuli from both modalities accompanied (Hidaka et al., 2011; Teramoto et al., 2011), which can be considered as an extension of the classical, unimodal contingent aftereffect that occurs between different dimensions in a single modality such as the McCollough effect (McCollough, 1965). It is natural to link this phenomenon to the Experiment 3a of our study:

in Experiment 3a, observers were exposed to the leftward head movement accompanied by the downward-shifting grating (or rightward head movement by upward grating) in the adaptation phase, and then when they were tested with the head movement only, they perceived the shift of the flash in the corresponding direction. A large number of studies regarding the cross-modal contingent aftereffects in the domain of motion perception have focused on the interaction between vision and other modalities, such as audition and tactile sense (see Hidaka et al., 2015 for a review); by contrast, the visual-vestibular interactions in this phenomenon has not yet been researched as extensively, particularly considering the typical multimodal nature of the vestibular sense (Angelaki & Cullen, 2008). The current findings of Experiment 3a might help in filling this gap.

In addition to the motion extrapolation hypothesis, several other theories have been proposed to explain the flash-lag effect (see Hubbard, 2014 for a review), such as the attention shift theory (Baldo & Klein, 1995), the differential latency theory (Whitney & Murakami, 1998), and the postdiction theory (Eagleman & Sejnowski, 2000a). Few among them, though, can account for the current findings. For example, consider the HMRM condition in Experiment 2. According to the attention shift theory, the moving stimulus would keep moving whilst attention shifted to the flash. Thus, the flash should have lagged the moving reference, resulting in a perceived offset of the flash in the same direction as the head rotation. However, we observed the reversed pattern. Similarly, according to the postdiction theory, if the onset of flash resets the time window for temporal integration, the moving reference should also have appeared ahead of the flash along its moving direction, which is inconsistent with the current observation. As for the differential latency theory, it suggests that a moving target is processed more quickly than a stationary flashed object. Predicting the present finding by this theory is only possible when the following two premises have been demonstrated. First, the processings of the retinally moving reference stimulus and visual-vestibular interactions have to be serial. Second, the serial processings should be able to slow down the processing of the reference stimulus so much that the flash reaches the perceptual awareness faster than the moving reference. Even if this is the case, the differential latency theory still cannot explain the findings in Experiment 3a in which the perceived shift of flash and the head rotation were perpendicular to each other. This is because the reference stimulus in that experiment was retinally static and so its retinal position (i.e., central fixation) was reliably accessible all the time during the head rotation. Thus, there was unlikely a delay of latency for processing the reference stimulus relative to the flash.

In fact, none of the classical explanations can account for the perpendicular direction of the flash-lag effect and

the absence of retinal motion for the reference stimulus in Experiment 3a. However, all the past and current head-rotation-induced flash-lag effects can be explained with the cross-modal bias hypothesis. Nevertheless, the cross-modal bias hypothesis should not be used to interpret any types of visual-vestibular interactions. Visual-vestibular signals are not always integrated. There are also circumstances in which one sensory system (e.g., visual) suppresses processing in another sensory system (e.g., vestibular) to avoid visual-vestibular conflicts if the signals do not agree (Berger & Bülthoff, 2009; Brandt et al., 1998; Frank et al., 2016a, 2016b, 2020). Recent work even reports that the suppression of visual processing can occur during the preparation of head movement (Dong & Bao, 2021).

In summary, the current study compared the role of vestibular with kinesthetic signals in the formation of the head-rotation-induced flash-lag effect, and highlighted the role of vestibular signals. Moreover, we examined the motion extrapolation hypothesis proposed and modified by Nijhawan (1994) and Schlag et al. (2000) using the stimuli imitating static objects in real life. Our findings suggested that the motion extrapolation hypothesis could not provide a complete explanation for the mechanism of various types of the head-rotation-induced flash-lag effects. Instead, the cross-modal bias hypothesis we previously proposed (Bai et al., 2020) was a more flexible candidate account that works for all kinds of the head-rotation-induced flash-lag effects involved in the present study. To test our hypothesis, we used a new paradigm in which the participants' horizontal head rotation was artificially associated with the vertical retinal motion. We then observed a head-rotation-induced flash-lag effect consistent with this briefly trained association. This result pattern was again beyond the prediction of the motion extrapolation hypothesis, but in nice accord with the cross-modal bias hypothesis. Taken together, our findings provided a new perspective on the visual–vestibular interactions, lending more credence to the cross-modal bias hypothesis (Bai et al., 2020).

Appendix: Summary of the pilot experiment

Method

Participants

Before we conducted the pilot experiment, the number of participants was predetermined on the basis of the sample size for the previous study (Schlag et al., 2000). Eight participants (age range 21–31 years; 5 males and 3 females)

participated in the pilot experiment, including the author J.B.

Design and procedure

Participants completed one pure-black-background (BB) session and one dark-scene (DS) session. In the BB condition all the stimuli were presented on the pure-black background. In case the background environment in Schlag et al.'s (2000) work was dimly lit up by the light-emitting diodes, we also conducted a DS session in which all the procedure was the same as in the BB condition except that the environmental images of the experimental room in darkness were captured by a camera and then presented to the participants in real-time as if they were viewing with their own eyes (i.e., a see-through mode). The session order was counterbalanced across participants. The procedures for both conditions were almost the same as in the HM condition of Experiment 1, as illustrated in Fig. 1 of the main text. The minor difference was that there were only five spatial offsets (instead of seven) of the flash relative to the reference bar: 4 or 6 pixels to the left or right, or 0 (i.e., physically aligned).

Statistical analysis

The statistical analysis was also the same as in Experiment 1.

Results

Similar to Schlag et al.'s (2000) finding, we observed a negative PSA ($- 2.23 \pm 1.84$ pixels, $t(7) = 3.43$, $p = 0.011$, $d = 1.21$, 95% CI $= [- 3.78, - 0.69]$ pixels) in the BB condition, suggesting that the flash appeared lagging behind the stationary reference bar when the two bars were physically aligned with each other. In the DS condition, the head-rotation-induced flash-lag effect was also observed (PSA $= -1.87 \pm 1.17$ pixels, $t(7) = 4.52$, $p = 0.003$, $d = 1.60$, 95% CI $= [- 2.85, - 0.89]$ pixels), with no significant difference as compared with the BB condition ($t(7) = 1.05$, $p = 0.328$, $d_z = 0.37$).

References

Alais, D., & Burr, D. (2003). The "flash-lag" effect occurs in audition and cross-modally. *Current Biology, 13*(1), 59–63. https://doi.org/10.1016/S0960-9822(02)01402-1

Angelaki, D. E., & Cullen, K. E. (2008). Vestibular system: The many facets of a multimodal sense. *Annual Review of Neuroscience, 31*(1), 125–150. https://doi.org/10.1146/annurev.neuro.31.060407.125555

Arnold, D. H., Ong, Y., & Roseboom, W. (2009). Simple differential latencies modulate, but do not cause the flash-lag effect. *Journal of Vision, 9*(5), 4–4. https://doi.org/10.1167/9.5.4

Bai, J., Bao, M., Zhang, T., & Jiang, Y. (2019). A virtual reality approach identifies flexible inhibition of motion aftereffects induced by head rotation. *Behavior Research Methods, 51*(1), 96–107. https://doi.org/10.3758/s13428-018-1116-6

Bai, J., He, X., Jiang, Y., Zhang, T., & Bao, M. (2020). Rotating one's head modulates the perceived velocity of motion aftereffect. *Multisensory Research, 33*(2), 189–212. https://doi.org/10.1163/22134808-20191477

Baldo, M. V. C., & Klein, S. A. (1995). Extrapolation or attention shift? *Nature, 378*(6557), 565–566. https://doi.org/10.1038/378565a0

Berger, D. R., & Bülthoff, H. H. (2009). The role of attention on the integration of visual and inertial cues. *Experimental Brain Research, 198*(2–3), 287–300. https://doi.org/10.1007/s00221-009-1767-8

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Brandt, T., Bartenstein, P., Janek, A., & Dieterich, M. (1998). Reciprocal inhibitory visual-vestibular interaction. Visual motion stimulation deactivates the parieto-insular vestibular cortex. *Brain, 121*(9), 1749–1758. https://doi.org/10.1093/brain/121.9.1749

Cantor, C. R. L., & Schor, C. M. (2007). Stimulus dependence of the flash-lag effect. *Vision Research, 47*(22), 2841–2854. https://doi.org/10.1016/j.visres.2007.06.023

Carriere, B. N., Royal, D. W., Perrault, T. J., Morrison, S. P., Vaughan, J. W., Stein, B. E., & Wallace, M. T. (2007). Visual deprivation alters the development of cortical multisensory integration. *Journal of Neurophysiology, 98*(5), 2858–2867. https://doi.org/10.1152/jn.00587.2007

Cellini, C., Scocchia, L., & Drewing, K. (2016). The buzz-lag effect. *Experimental Brain Research, 234*(10), 2849–2857. https://doi.org/10.1007/s00221-016-4687-4

Cuturi, L. F., & MacNeilage, P. R. (2014). Optic flow induces nonvisual self-motion aftereffects. *Current Biology, 24*(23), 2817–2821. https://doi.org/10.1016/j.cub.2014.10.015

Dong, X., Bai, J., & Bao, M. (2017). Robust size illusion produced by expanding and contracting flow fields. *Vision Research, 133*, 87–94. https://doi.org/10.1016/J.VISRES.2017.01.003

Dong, X., & Bao, M. (2021). The growing sensory suppression on visual perception during head-rotation preparation. *PsyCh Journal*. https://doi.org/10.1002/pchj.438 Epub ahead of print.

Drewing, K., Hitzel, E., & Scocchia, L. (2018). The haptic and the visual flash-lag effect and the role of flash characteristics. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0189291

Eagleman, D. M., & Sejnowski, T. J. (2000a). Motion integration and postdiction in visual awareness. *Science, 287*(5460), 2036–2038. https://doi.org/10.1126/SCIENCE.287.5460.2036

Eagleman, D. M., & Sejnowski, T. J. (2000b). The position of moving objects. *Science, 289*(5482), 1107a. https://doi.org/10.1126/science.289.5482.1107a

Faisal, A. A., Selen, L. P. J., & Wolpert, D. M. (2008). Noise in the nervous system. *Nature Reviews Neuroscience, 9*(4), 292–303. https://doi.org/10.1038/nrn2258

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. https://doi.org/10.3758/BF03193146

Frank, S. M., Pawellek, M., Forster, L., Langguth, B., Schecklmann, M., & Greenlee, M. W. (2020). Attention networks in the parietooccipital cortex modulate activity of the human vestibular cortex during attentive visual processing. *Journal of Neuroscience, 40*(5), 1110–1119. https://doi.org/10.1523/JNEUROSCI.1952-19.2019

Frank, S. M., Sun, L., Forster, L., Tse, P. U., & Greenlee, M. W. (2016a). Cross-modal attention effects in the vestibular cortex during attentive tracking of moving objects. *Journal of Neuroscience, 36*(50), 12720–12728. https://doi.org/10.1523/JNEUROSCI.2480-16.2016

Frank, S. M., Wirth, A. M., & Greenlee, M. W. (2016b). Visual-vestibular processing in the human Sylvian fissure. *Journal of Neurophysiology, 116*(2), 263–271. https://doi.org/10.1152/jn.00009.2016

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain, 119*(2), 593–609. https://doi.org/10.1093/brain/119.2.593

Hayashi, R., & Murakami, I. (2019). Distinct mechanisms of temporal binding in generalized and cross-modal flash-lag effects. *Scientific Reports, 9*(1), 1–12. https://doi.org/10.1038/s41598-019-40370-7

Hebb, D. O. (1949). *The organization of behavior*. Wiley.

Hidaka, S., Teramoto, W., Kobayashi, M., & Sugita, Y. (2011). Sound-contingent visual motion aftereffect. *BMC Neuroscience, 12*(1), 1–6. https://doi.org/10.1186/1471-2202-12-44

Hidaka, S., Teramoto, W., & Sugita, Y. (2015). Spatiotemporal processing in crossmodal interactions for perception of the external world: A review. *Frontiers in Integrative Neuroscience, 9*, 62. https://doi.org/10.3389/fnint.2015.00062

Hubbard, T. L. (2014). The flash-lag effect and related mislocalizations: Findings, properties, and theories. *Psychological Bulletin, 140*(1), 308–338. https://doi.org/10.1037/a0032899

Hubbard, T. L. (2018). The Flash-lag Effect. In T. L. Hubbard (Ed.), *Spatial biases in perception and cognition*. Cambridge University Press. https://doi.org/10.1017/9781316651247

Keppel, G. (1991). *Design and analysis: A researcher's handbook* (3rd ed.). Prentice-Hall Inc.

Kwon, O.-S., Tadin, D., & Knill, D. C. (2015). Unifying account of visual motion and position perception. *Proceedings of the National Academy of Sciences of the United States of America, 112*(26), 8142–8147. https://doi.org/10.1073/pnas.1500361112

Lappe, M., & Krekelberg, B. (1998). The position of moving objects. *Perception, 27*(12), 1437–1449. https://doi.org/10.1068/p271437

Mackay, D. M. (1958). Perceptual stability of a stroboscopically lit visual field containing self-luminous objects. *Nature, 181*(4607), 507–508. https://doi.org/10.1038/181507a0

McCollough, C. (1965). Color adaptation of edge-detectors in the human visual system. *Science, 149*(3688), 1115–1116. https://doi.org/10.1126/SCIENCE.149.3688.1115

Nijhawan, R. (1992). Misalignment of contours through the interaction of apparent and real motion systems. *Investigative Ophthalmology & Visual Science, 33*(4), 974–974.

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature, 370*(6487), 256–257. https://doi.org/10.1038/370256b0

Nijhawan, R., & Kirschfeld, K. (2003). Analogous mechanisms compensate for neural delays in the sensory and the motor pathways: Evidence from motor flash-lag. *Current Biology, 13*(9), 749–753. https://doi.org/10.1016/S0960-9822(03)00248-3

Öğmen, H., Patel, S. S., Bedell, H. E., & Camuz, K. (2004). Differential latencies and the dynamics of the position computation process for

moving targets, assessed with the flash-lag effect. *Vision Research, 44*(18), 2109–2128. https://doi.org/10.1016/j.visres.2004.04.003

Purushothaman, G., Patel, S. S., Bedell, H. E., & Ogmen, H. (1998). Moving ahead through differential visual latency. *Nature, 396*(6710), 424–424. https://doi.org/10.1038/24766

Schlag, J., Cai, R. H., Dorfman, A., Mohempour, A., & Schlag-Rey, M. (2000). Extrapolating movement without retinal motion. *Nature, 403*(6765), 38–39. https://doi.org/10.1038/47402

Shams, L., Kamitani, Y., & Shimojo, S. (2000). What you see is what you hear. *Nature, 408*(6814), 788. https://doi.org/10.1038/35048669

Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative and Physiological Psychology, 43*(6), 482–489. https://doi.org/10.1037/h0055479

Teramoto, W., Kobayashi, M., Hidaka, S., & Sugita, Y. (2011). Cross-modal contingent aftereffect. *I-Perception, 2*(8), 880–880. https://doi.org/10.1068/IC880

von Holst, E., & Mittelstaedt, H. (1950). Das Reafferenzprinzip (Wechselwirkungen zwischen Zentralnervensystem und Peripherie). *Naturwissenschaften, 37*(20), 464–476. https://doi.org/10.1007/BF00622503

Wallace, M. T., Perrault, T. J., Hairston, W. D., & Stein, B. E. (2004). Visual experience is necessary for the development of multisensory integration. *Journal of Neuroscience, 24*(43), 9580–9584. https://doi.org/10.1523/JNEUROSCI.2535-04.2004

Wallace, M. T., & Stein, B. E. (2007). Early experience determines how the senses will interact. *Journal of Neurophysiology, 97*(1), 921–926. https://doi.org/10.1152/jn.00497.2006

Whitney, D., & Cavanagh, P. (2000). The position of moving objects. *Science, 289*(5482), 1107a. https://doi.org/10.1126/science.289.5482.1107a

Whitney, D., & Murakami, I. (1998). Latency difference, not spatial extrapolation. *Nature Neuroscience, 1*(8), 656–657. https://doi.org/10.1038/3659