Human estimates of descending objects' motion are more accurate than those of ascending objects regardless of gravity information

Takashi Hirata
Department of Cognitive and Psychological Sciences,
Nagoya University Graduate School of Informatics,
Nagoya, Aichi, Japan
JSPS Research Fellowships for Young Scientists,
Tokyo, Japan

Yutaka Hirata
Department of Artificial Intelligence and Robotics,
Chubu University College of Science and Engineering,
Kasugai, Aichi, Japan
Academy of Emerging Sciences, Chubu University,
Kasugai, Aichi, Japan
Center for Mathematical Science and Artificial
Intelligence, Chubu University, Kasugai, Aichi, Japan

Nobuyuki Kawai
Department of Cognitive and Psychological Sciences,
Nagoya University Graduate School of Informatics,
Nagoya, Aichi, Japan
Academy of Emerging Sciences, Chubu University,
Kasugai, Aichi, Japan

Humans can accurately estimate and track object motion, even if it accelerates. Research shows that humans exhibit superior estimation and tracking performance for descending (falling) than ascending (rising) objects. Previous studies presented ascending and descending targets along the gravitational and body axes in an upright posture. Thus, it is unclear whether humans rely on congruent information between the direction of the target motion and gravity or the direction of the target motion and longitudinal body axes. Two experiments were conducted to explore these possibilities. In Experiment 1, participants estimated the arrival time at a goal for both upward and downward motion of targets along the longitudinal body axis in the upright (both axes of target motion and gravity congruent) and supine (both axes incongruent) postures. In Experiment 2, smooth pursuit eye movements were assessed while tracking both targets in the same postures. Arrival time estimation and smooth pursuit eye movement performance were consistently more accurate for downward target motion than for upward motion, irrespective of posture. These findings suggest that the visual experience of seeing an object moving along an observer's leg side in everyday life may influence the ability to accurately estimate and track the descending object's motion.

Introduction

Accurate estimation of the position of a moving object is essential for successfully catching, hitting, or intercepting moving objects. Humans possess this estimation ability, as evidenced by their capacity to successfully catch or intercept a descending object at an appropriate moment (Brenner, Driesen, & Smeets, 2014; Lacquaniti et al., 2015; Lacquaniti & Maioli, 1989; Zago et al., 2004). Interestingly, our performance in catching or intercepting a moving object is asymmetrically accurate in the vertical (up and down) directions. For example, Zago, La Scaleia, Miller, and Lacquaniti (2011) found that participants exhibited superior interception performance when a visual target moved downward rather than upward. This study suggests that humans have a superior ability to estimate the descending motion states, such as their

Citation: Hirata, T., Hirata, Y., & Kawai, N. (2024). Human estimates of descending objects' motion are more accurate than those of ascending objects regardless of gravity information. *Journal of Vision*, *24*(3):2, 1–15, https://doi.org/10.1167/jov.24.3.2.

https://doi.org/10.1167/jov.24.3.2 Received July 5, 2023; published March 4, 2024 ISSN 1534-7362

position and velocity. The vertical asymmetry in the motion estimation for ascending and descending objects seems to be related to our experiences on Earth, in which we see downward motion such as a falling object, a thrown or batted ball, and raindrops in everyday life. Specifically, the sensory signals used to estimate the motion of a descending object in everyday life may play a crucial role in demonstrating this superior ability.

Daily experiences indicate that a descending object moves toward the direction of gravity. This direction can be perceived from multisensory signals of visual, vestibular, and somatosensory cues (Angelaki, McHenry, Dickman, Newlands, & Hess, 1999; Glasauer, 1992; Merfeld, Zupan, & Peterka, 1999). With visual cues, individuals can estimate "up and down" from visual information such as "sky and ground" and "ceiling and floor." When sitting or standing in the upright position, the sky or ceiling covers the observer's upper part of the visual field, whereas the ground or floor occupies the lower visual field. However, even without background scenes, such as sky and ground, individuals naturally predict that an object falls from the upper to lower visual field when no external forces are imposed. With vestibular cues, individuals can recognize the direction of gravity through vestibular signals. The up and down directions can be discerned through the otolith in their vestibular organs, which detects the observer's body orientation relative to the Earth's gravity axis. Previous studies have reported that gravitational sensory signals from vestibular organs may be used to estimate time to contact (TTC) for ascending and descending objects (Baurés & Hecht, 2011; Le Séac'h, Senot, & McIntyre, 2010; Senot, Zago, Lacquaniti, & McIntyre, 2005; Senot et al., 2012). These studies indicate that estimating a descending object's position and velocity may be facilitated when the directional information of gravity detected by vestibular cues and the directional information of descending object motion detected by visual cues are congruent. Therefore, our brain may accurately estimate the descending object's motion using two congruent pieces of information between vestibular and visual cues, and this methodology might be also applied to ascending motion.

Although previous studies have suggested that congruence of vestibular and visual information is necessary for the superior estimation of descending object motion, visual cues alone may provide sufficient information to estimate descending object motion. When observing the descending objects, we typically perceive them in an upright posture in which the downward motion of an object aligns with the observer's longitudinal body axis. The congruent pieces of information between the direction of object motion from the visual cues and the downward direction of egocentric (longitudinal body) axis may be used to estimate the vertical component of a descending object motion. Miwa, Hisakata, and Kaneko (2019) found that the vertical asymmetry of speed perception for the upward and downward motion of objects depends on the longitudinal body axis of the participants, rather than the gravity axis. This study suggests that humans perceive an object moving toward the observer's leg direction along the frontoparallel plane as descending object motion.

If congruence between the visual and vestibular cues is necessary for a superior estimation of descending object position and velocity, the vertical asymmetry of position and velocity estimation depends on the gravity axis. In previous studies on TTC estimation (Brenner et al., 2014; Lacquaniti et al., 2015; Lacquaniti & Maioli, 1989; Senot et al., 2005, Senot et al., 2012), the ascending and descending object stimuli were presented along the observer's longitudinal body and gravity axes in an upright posture. However, Indovina et al. (2005) and Miller et al. (2008) displayed upward and downward motion of visual stimuli along the egocentric vertical (observer's longitudinal body) axis in the supine posture. In this situation, the direction of the visual stimuli motion was not along the direction of the gravity axis, but rather along the observer's longitudinal body axis. Although these studies disentangled the direction of object motion and the direction of the gravity axis, they evaluated TTC estimation only for the downward motion of targets, not for both the upward and downward motion of targets. Nagai, Kazai, and Yagi (2002) also presented upward and downward motion of targets along the egocentric vertical axis in the upright and prompt postures. However, they did not measure TTC for the upward and downward motion of targets, but instead measured the participant's judgment of the vanishing position for both targets. It remains unclear whether the direction of gravity, detected by vestibular cues, is used for a superior estimation of downward object motion.

While estimating the timing of catching or intercepting moving objects, our gaze instinctively tracks the object's trajectory to accurately estimate it using smooth pursuit eye movements (SPEMs) and saccadic eye movements (Brenner & Smeets, 2009, Brenner & Smeets, 2011; Dorr, Martinetz, Gegenfurtner, & Barth, 2010). A previous study showed that TTC estimation was more accurate when participants tracked moving objects by SPEM than when they did not (Spering, Schütz, Braun, & Gegenfurtner, 2011). Moreover, SPEM is more accurate in tracking downward than upward (Akao, Kumakura, Kurkin, Fukushima, & Fukushima, 2007; Grönqvist, Gredebäck, & von Hofsten, 2006; Ke, Lam, Pai, & Spering, 2013; Takeichi et al., 2003). However, many studies on vertical SPEM performance have been conducted by presenting stimuli along the gravity axis in an upright posture. Therefore, the following question remains: Is congruent information regarding the direction of the object's motion and the observer's leg side used more effectively than the direction of gravity for superior TTC estimation and superior SPEM for descending objects?

In the current study, we presented the upward and downward motion of targets along the observer's longitudinal body (egocentric) axis in upright and supine postures. The upward and downward motion of targets were aligned with both gravity and the observer's longitudinal body axes in the upright posture. In contrast, in the supine posture, both target motions aligned with the longitudinal body axes, but were orthogonal to the gravity axis. We measured the accuracy of the arrival time estimation at a goal for the upward and downward motion of targets in upright and supine postures in Experiment 1. In Experiment 2, we assessed the performance of SPEM while participants tracked both targets in the upright and supine postures.

Experiment 1

A previous study on TTC estimation evaluated the estimation of arrival time at a particular location for a descending target (Miller et al., 2008). They recorded the time of button press when participants expected the target's arrival at the particular location. The time difference between the button press time and the actual arrival time of the target represents a motor correlate of the TTC estimation (Miller et al., 2008). Following Miller et al. (2008), we recorded the key responses to evaluate the accuracy of the arrival time estimation at a goal for egocentric upward and downward motion of targets in upright or supine postures. The target was displayed to move along the vertical axis in both the participant's upright and supine postures using a head-mounted display (HMD). As previous studies have shown vertical asymmetry of TTC estimation for 1 G (acceleration), 0 G (constant velocity), and −1 G (deceleration) object motion (Le Séac'h et al., 2010; Senot et al., 2005; Zago et al., 2011), we adopted these three acceleration conditions (1 G, 0 G, and −1 G) for the upward and downward motion of targets.

Methods

Participants

Twenty-three college students participated in Experiment 1 (12 females and 11 males, mean age 20.0 ± 0.5 years). All participants had normal or corrected vision. They provided written informed consent after the aims and procedures of the experiment were explained to them. All experimental procedures were designed and conducted per the Declaration of Helsinki and were approved by the Committee for Human Research at Nagoya University (210805-C-02-02).

Apparatus

Figure 1 summarizes a schematic representation of Experiment 1. Virtual reality (VR) visual stimuli were presented using an HMD (VIVE Pro Eye, 2,880 × 1,600 pixels; HTC, New Taipei City, Taiwan) at a refresh rate of 90 Hz. VR stimuli were created in Unity (Unity Technologies, San Francisco, CA) using customized programs written in C#. The participants were seated on a chair in an upright posture (Figure 1A, left) or lied down on a bed in a supine posture (Figure 1A, right). In the upright posture, a chinrest was provided to ensure a stable posture. A numeric keypad was placed by the dominant hand of the participant to collect responses.

Postures and stimuli

All participants experienced both upright and supine postures and observed two types of VR visual scenes: upward and downward motion scenes (Figure 1B). In both upward and downward motion scenes, a white ball target with a diameter of 0.2 m (visual angle of 0.8°) and a white ring-shaped goal were presented virtually at 13 m away from the participant's viewpoint along the frontoparallel plane. There were no depth cues in either the upward or downward motion scenes. In the upward motion scene, the target and goal were located 21.03° (visual angle) below and above the center of the screen, respectively (Figure 1B left). The target moved upward along the egocentric vertical axis (observer's longitudinal body axis) toward the goal positioned in the participant's head side in the upright and supine postures. In the downward motion scene, the location of the target and the goal, as well as the motion of the target, were opposite to those in the upward motion scene (Figure 1B right). Hence, the target moved downward along the egocentric vertical axis toward the goal positioned in the participant's leg side in both postures.

When the target had arrived at the goal in both scenes, the object overlapped the ring-shaped goal, resembling the appearance of Saturn. The target did not make any physical contact with the ring in either scene. The acceleration condition of the target motion was either 1 G (accelerated by 9.81 m/s^2 from the initial velocity at 0 m/s), 0 G (constant velocity at 7 m/s), or −1 G (decelerated by 9.81 m/s^2 from the initial velocity at 14 m/s) in both upward and downward motion scenes (Figure 1C). The constant velocity of the 0 G target was set to equal the average velocity of 1 G and −1 G targets. The initial velocity of −1 G target was set to be the same as the final velocity of 1 G target. Thus, under all acceleration conditions, the time from the beginning of the target motion to when the center of the target

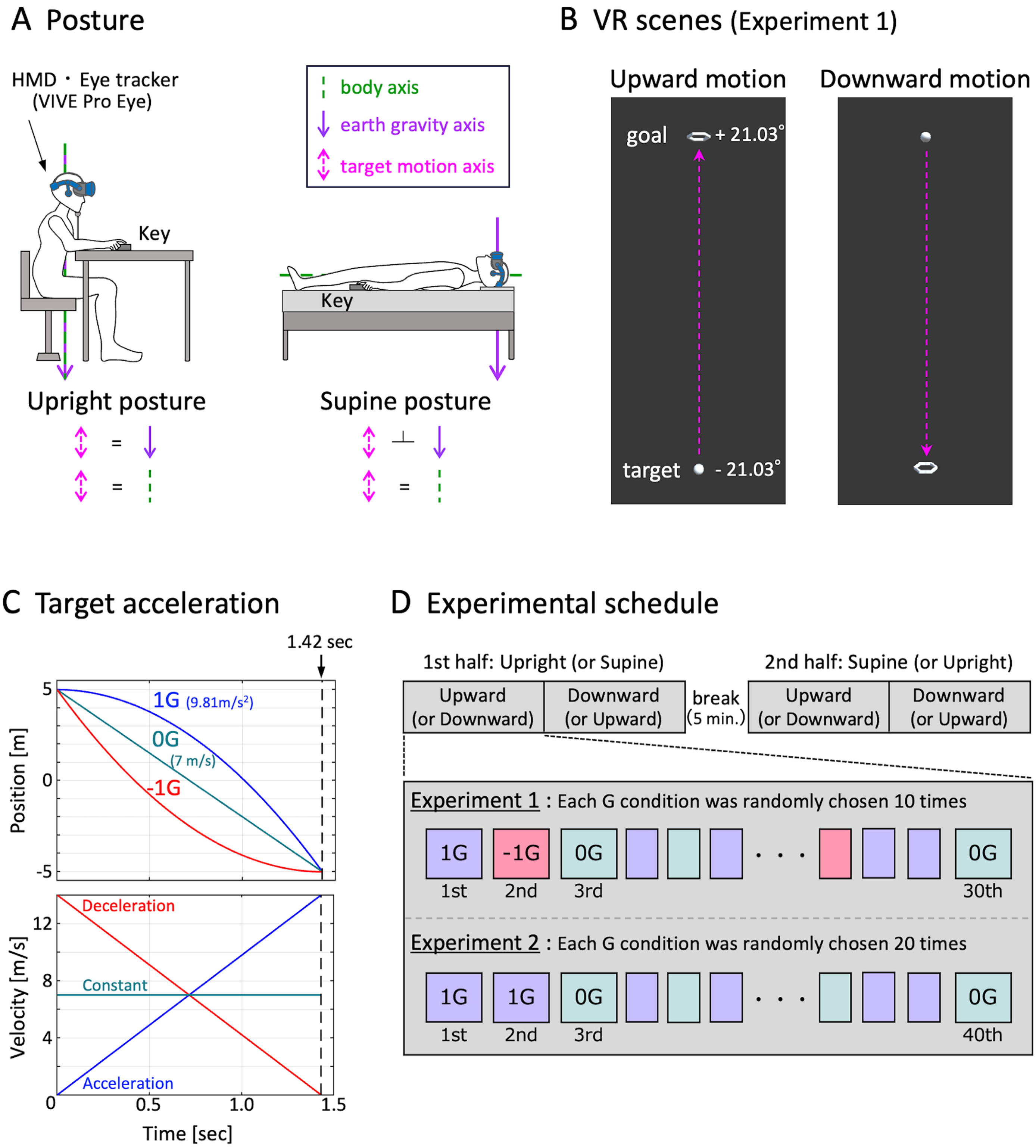

Figure 1. Experimental design. (**A**) Experimental setup and postures. The purple and green lines represent the Earth's vector and longitudinal body axes, respectively. The magenta line indicates the axis of the visual target stimulus. (**B**) VR scenes depicting upward motion (left) and downward motion (right) with the target trajectory indicated by the magenta line. (**C**) Target acceleration pattern in Experiment 1. The *x* axis represents the present time for the targets in upward and downward motion scenes, whereas the *y* axis represents the target position (above) and target velocity (below). The blue, green, and red traces correspond with 1 G, 0 G, and −1 G, respectively. (**D**) Experimental schedule in Experiments 1 and 2.

arrives at the center of the goal was the same: 1.42 seconds. Participants were not informed of the three acceleration types for the target or the elapsed time of target motion.

Procedure

The experimental procedure is illustrated in Figure 1D. Each participant performed the experimental task (arrival time estimation and key-pressing) in both upright and supine postures with the upward and downward motion scenes. The experiment was composed of two halves, and the upward and downward motion scenes were presented for each half. Either upright or supine postures were used in the first half. A different posture from the first half was used in the second half. The order of posture was randomized for each participant. The upward motion blocks for presenting the upward motion scene and the downward motion blocks for presenting the downward motion scene were included in both halves (Figure 1D). The order of each block was randomized for each participant, but the order of blocks was aligned in the first and second halves.

In each scene, the target was randomly presented ten times for each of the three acceleration conditions (30 times in total). During the presentation of each target, the target appeared on the HMD and remained stationary for 1 second. After this pause, the target moved toward the goal along the egocentric vertical axis in each scene. Once the target arrived at the goal, it disappeared for 3.58 s (intertrial interval [ITI]), resulting in a total of a 6-second trial (1-second pause + 1.42-second motion + 3.58-second ITI).

The participants experienced 12 experimental conditions (2 postures × 2 VR scenes × 3 accelerations). For each condition, the participants were instructed to press the number 0 on the numeric key with the index finger of their dominant hand as accurately as possible when the center of the target arrived at the goal. Feedback on arrival time estimation was not provided to participants. They were also instructed to follow the target with their eyes (SPEM) without blinking during target motion. Blinking was allowed only during the ITI. Participants were instructed not to move their heads during the target motion. To ensure that the participants maintained high concentration levels, they were given a 5-minute break between the different posture conditions.

Before the experiment, participants adjusted the chair and chin rest height to comfortable positions. They also adjusted their head position to a comfortable position in the supine posture. After the adjustments, they wore the HMD and underwent a practice session with the same VR scene and same posture in the first trial. The target moved at a constant velocity of 5 m/s in the practice trial to prevent the effect of the practice session on the first trial.

Data analysis

During the experiment, participants completed ten trials for each of the 12 experimental conditions (120 trials in total). We calculated the timing difference (TD) for each trial to evaluate the accuracy of the arrival time estimation. The TD was defined as the difference between the time from the beginning of the target motion to the arrival at the goal (1.42 seconds) and the time of the keypress. A positive TD indicates that the key is pressed before the target arrives at the goal. Outliers were identified and removed from each experimental condition if the TD values were outside the range of ±3 standard deviations from the mean TD for all the participants. A three-way analysis of variance (ANOVA) was performed to evaluate differences in TD with target accelerations (1 G, 0 G, and −1 G), target directions (upward and downward), and postures (upright and supine). We also conducted one-sample t tests for each mean TD to determine differences between the mean TD and zero. The significance level for the one-sample t test was set at 0.42% (5%/12 conditions).

Results

Figure 2 compares the mean TDs (error bars are standard errors) under different conditions for the 23 participants. The criterion of ±3 standard deviations excluded approximately 1.8% of the data in each experimental condition as outliers. A three-way ANOVA with target acceleration, target direction, and posture conditions showed a significant main effect for the target acceleration condition, $F(2, 272) = 75.32$, $p < 0.01$, $\eta^2 = 0.05$, and the target directions, $F(1, 272) = 22.07$, $p < 0.01$, $\eta^2 = 0.34$. After multiple comparisons for the acceleration condition, the average TD under the −1 G condition significantly differed from those under 1 G and 0 G conditions ($p < 0.01$). There was no main effect for the posture condition, $F(1, 272) < 1$, not significant, $p = 0.482$, $\eta^2 < 0.01$, or for any interactions. In additional tests, all mean TDs, except for the upward motion of −1 G condition in upright and supine postures, were significantly different from zero, $p < 0.0042$.

The mean TDs for the downward motion of the target at 1 G differed from those for the upward motion under both upright and supine conditions. This upward and downward asymmetry of mean TDs was observed in the 0 G condition in both upright and supine postures. In the −1 G condition, a vertical asymmetry of mean TDs was also observed in both upright and

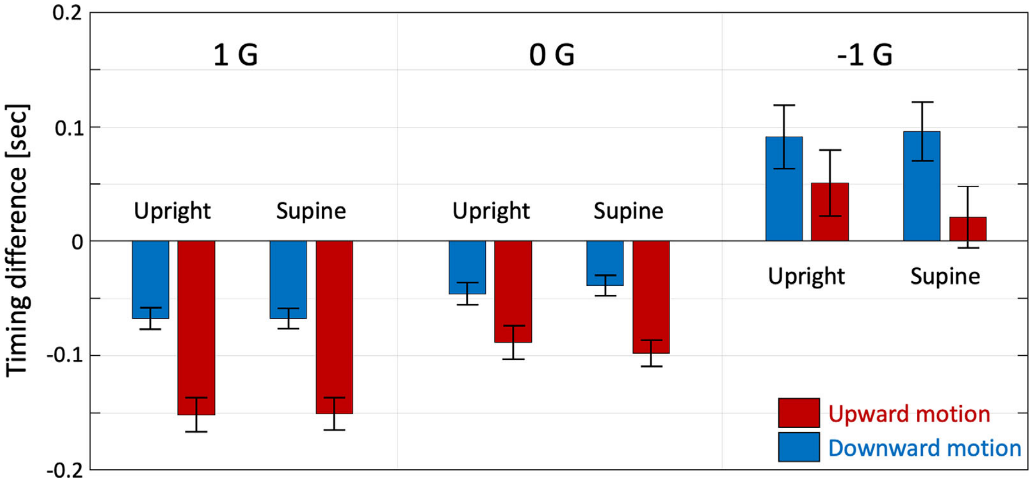

Figure 2. TD for the upward and downward motion of targets in upright and supine postures. TD for the upward and downward motion of targets in upright and supine postures under the 1 G, 0 G, and −1 G conditions. The red and blue bars represent the mean TDs for the upward and downward motion conditions, respectively. Each 1 G, 0 G, and −1 G figure's left and right side depict the upright and supine conditions, respectively. Error bars indicate 1 standard error.

supine postures, but the tendency of the asymmetry was opposite to that observed in the 1 G and 0 G conditions.

Discussion

The mean TDs for the 1 G and 0 G conditions indicated that the estimation of the arrival time for the downward motion of targets was more accurate than that for the upward motion of targets in both the upright and supine postures. In the −1 G condition, the estimation of the arrival time for the upward motion of the target was more accurate than that for the downward motion, regardless of posture. These findings support the notion that the direction of gravity, as indicated by the vestibular cue, did not influence the accurate estimation of the arrival time for the upward and downward motion of targets. Instead, the congruence with the direction of the target motion and the direction toward the observer's leg side in the frontoparallel plane played a significant role in the accurate estimation of the arrival time for the downward motion of the target.

This accurate arrival time estimation for the downward motion of a target has also been reported in previous studies, where the interception performance and TTC estimation for the descending object were more accurate than that for the ascending ones under 1 G and 0 G object motion in the upright posture (Le Séac'h et al., 2010; Senot et al., 2005; Zago et al., 2011). In addition, another study replicated the finding that the arrival time estimation was more accurate for a 1 G descending target than that for a 1 G ascending target in an upright posture (Hirata & Kawai, 2023). The result that the arrival time estimation for the upward motion of the target was more accurate than that for downward motion in the −1 G condition was also consistent with previous studies on TTC estimation for −1 G ascending and descending objects (Le Séac'h et al., 2010; Senot et al., 2005). Importantly, our study observed these previous findings of upward and downward asymmetry under 1 G, 0 G, and −1 G conditions in both the upright and supine postures, suggesting that vestibular cues do not influence asymmetric accuracies of the arrival time estimation for the upward and downward motion of targets.

Only the mean TDs of the −1 G target for both upward and downward motion were positive, indicating that participants pressed the key before the target arrival at the goal. A previous study reported that the estimation of TTC error for a moving target was influenced by the target's speed, with higher speeds leading to delayed responses (Peterken, Brown, & Bowman, 1991). Moreover, Bennett, Baures, Hecht, and Benguigui (2010) showed an early response in TTC for slow-moving targets when compared with medium and fast moving targets. In the current study, the mean TD results indicated that the key response was earlier in the following order: 1 G, 0 G, and −1 G targets. Furthermore, the final velocity of each target followed the same order: 1 G (53.7°/s), 0 G (27°/s), and −1 G (0°/s). Therefore, the positive mean TDs in the −1 G condition may be attributed to the fact that the final target velocity was slower than in the other acceleration conditions.

Under the 1 G and 0 G conditions, the mean TDs for the downward motion condition were smaller than those for the upward motion condition. These studies indicate that the timing of the key responses for the upward motion of the target was later than that for the downward motion. Although we did not measure velocity perception, previous studies showed that the response for TTC estimation was delayed in fast target motion (Bennett et al., 2010; Peterken et al., 1991). The fast velocity perception of upward motion could potentially be the cause of the delayed response for the downward motion of 1 G and 0 G targets.

A previous study reported that large retinal motion images induced the perception of fast target motion (Miyamoto, Numasawa, & Ono, 2022). The size of the retinal motion depends on how people track the target on the fovea during SPEMs. It is known that human SPEM performance for tracking an ascending object is less accurate compared with tracking a descending object (Brenner & Smeets, 2009, Brenner & Smeets, 2011; Dorr et al., 2010). Thus, the delayed response and the perception of faster speed in the upward motion of 1 G and 0 G targets might be caused by inaccurate SPEM performance and a large retinal motion.

A similar phenomenon related to keypress responses in arrival time estimation for a moving object has been reported as representational momentum (Freyd & Finke, 1984; Hubbard, 2005, Hubbard, 2018) and representational gravity (Gray & Thornton, 2001; Hubbard, 1995, Hubbard, 1998, Hubbard, 2001; Nagai et al., 2002). In representational momentum, people viewing an object moving from one location to another that suddenly disappears tend to report that they saw it slightly further along its path than where it vanished (Freyd & Finke, 1984). With representational gravity, people tend to more frequently perceive the final position of an object in a downward motion as being farther forward than the actual position of the object compared with an upward motion object (Hubbard, 1995, Hubbard, 1998, Hubbard, 2020). Early key responses for the downward motion of a −1 G target might be interpreted as an overestimation of the position of the −1 G targets owing to representational momentum and representational gravity.

The mean TDs for the −1 G condition tended to be opposite those for the 1 G and 0 G conditions; namely, arrival time estimation for upward motion was more accurate than that for downward motion. This difference is assumed to be caused by familiarity with vertically moving objects in the natural environment. The upward motion of a −1 G target and the downward motion of a 1 G target are consistent with naturally rising and falling objects on Earth, respectively. In contrast, the downward motion of a −1 G target and the upward motion of a 1 G target are inconsistent with natural object motion on Earth. In other words, in our visual experiences, we more often see rising and falling objects decelerating (−1 G) and accelerating (+1 G) at gravitational acceleration, respectively, than the other way around. For the 0 G condition, where we found that downward was more accurate than upward, it may be the case that we more often see falling objects than rising ones in the natural environment. The superior arrival time estimation for the −1 G upward motion and 1 G and 0 G downward motion might be attributed to visual experience with object motion on Earth.

Speed might be another factor causing the asymmetrical accuracy of arrival time estimation. Under 1 G and 0 G, the object's speed when it passed through the goal was 53.7°/s, and 27°/s, respectively, for both upward and downward motion, whereas under the −1 G condition, the object's speed was 0 when it passed through the goal. This significant difference in the target's speed near the goal might have caused the opposite directional asymmetries under 1 G and 0 G (>27°/s) and −1 G (close to 0°/s) conditions.

It is important to acknowledge the possibility of differences in vertical SPEM performance during arrival time estimation for the upward and downward motion of targets, because the participants were required to track them. Generally, during visuomotor tasks as conducted in our experiment, both SPEM and saccadic eye movements are generated (Spering et al., 2011; Orban de Xivry & Lefevre, 2007). People can perceive object motion during SPEM, whereas saccade suppresses the perception of object motion owing to saccadic suppression (Krekelberg, 2010). Therefore, SPEM plays a crucial role in estimating object motion and in intercepting or catching moving objects. We predicted that the performance of SPEM would contribute to accurately estimating the arrival time for the downward motion of a target. Consequently, asymmetry of SPEM in the vertical direction can be observed in both the upright and supine postures. In Experiment 2, we evaluated the performance of SPEM while participants tracked the upward and downward motion of targets along their longitudinal body axis in the upright and supine postures.

Experiment 2

Experiment 2 aimed to determine whether the tracking performance of SPEM for the downward motion of a target was more accurate than that for the upward motion when the target's direction of movement was not aligned with the gravity axis.

Methods

Participants

Thirty-six people participated in Experiment 2 (19 females and 17 males, mean age 29.6 ± 9.8 years). All participants had normal or corrected vision and did not participate in Experiment 1. They provided written informed consent after the aims and procedures of the experiment were explained to them. All experimental procedures were designed and conducted per the Declaration of Helsinki and were approved by the Committee for Human Research at Nagoya University (210805-C-02-02).

Apparatus

The same HMD as in Experiment 1 was used to present the VR visual stimuli and record the eye movements. The VR stimuli were created in Unity using customized programs written in C#. The eye movements were recorded using an eye camera (Tobii, Stockholm, Sweden) equipped with the HMD (VIVE Pro Eye) at a sampling rate of 120 Hz. Upright and supine postures were produced using the same desk, chinrest, chair, and bed.

Posture and stimulus

All participants watched the upward and downward motion scenes in upright or supine postures. The same white ball target with a diameter of 0.2 m (visual angle 0.8°) as that in Experiment 1 was presented virtually 13 m away from the participants' viewpoints in both scenes. A ring-shaped goal was not presented in Experiment 2 because we did not record key responses.

Naturally, people use SPEMs for tracking moving objects. However, when the object velocity exceeds the capability of SPEMs (eye velocity usually $<50°/s$), saccadic eye movements occur (Pola & Wyatt, 1991). To obtain a better visual angle and angular velocity to record vertical SPEM, we changed the target position and acceleration from Experiment 1. Each target in the upward and downward motion scenes was located 15° (visual angle) above or below the center of the screen, respectively. The target moved upward in the direction of the participant's head side along a frontoparallel plane in the upward motion scene. In the downward motion scene, the target moved downward in the direction of the participant's leg side along a frontoparallel plane. The acceleration of the target was either 1 G (the same acceleration as in Experiment 1) or 0 G (constant angular velocity of 24.2°/s) in both scenes. The velocity of the 0 G target was set to be the same average velocity of 1 G target. The target moved for 1.19 seconds under the 1 G and 0 G acceleration conditions.

To determine the parameters of target velocity, position, and length of trajectory in Experiment 2, a preliminary experiment was conducted beforehand. In the preliminary experiment, we recorded the SPEM while the participants followed the upward and downward motion of -1 G targets (initial velocity 53.7°/s). The SPEM results showed a large catch-up saccade under the upward and downward motion conditions because the initial speed of -1 G target exceeded the tracking capability of vertical SPEMs. In addition, most participants tracked the -1 G target using saccadic eye movements rather than SPEMs. We considered that the deceleration motion of the -1 G target was not appropriate for evaluating SPEM. Therefore, we did not include the -1 G condition in Experiment 2.

Procedure

The procedure for Experiment 2 is shown in Figure 1D. The experimental setup was identical to that described for Experiment 1, except that the target was presented 20 times for each acceleration condition in both the upward and downward motion scenes. During the presentation of each target, the target appeared on the HMD and remained stationary for one second. After this pause, the target moved toward the opposite end along the egocentric vertical axis in both upright and supine postures. Once the target moved 1.19 s, it disappeared for 3.81 seconds (ITI), resulting in a total motion cycle of 6 seconds (1-second pause + 1.19-seconds motion + 3.81-seconds ITI). The participants experienced eight experimental conditions (2 postures × 2 VR scenes × 2 accelerations) throughout the experiment. For each condition, the participants were instructed to track the target as much as possible without blinking. Blinking was allowed only during ITI. To ensure that the participants maintained high concentration levels, they were given a 5-minute break between the different posture conditions. The same procedure as that in Experiment 1 was performed before the start of the experiment.

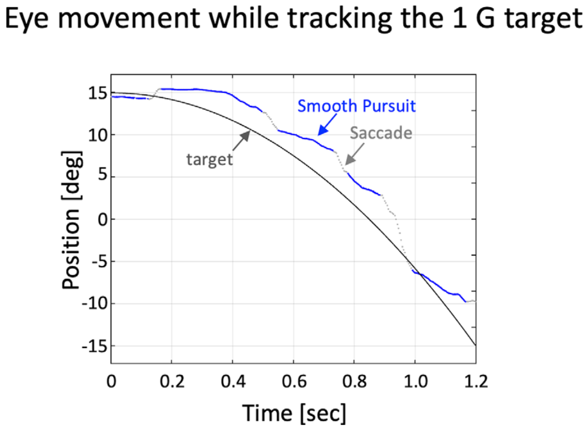

Figure 3. Vertical eye movement while a participant tracked the downward motion of a 1 G target in an upright posture. The horizontal axis represents the time after the beginning of the target motion, and the vertical axis represents the position (°). The black line shows the position of the 1 G target stimulus. The blue portion of the line shows SPEMs, and the gray portion shows saccade movements detected by the desaccading algorithm.

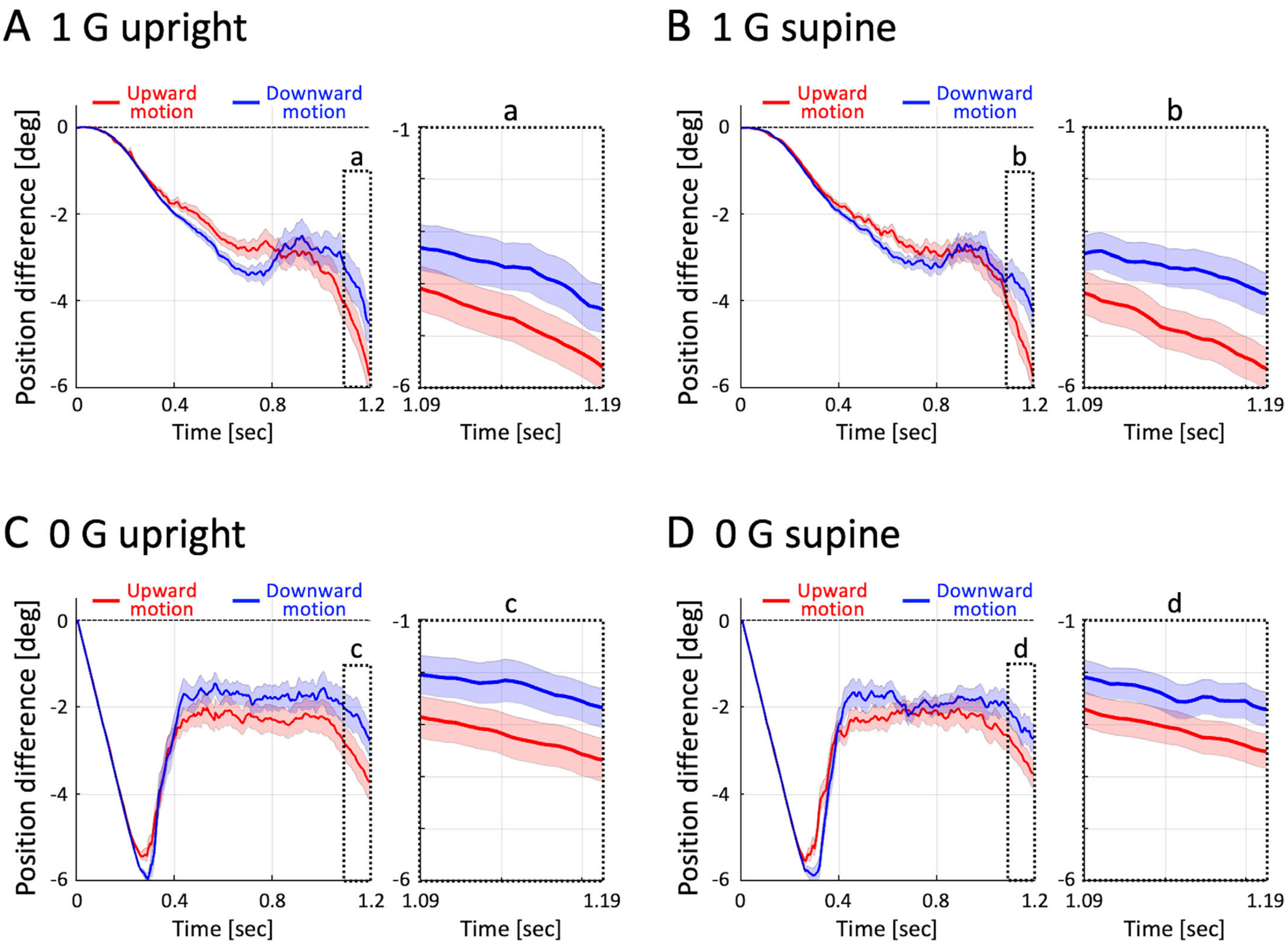

Figure 4. SPEM position difference under each experimental condition. The average SPEM position difference during target moving, with enlarged data between 1.09 and 1.19 seconds (a, b, c, d). The horizontal axis represents the time after the beginning of the target motion, and the vertical axis represents the position difference (°). The red and blue lines represent the mean SPEM position difference for the upward and downward motion conditions, respectively. Shadows indicate standard errors. (**A**) The 1 G upright condition. (**B**) The 1 G supine condition. (**C**) The 0 G upright condition. (**D**) The 0 G supine condition.

Data analysis

All eye movement analyses were conducted offline using MATLAB (MathWorks, Natick, MA). In total, 20 vertical eye position data were recorded for each of the eight experimental conditions. Vertical eye position data containing eye blinks were excluded. The eye position data were resampled from 120 to 240 Hz using the MATLAB interp1 function (cubic interpolation), and the eye velocity data were calculated by applying a three-point low-pass differentiation once. Saccades were eliminated from the eye movements using an automated desaccading algorithm with a velocity threshold (Hirata & Highstein, 2001). Figure 3 illustrates an example of a desaccading vertical eye position while tracking the downward motion of a 1 G target.

Excluded data were not used for further analyses. To evaluate the desaccading vertical SPEM performance under each experimental condition, the position difference (PD) between the target position and the smooth pursuit eye position was calculated using the following equation:

$$PD(t) = P_{stim}(t) - P_{eye}(t) \qquad (1)$$

where PD is the position difference and P_{stim} and P_{eye} are the target position (in degrees) and vertical smooth pursuit eye position (in degrees), respectively. t denotes the time from the beginning of the target movement.

Figure 4 shows the average smooth pursuit PDs under each experimental condition. In the upward motion condition, the PD were inverted to align the

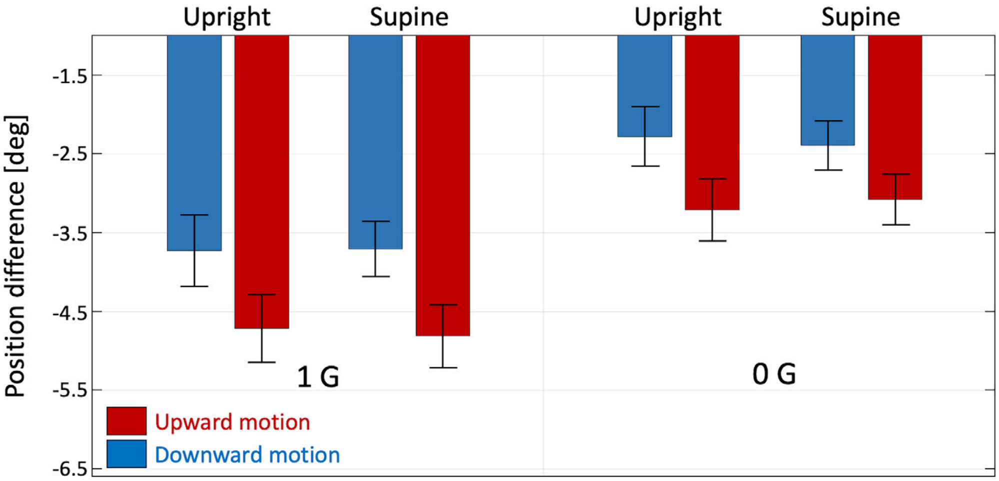

Figure 5. SPEM position difference for the upward and downward motion of targets in upright and supine postures. The average SPEM position difference for the upward (red) and downward (blue) motion of targets under 1 G and 0 G conditions. The left side of each 1 G and 0 G figure shows the upright condition, and the right side shows the supine condition. Error bars indicate one standard error.

direction of movement with the PD in the downward motion condition. We focused on the period between 1.09–1.19 s in the average PD for each experimental condition (Figure 4), because this period closely matched the time it took for the target to reach the goal in Experiment 1. To evaluate the SPEM performance during 1.09 to 1.19 seconds under each experimental condition, a three-way ANOVA was performed to evaluate differences in PD with target acceleration (1 G and 0 G), target direction (upward and downward), and postures (upright and supine). We also conducted one-sample t tests for each mean PD to determine the differences from the target position (zero). The significance level was set at 0.63% (5%/8 conditions) for the one-sample t tests.

Results

Several participants were excluded because their mean smooth pursuit eye positions were outside the range of ±3 standard deviations across all conditions: one participant was excluded from the Upright-upward motion of 0 G and Supine-downward motion of 1 G conditions. Two participants were excluded from the Upright-downward motion of 1 G and 0 G and the Supine-upward motion of 1 G and 0 G conditions. Three participants were excluded from the Upright-upward motion of 1 G and Supine-downward motion of 1 G conditions.

Figure 5 compares the mean PDs from 1.09 to 1.19 seconds between upward and downward motion conditions in upright and supine postures. A three-way ANOVA with target acceleration (1 G and 0 G), target direction (upward and downward), and posture (upright and supine) conditions showed a significant main effect of target acceleration, $F(1, 261) = 30.72$, $p < 0.01$, $\eta^2 = 0.1$, and target direction, $F(1, 261) = 11.74$, $p < 0.01$, $\eta^2 = 0.04$. There was no significant main effect of posture, $F(1, 261) < 1$, not significant, $p = 0.95$, $\eta^2 < 0.01$, and no interactions. The results of the one-sample t tests showed significant differences between the mean PDs and zero in all conditions, $p < 0.0063$. In the 1 G condition, the mean PDs for downward motion differed from those for the upward motion in upright and supine postures. Similarly, the mean PDs for downward motion also differed from those for upward motion in both upright and supine postures under the 0 G condition.

Discussion

In Experiment 2, participants could track the downward motion of 1 G and 0 G targets more precisely than the upward motion of 1 G and 0 G targets in both the upright and supine postures. This observed upward and downward asymmetry in SPEM performance aligns with previous studies (Brenner & Smeets, 2009, Brenner & Smeets, 2011; Dorr et al., 2010) that focused solely on the upright posture. Our findings suggest that the gravity direction from the vestibular cues has no impact on the tracking of descending objects using SPEMs.

Moreover, the results of Experiment 2 indicate that the unity between Earth's gravitational acceleration (1 G) and a target's acceleration does not provide an advantage in achieving superior SPEM performance. Notably, the mean PDs in the 1 G and 0 G conditions revealed that the participants tracked 0 G targets more accurately than 1 G targets, as evidenced by the observed PDs between the two acceleration conditions.

A major factor of this result would be the slower target's velocity in 0 G (24.2°/s) than in 1 G (39.4–41.8°/s during the evaluated period), suggesting that the participants tracked 0 G targets more easily than the 1 G targets.

We assume that the upward and downward asymmetry in SPEM performance may be acquired to adapt to the Earth's environment. On Earth, tracking descending objects requires higher performance than tracking ascending objects because free-falling objects move faster than ascending objects. When we pursue these free-falling objects, our eyes typically move in a downward direction. Therefore, eye movement habits may achieve superior performance in the downward SPEM when tracking descending objects. Although precise SPEMs for the downward motion of objects seems adaptive, whether this ability evolved because of the Earth's gravitational environment and whether humans innately possess it or acquire it through postnatal experiences remains uncertain.

General discussion

This study aimed to clarify whether superior estimation of arrival time (Experiment 1) and the performance of SPEM (Experiment 2) for a descending object require directionally congruent information between the object's motion and the observer's leg side. In Experiment 1, the arrival time estimation at the goal for the downward motion of 1 G and 0 G targets exhibited higher accuracy than that for the upward motion of 1 G and 0 G targets in the upright and supine postures. In Experiment 2, the performance of SPEM for the downward motion of 1 G and 0 G targets was more accurate than that for the upward motion in both postures. These results suggest that congruence between the direction of target motion detected by visual cues and the direction of the observer's leg side along the frontoparallel plane is required to perform accurate arrival time estimation and SPEM for objects in downward motion. In other words, the direction of gravity from vestibular cues does not assist in estimating the arrival time or the performance of SPEM for objects in downward motion. These findings represent the first investigation to demonstrate that vertical asymmetry in the arrival time estimation and in SPEM performance depends not on the gravity axis, but on the egocentric vertical (observer's longitudinal body) axis.

Both arrival time estimation and SPEMs demonstrated superior performance under 1 G and 0 G conditions when the target moved downward, as compared with when it moved upward (Figures 2 and 5). These results suggest that the superior arrival time estimation for the downward motion of targets under 1 G and 0 G in Experiment 1 may be induced by the better performance of downward SPEM. Better SPEM enhances visual clarity, particularly when tracking moving objects (Fooken, Yeo, Pai, & Spering, 2016). Additionally, previous studies have demonstrated that interception performance for a moving object is more accurate when participants actively track the object (Spering et al., 2011). Thus, the vertical asymmetry in SPEM performance may be a major factor in inducing the vertical asymmetry in arrival time estimation, TTC estimation, or interception performance for the upward and downward motion of objects.

Previous studies noted that the response to TTC for an approaching object was observed before the object reached the contact area (Benguigui, Ripoll, & Broderick, 2003; Mcleod & Ross, 1983; Schiff & Oldak, 1990; Neuhoff, 2001). In contrast, our result from Experiment 1 under 1 G and 0 G conditions showed that participants responded after a target arrived at the goal. The difference between the current study and previous studies may be attributed to whether the target motion is approaching, in conjunction with the requirements of SPEM and its performance. The early response (underestimation) of TTC for an approaching object has been interpreted by the "margin of safety theory" as an adaptive response that allows the observer to have enough time to engage in an appropriate response to the approaching object (Neuhoff, 1998, Neuhoff, 2001; Vagnoni, Lingard, Munro, & Longo, 2020). This underestimation is advantageous for survival because a response for an approaching object that is too late is far more dangerous than responding too early (Haselton & Nettle, 2006; Vagnoni, et al., 2020). In our study, the target was presented along the frontoparallel plane of a participant, suggesting that the "margin of safety theory" did not contribute to the arrival time estimation. In contrast, SPEM performance during tracking might be a cause of delayed responses. The mean PDs in Experiment 2 were negative, indicating that SPEM performance was insufficient and the participant's gaze did not catch up to the moving target (Figures 4 and 5). A similar situation must have occurred in Experiment 1, where the target velocities under the 1 G and 0 G conditions near the goal were faster than those in Experiment 2. Thus, the insufficient SPEM performance may have caused the late key response under the 1 G and 0 G conditions in Experiment 1. In the case of the looming target motion to which observers do not have to make SPEM to track the object, it is conceivable that the eyes were on the target. Taken together, underperformance of SPEM, which potentially causes faster motion perception under 1 G and 0 G conditions of the current experiment, may be the major factor in late responses.

The results of Experiments 1 and 2 suggest no impact of vestibular gravitational information on the performance of arrival time estimation and SPEM for the downward motion of a target. However, several previous studies have indicated that vestibular information can be used for TTC estimation (Le Séac'h et al., 2010; Senot et al., 2005; Zago et al., 2011). The experimental setup used in these studies differs from that used in the present study, including whether the vestibular information is used for TTC estimation or not. In previous studies, the participants looked up to see the downward motion of a target and looked down to see its upward motion, with the target approaching their faces. The relative position of the approaching target on the observer's retinal fovea was the same for both upward and downward motion. Moreover, the participants could not visually discriminate between the upward and downward motion. In contrast, vestibular signals are the information available to differentiate between upward and downward motion because they are different when looking up to see the downward motion of the target and down to see the upward motion of the target. In our study, the upward and downward motion of targets were presented in the participant's frontoparallel plane. Because the participants did not move their heads to see both targets, the vestibular signals were the same when they saw both targets. Therefore, they could not discriminate the target direction from the vestibular information. In contrast, visual information indicated that the targets moved toward the observer's head direction for the upward motion and toward the observer's leg direction for the downward motion. Participants in our study might not have relied on vestibular cues, but instead used visual cues to estimate the arrival time for upward and downward motion of targets.

Previous studies have suggested that vestibular inputs influence representational gravity (de sá Teixeira & Hecht, 2014; Nagai et al., 2002). These previous studies demonstrated that representational gravity was observed in the upright posture (i.e., gravity-congruent condition) in some conditions, but not in the other posture. Although our study parallels some of the methods and designs in previous studies on representational gravity, there are critical differences in the method. Participants in our study estimated the arrival time at the goal or tracked the moving target, whereas participants judged the final (vanishing) location for the moving target in the previous study. These methodological differences may produce divergent results.

Some previous studies have considered that the body's longitudinal axis seems to be the main vertical reference in the representational gravity literature (de sá Teixeira et al., 2017). de sá Teixeira et al. (2017) have shown that people perceive the final position of a downward motion of a target as being farther forward than the actual position of the object compared with an upward motion of a target. This vertical asymmetry in the final positional judgment of the upward and downward motion of targets depends on an egocentric vertical axis (de sá Teixeira et al., 2017). Our study also observed that the vertical asymmetry in arrival time estimation and SPEM performance for upward and downward motion depends on the egocentric vertical axis. These results suggest that people might perceive the direction of ascending or descending motion based on their longitudinal body's axis and visual cues.

It is known that people perceive vertical direction and body orientation by multisensory inputs from visual, vestibular, and somatosensory organs (Harris, Herpers, Hofhammer, & Jenkin, 2014; Kersten, Mamassian, & Yuille, 2004; Miwa et al., 2019). In multisensory inputs, sensory information is weighted depending on the reliability of each piece of information (Harris et al., 2014; Kersten et al., 2004; Miwa et al., 2019). Our results on the arrival time estimation and SPEM performance for each target's motion did not show differences between upright and supine postures. This finding suggests that, under the current experimental conditions where tasks require visual information processing, the weight of visual information may be greater than that of vestibular information. Humans may learn that the visual information of a target moving in the direction of the observer's head or leg is sufficient sensory input to perceive the motion of ascending and descending objects through everyday life.

One study reported that people begin to learn the features of descending object motion in early childhood (Kim & Spelke, 1992). This finding supports the idea that perceptual learning, rather than vestibular information, contributes to understanding the directions of a moving object throughout one's lifetime. de sá Teixeira (2014) reported that the human reference frame for ascension and descension is the body-referenced downward direction (Hubbard, 2020). Therefore, in a zero gravity environment, if people perceive an object's motion toward the leg direction, they may perceive it as the motion of a descending object.

Conclusions

This study aimed to clarify whether our superior ability to estimate descending object motion relies on the congruence between the direction of target motion and the leg side or gravity. Our findings suggest that both arrival time estimation and SPEM performance for the downward motion of targets were more accurate than those for the upward motion of targets under the 1 G and 0 G conditions, irrespective of whether the target moved along the Earth's gravity axis. These results indicate that congruent information between the object direction detected by visual cues and the direction toward the observer's leg side may

be required for a superior estimation of descending object motion. Furthermore, the direction of gravity from the vestibular information may not be necessary for estimating descending object motion. Finally, our results indicate that the visual experience of observing a descending object's motion toward the observer's leg direction in everyday life may influence the ability to estimate object motion accurately.

Keywords: time to contact, smooth pursuit eye movements, gravity information, body orientation

Acknowledgments

Supported by JSPS Research Fellow (JP22J1402) and JST CREST (Grant No. JPMJCR225P5).

Commercial relationships: none.

Corresponding author: Nobuyuki Kawai.
Email: kawai@is.nagoya-u.ac.jp.
Address: Nagoya University, Chikusaku, Furoucho, Nagoya, Aichi 464-8601, Japan.

References

Akao, T., Kumakura, Y., Kurkin, S., Fukushima, J., & Fukushima, K. (2007). Directional asymmetry in vertical smooth-pursuit and cancellation of the vertical vestibulo-ocular reflex in juvenile monkeys. *Experimental Brain Research, 182*(4), 469–478, https://doi.org/10.1007/s00221-007-1005-1.

Angelaki, D. E., McHenry, M. Q., Dickman, J. D., Newlands, S. D., & Hess, B. J. (1999). Computation of inertial motion: Neural strategies to resolve ambiguous otolith information. *Journal of Neuroscience, 19*(1), 316–327, https://doi.org/10.1523/JNEUROSCI.19-01-00316.1999.

Baurès, R., & Hecht, H. (2011). The effect of body posture on long-range time-to-contact estimation. *Perception, 40*(6), 674–681, https://doi.org/10.1068/p6945.

Benguigui, N., Ripoll, H., & Broderick, M. P. (2003). Time-to-contact estimation of accelerated stimuli is based on first-order information. *Journal of Experimental Psychology: Human Perception and Performance, 29*(6), 1083–1101, https://doi.org/10.1037/0096-1523.29.6.1083.

Bennett, S. J., Baures, R., Hecht, H., & Benguigui, N. (2010). Eye movements influence the estimation of time-to-contact in prediction motion. *Experimental Brain Research, 206*(4), 399–407, https://doi.org/10.1007/s00221-010-2416-y.

Brenner, E., & Smeets, J. B. (2009). Sources of variability in interceptive movements. *Experimental Brain Research, 195*, 117–133, https://doi.org/10.1007/s00221-009-1757-x.

Brenner, E., & Smeets, J. B. (2011). Continuous visual control of interception. *Human Movement Science, 30*(3), 475–494, https://doi.org/10.1016/j.humov.2010.12.007.

Brenner, E., Driesen, B., & Smeets, J. B. (2014). Precise timing when hitting falling balls. *Frontiers in Human Neuroscience, 8*, 342, https://doi.org/10.3389/fnhum.2014.00342.

de sá Teixeira, N. A. (2014). Fourier decomposition of spatial localization errors reveals an idiotropic dominance of an internal model of gravity. *Vision Research, 105*, 177–188, https://doi.org/10.1016/j.visres.2014.10.024.

de sá Teixeira, N. A., & Hecht, H. (2014). The dynamic representation of gravity is suspended when the idiotropic vector is misaligned with gravity. *Journal of Vestibular Research, 24*(4), 267–279, https://doi.org/10.3233/VES-140511.

de sá Teixeira, N. A., Hecht, H., Diaz Artiles, A., Seyedmadani, K., Sherwood, D. P., & Young, L. R. (2017). Vestibular stimulation interferes with the dynamics of an internal representation of gravity. *Quarterly Journal of Experimental Psychology, 70*(11). 2290–2305, https://doi.org/10.1080/17470218.2016.1231828.

Dorr, M., Martinetz, T., Gegenfurtner, K. R., & Barth, E. (2010). Variability of eye movements when viewing dynamic natural scenes. *Journal of Vison, 10*(10), 28–28, https://doi.org/10.1167/10.10.28.

Fooken, J., Yeo, S. H., Pai, D. K., & Spering, M. (2016). Eye movement accuracy determines natural interception strategies. *Journal of Vision, 16*(14), 1–1, https://doi.org/10.1167/16.14.1.

Freyd, J. J., & Finke, R. A. (1984). Representational momentum, *Journal of Experimental Psychology: Learning, Memory, and Cognition. 10*(1), 126–132, https://doi.org/10.1037/0278-7393.10.1.126.

Glasauer, S. (1992). Interaction of semicircular canals and otoliths in the processing structure of the subjective zenith. *Annals of the New York Academy of Sciences, 656*, 847–849, https://doi.org/10.1111/j.1749-6632.1992.tb25272.x.

Gray, R., & Thornton, I. M. (2001). Exploring the link between time to collision and representational momentum. *Perception, 30*(8), 1007–1022, https://doi.org/10.1068/p3220.

Grönqvist, H., Gredebäck, G., & von Hofsten, C. (2006). Developmental asymmetries between horizontal and vertical tracking. *Vision Research, 46*(11), 1754–1761, https://doi.org/10.1016/j.visres.2005.11.007.

Harris, L. R., Herpers, R., Hofhammer, T., & Jenkin, M. (2014). How much gravity is needed to establish the perceptual upright? *PLoS One, 9*(9), e106207, https://doi.org/10.1371/journal.pone.0106207.

Haselton, M. G., & Nettle, D. (2006). The paranoid optimist: An integrative evolutionary model of cognitive biases. *Personality and Social Psychology Reviewer, 10*(1), 47–66, https://doi.org/10.1207/s15327957pspr1001_3.

Hirata, Y., & Highstein, S. M. (2001). Acute adaptation of the vestibuloocular reflex: Signal processing by floccular and ventral parafloccular Purkinje cells. *Journal of Neurophysiology 85*(5), 2267–2288, https://doi.org/10.1152/jn.2001.85.5.2267.

Hirata, T., & Kawai., N. (2023). Enhanced estimation of object motion in the lower visual field compared the upper visual field. *The 40th Annual Meeting of Japanese Cognitive Science Society,* 617–620.

Hubbard, T. L. (1995). Cognitive representation of motion: Evidence for friction and gravity analogues. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(1), 241–254, https://doi.org/10.1037/0278-7393.21.1.241.

Hubbard, T. L. (1998). Some effects of representational friction, target size, and memory averaging on memory for vertically moving targets. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale, 52*(1), 44, https://doi.org/10.1037/h0087278.

Hubbard, T. L. (2001). The effect of height in the picture plane on the forward displacement of ascending and descending targets. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale, 55*(4), 325, https://doi.org/10.1037/h0087380.

Hubbard, T. L. (2005). Representational momentum and related displacements in spatial memory: A review of the findings. *Psychonomic Bulletin & Review, 12*, 822–851, https://doi.org/10.3758/bf03196775.

Hubbard, T. L. (2018). *Spatial Biases in Perception and Cognition*. Cambridge, UK: Cambridge University Press, https://doi.org/10.1017/9781316651247.

Hubbard, T. L. (2020). Representational gravity: Empirical findings and theoretical implications. *Psychonomic Bulletin & Review, 27*(1), 36–55, https://doi.org/10.3758/s13423-019-01660-3.

Indovina, I., Maffei, V., Bosco, G., Zago, M., Macaluso, E., & Lacquaniti, F. (2005). Representation of visual gravitational motion in the human vestibular cortex. *Science, 308*(5720), 416–419, https://doi.org/10.1126/science.1107961.

Ke, S. R., Lam, J., Pai, D. K., & Spering, M. (2013). Directional asymmetries in human smooth pursuit eye movements. *Investigative Ophthalmology & Visual Science, 54*(6), 4409–4421, https://doi.org/10.1167/iovs.12-11369.

Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology, 55,* 271–304, https://doi.org/10.1146/annurev.psych.55.090902.142005.

Kim, I. K., & Spelke, E. S. (1992). Infants' sensitivity to effects of gravity on visible object motion. *Journal of Experimental Psychology: Human Perception and Performance, 18*(2), 385, https://doi.org/10.1037/0096-1523.18.2.385.

Krekelberg, B. (2010). Saccadic suppression. *Current Biology, 20*(5), R228–229, https://doi.org/10.1016/j.cub.2009.12.018.

Lacquaniti, F., & Maioli, C. (1989). The role of preparation in tuning anticipatory and reflex responses during catching. *Journal of Neuroscience, 9*(1), 134–148, https://doi.org/10.1523/JNEUROSCI.09-01-00134.1989.

Lacquaniti, F., Bosco, G., Gravano, S., Indovina, I., La Scaleia, B., Maffei, V., . . . Zago, M. (2015). Gravity in the brain as a reference for space and time perception. *Multisensory Research, 28*(5-6), 397–426, https://doi.org/10.1163/22134808-00002471.

Le Séac'h, A. B. H., Senot, P., & McIntyre, J. (2010). Egocentric and allocentric reference frames for catching a falling object. *Experimental Brain Research, 201*, 653–662, https://doi.org/10.1007/s00221-009-2081-1.

McLeod, R. W., & Ross, H. E. (1983). Optic-flow and cognitive factors in time-to-collision estimates. *Perception, 12*(4), 417–423, https://doi.org/10.1068/p120417.

Merfeld, D. M., Zupan, L., & Peterka, R. J. (1999). Humans use internal models to estimate gravity and linear acceleration. *Nature, 398*(6728), 615–618, https://doi.org/10.1038/19303.

Miller, W. L., Maffei, V., Bosco, G., Iosa, M., Zago, M., Macaluso, E., . . . Lacquaniti, F. (2008). Vestibular nuclei and cerebellum put visual gravitational motion in context. *Journal of Neurophysiology, 99*(4), 1969–1982, https://doi.org/10.1152/jn.00889.2007.

Miwa, T., Hisakata, R., & Kaneko, H. (2019). Effects of the gravity direction in the environment and the visual polarity and body direction on the perception of object motion. *Vision Research, 164*, 12–23, https://doi.org/10.1016/j.visres.2019.08.005.

Miyamoto, M., Numasawa, K., & Ono, S. (2022). Changes in visual speed perception induced by anticipatory smooth eye movements. *Journal of Neurophysiology, 127*(4), 1198–1207, https://doi.org/10.1152/jn.00498.2021.

Nagai, M., Kazai, K., & Yagi, A. (2002). Larger forward memory displacement in the direction of gravity. *Visual Cognition, 9*(1–2), 28–40, https://doi.org/10.1080/13506280143000304.

Neuhoff, J. G. (1998). Perceptual bias for rising tones. *Nature, 395*(6698), 123–124, https://doi.org/10.1038/25862.

Neuhoff, J. G. (2001). An adaptive bias in the perception of looming auditory motion. *Ecological Psychology, 13*(2), 87–110, https://doi.org/10.1207/S15326969ECO1302_2.

Orban de Xivry, J. J., & Lefevre, P. (2007). Saccades and pursuit: Two outcomes of a single sensorimotor process. *Journal of Physiology, 584*(1), 11–23, https://doi.org/10.1113/jphysiol.2007.139881.

Peterken, C., Brown, B., & Bowman, K. (1991). Predicting the future position of a moving target. *Perception, 20*(1), 5–16, https://doi.org/10.1068/p200005.

Pola, J., & Wyatt, H. J. (1991). *Smooth pursuit: Response characteristics, stimuli and mechanisms. Eye Movements: Vision and Visual Dysfunction.* Macmillan Press.

Schiff, W., & Oldak, R. (1990). Accuracy of judging time to arrival: Effects of modality, trajectory, and gender. *Journal of Experimental Psychology: Human Perception and Performance, 16*(2), 303, https://doi.org/10.1037/0096-1523.16.2.303.

Senot, P., Zago, M., Lacquaniti, F., & McIntyre, J. (2005). Anticipating the effect of gravity when intercepting moving objects: Differentiating up and down based on nonvisual cues. *Journal of Neurophysiology, 94*(6), 4471–4480, https://doi.org/10.1152/jn.00527.2005.

Senot, P., Zago, M., Le Séac'h, A., Zaoui, M., Berthoz, A., Lacquaniti, F., . . . McIntyre, J. (2012). When up is down in 0g: How gravity sensing affects the timing of interceptive action. *Journal of Neuroscience, 32*(6), 1969–1973, https://doi.org/10.1523/JNEUROSCI.3886-11.2012.

Spering, M., Schütz, A. C., Braun, D. I., & Gegenfurtner, K. R. (2011). Keep your eyes on the ball: Smooth pursuit eye movements enhance prediction of visual motion. *Journal of Neurophysiology, 105*(4), 1756–1767, https://doi.org/10.1152/jn.00344.2010.

Takeichi, N., Fukushima, J., Kurkin, S., Yamanobe, T., Shinmei, Y., & Fukushima, K. (2003). Directional asymmetry in smooth ocular tracking in the presence of visual background in young and adult primates. *Experimental Brain Research, 149*(3), 380–390, https://doi.org/10.1007/s00221-002-1367-3.

Vagnoni, E., Lingard, L., Munro, S., & Longo, M. R. (2020). Semantic modulation of time-to-collision judgments. *Neuropsychologia, 147*, 107588, https://doi.org/10.1016/j.neuropsychologia.2020.107588.

Zago, M., Bosco, G., Maffei, V., Iosa, M., Ivanenko, Y. P., & Lacquaniti, F. (2004). Internal models of target motion: Expected dynamics overrides measured kinematics in timing manual interceptions. *Journal of Neurophysiology, 91*(4), 1620–1634, https://doi.org/10.1152/jn.00862.2003.

Zago, M., La Scaleia, B., Miller, W. L., & Lacquaniti, F. (2011). Coherence of structural visual cues and pictorial gravity paves the way for interceptive actions. *Journal of Vison, 11*(10), 13–13, https://doi.org/10.1167/11.10.13.