



OPEN Psychophysical evidence for an internal model of gravity in the visual and vestibular estimates of vertical motion duration

Sergio Delle Monache^{1,2,5}✉, Barbara La Scaleia^{1,5}✉, Anna Finazzi Agrò^{1,3,4},
Francesco Lacquaniti^{1,2,6} & Myrka Zago^{1,2,6}

The motion of objects and ourselves along the vertical is affected by gravitational acceleration. However, the visual system is poorly sensitive to accelerations, and the otolith organs do not disassociate gravitational and inertial accelerations. Here, we tested the hypothesis that the brain estimates the duration of vertical visual motion and self-motion by means of an internal model of gravity predicting that downward motions are accelerated and upward motions are decelerated by gravity. In visual sessions, a target moved up or down while participants remained stationary. In vestibular sessions, participants were moved up or down in the absence of a visual target. In visual-vestibular sessions, participants were moved up or down while the visual target remained fixed in space. In all sessions, we verified that participants looked straight-ahead. We found that downward motions of either the visual target or the participant were systematically perceived as lasting less than upward motions of the same duration, and vice-versa for the opposite direction of motion, consistent with the predictions of the internal model of gravity. In visual-vestibular sessions, there was no significant difference in the average estimates of duration of downward and upward motion of the participant. However, there was large inter-subject variability of these estimates.

Keywords Internal model, Gravity, Otoliths, Psychophysics, Bias, Human

We move and experience object motion in a 3D-environment. Our spatial perception and orientation are typically anchored to the gravitational vertical and the horizon, and depend heavily on visual and vestibular cues. Thus, there is as much image structure at vertical orientation as at horizontal orientation and least at obliques in the natural visual scenes commonly experienced¹. With regards to the vestibular inputs experienced daily, small-amplitude self-motion in the vertical direction is as common as that in the horizontal direction². Here, we are concerned with the processing of temporal duration of motion of objects or ourselves in the vertical direction.

The issue of whether time estimates depend on a centralized supramodal clock or distributed, modality-specific mechanisms is still unsettled^{3–5}. Irrespectively, an internal model of gravity effects might contribute to time estimates for motions with a vertical component⁶. In fact, the visual system is poorly sensitive to accelerations⁷, while the otolith organs of the vestibular system cannot disentangle gravitational and inertial accelerations of self-motion⁸. However, brain mechanisms predicting gravitational kinematics in a probabilistic manner^{9–15} may contribute to the corresponding time estimates^{16–18}.

There is ample evidence for the motor utilization of a gravity model in the implicit time estimates required by motor actions^{9,12,19–21}. Thus, humans catch a ball falling under gravity (1g) in the absence of visual cues when interception time and location are predictable^{22,23}. On the other hand, erroneous assumption of 1g leads astronauts to move too early to catch a ball descending at constant speed (0g) in weightlessness²⁴, as it does for participants asked to intercept a virtual target descending at 0g on Earth²⁵. In both cases, the response timing corresponds to the expectation of 1g effects^{24,25}. In the extrapolation of vertical motion through a visual

¹Laboratory of Neuromotor Physiology, IRCCS Santa Lucia Foundation, Rome 00179, Italy. ²Department of Systems Medicine and Centre of Space Bio-medicine, University of Rome Tor Vergata, Rome 00133, Italy. ³Centre of Space Bio-Medicine, University of Rome Tor Vergata, Rome 00133, Italy. ⁴Department of Physics, University of Trento, Trento 38123, Italy. ⁵Sergio Delle Monache and Barbara La Scaleia contributed equally to this work. ⁶These authors jointly supervised this work: Francesco Lacquaniti and Myrka Zago ✉email: sergio.dellemonache@uniroma2.it; b.lascaleia@hsantalucia.it

occlusion, the interception of virtual targets accelerated by gravity (1g) is more accurate than that of targets decelerated by reversed gravity (-1g), or moving at constant speed (0g)²⁶. Moreover, there is a response bias as a function of the upward or downward direction of target motion: whatever the acceleration of the virtual target, subjects trigger movements to intercept it earlier when it comes from above instead of below, consistent with the prior assumption that downward motion is accelerated by gravity while upward motion is decelerated by gravity²⁷. The response bias for accelerated 1g motion in the upward direction (unnatural motion) versus the downward direction (natural motion) depends on the familiarity with the target trajectory²⁸. Vestibular inputs also play a role in interception timing²⁹, since the response bias reversed sign between the above and below conditions in parallel with the sign reversal of otolith signals at the transition from hypergravity to hypogravity during parabolic flight³⁰. Eye movements also betray the implementation of a gravity model, as shown by the higher gain of smooth pursuit when tracking targets moving at 1g than those moving at 0g or other g levels^{31–33}, and by faster and smoother pursuit in response to downward versus upward motion³⁴. As for the vestibular system, reflexive eye movements in response to tilts and translations appear related to central estimates of linear acceleration derived from an internal model of gravity, rather than related to the net gravito-inertial force encoded by the otoliths^{35–37}.

In contrast with the implicit time estimates of motor tasks, there is considerably less evidence for the gravity model in the explicit time estimates required by the discrimination of vertical motion duration³⁸. Implicit or procedural knowledge of physics (as involved in motor actions) and explicit or declarative knowledge (as involved in perceptual judgments) may depart substantially from one another^{39–41}. Therefore, there is no guarantee that visual and vestibular estimates of motion duration share the same prior assumption about the effects of gravity as motor actions.

However, some evidence for the internal model of gravity in perceptual judgments of visual motion duration has been produced. Thus, it has been shown that the discrimination precision of motion duration⁴² (or speed⁴³) is significantly better for motions accelerating downwards than accelerating upwards. Moreover, downward visual motions at constant speed are misperceived as faster than upward motions at the same speed⁴⁴.

As for the vestibular system, there is indirect evidence that it may contribute to judgments of elapsed time. Thus, estimates of a 1-minute interval made by astronauts in flight – when their vestibular system is understimulated – are significantly longer than those made pre- or postflight on Earth^{45,46}. With regards to the vestibular discrimination of motion duration, healthy persons do not exhibit any significant bias when estimating in the dark the duration of rightward rotations relative to leftward rotations about a yaw axis (horizontal plane)⁴⁷. However, to our knowledge, the vestibular discrimination of duration of vertical motions has not been tested so far.

To address these open issues, we tested visual and vestibular discrimination of vertical motions duration. Participants sat upright on a motion platform, wearing a virtual-reality headset. In visual sessions (VI), a visual target moved up or down while participants remained stationary. In vestibular sessions (VE), participants were moved up or down and were asked to fixate an imaginary target moving together with the subject. In visual-vestibular sessions (VV), participants were moved up or down while the visual target was displayed at a fixed location, thus shifting relative to the subject in the direction opposite to their movement. In a two-alternative forced choice task, they were asked to judge which one of two stimuli had the shorter duration, i.e., a comparison stimulus moving up or down and a reference stimulus moving in the opposite direction (down or up). However, since stimulus speed and acceleration covaried with duration, they could use these kinematic cues to estimate duration.

The hypothesis of an internal model of gravity predicts that the downward motion of the target in the visual (VI) sessions and that of the participant in the vestibular (VE) sessions should be perceived as lasting significantly less than the upward motions of the same duration, and vice-versa for the opposite direction of motion. This prediction would be consistent with the prior assumption that downward motion is accelerated by gravity while upward motion is decelerated by gravity²⁷.

The prediction for the visual-vestibular sessions (VV), however, is not univocal due to the potential ambiguity in the subjective interpretation of the stimuli, given that the participant was moved up or down while the visual target was displayed at a fixed location. Therefore, the target shifted relative to the participant in the opposite direction to the participant motion. If the participant responded mainly to either the vestibular or the visual information, we should find a bias similar to that of the corresponding VE session or the VI session, respectively. If, however, the participant responded to both the vestibular and the visual cues at the same time, the visual and the vestibular bias should roughly cancel out, given that the resulting retinal slip would tend to compensate the head motion.

Methods

Participants

Twenty young adults (10 women, 10 men, mean age 29.7 ± 6.0 SD), with normal or corrected-to-normal vision, no history of psychiatric, neurological or vestibular symptoms, dizziness or vertigo, motion-sickness susceptibility, major health problems or medications potentially affecting vestibular function, volunteered to participate in the experiments. Sample size was calculated to detect an effect size of 0.83 (Cohen's d, estimated from previous studies with comparable conditions – Moscatelli et al.⁴⁴ and Kobel et al.⁴⁸ – as well as from our preliminary data) with power of 0.8 and alpha of 0.017 (i.e. $0.050 / 3$, where 3 is the correction for the expected levels of comparison corresponding to the number of sessions. G*power version 3.1). Data were anonymized after collection for subsequent analysis. To avoid experimenter-expectancy effects, the analysis was carried out on blinded data by experimenters different from those involved in data collection. Experimental procedures were approved by the Institutional Review Board of Santa Lucia Foundation (protocol n. CE/PROG0.757), and were performed in accordance with the Declaration of Helsinki (World Medical Association) regarding the use

of human participants in research. All participants gave written informed consent to participate to experimental procedures. Written informed consent was obtained from the participant to publish videos and figures in an online open access publication.

Setup

Participants sat on a gaming chair placed on a six-degrees-of-freedom motion platform (MB-E-6DOF/12/1000, Moog, USA). A 4-point harness held their trunk securely in place, while a medium density foam pad under the feet minimized plantar cues about body displacement. For acoustic isolation, participants wore earplugs (1100 series, 3M, USA) and on-ear headphones (Anker Soundcore Q20, China). They wore a 3D VR-headset system (Quest Pro, Meta, USA) with a refresh rate of 90 Hz and a horizontal Field of View of 106°. The VR-headset system also acquired the position of the head and eyes at 90 Hz, the latter with an average accuracy better than 2 degrees^{49,50}. Data from platform (chair) and VR-headset (head and eyes position and orientation) were referred to a fixed world reference system (W_{RF} see Fig. 1). We estimated the roto-translation matrix required to map the 3D data acquired in the VR-headset (head) reference frame and in the platform reference frame into W_{RF} by performing a spatial calibration procedure before each experimental session with the VR-headset mounted on the platform in a stable position. All details about the general spatial calibration procedures can be found in La Scaleia et al.⁵¹. By performing a cross-correlation analysis between the position and orientation data of the VR-headset (once it was rigidly attached to the platform) of the chair during predetermined movements, we estimated that there is a delay of 40 ms between the two systems. In managing the apparatuses to generate visuovestibular stimuli that are both congruent and temporally aligned, we compensated for any temporal misalignment. The data relative to the position of the chair, head and eyes acquired during the experiment and referred to W_{RF} were used to assess that the participants kept the head and eyes roughly aligned with the straight-ahead direction (see below) during the execution of the task.

Scenario

With the VR-headset system, participants were presented with a three-dimensional scenario programmed by means of Unity engine (editor version 2021.3.25f1, packed with XR Plugin Management 4.3.3, Oculus XR Plugin 3.3.0 and Meta Movement SDK 1.4.1). Prior to the beginning of the experiment, each subject was asked to slide the lenses of the VR-headset closer or further apart from one another so as to have a good VR experience and to achieve the best image clarity. This was required to take into account the interpupillary distance of the participant. To this end, they had to fuse binocularly a text presented in stereo at the same distance as the experimental scene. Afterwards, participants underwent the eye-tracking calibration routine included in the VR-headset software.

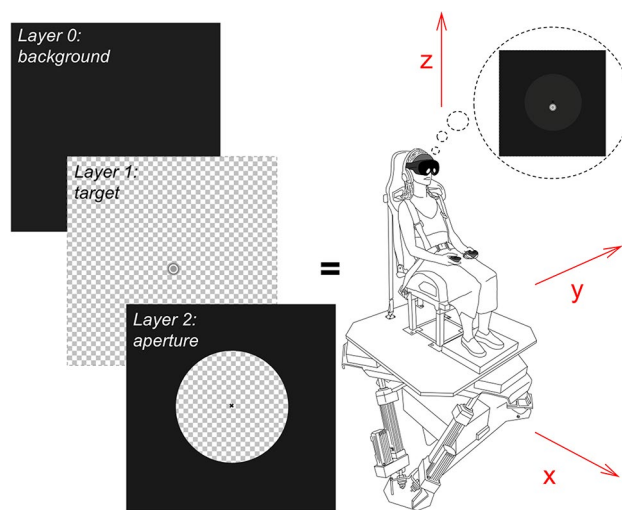


Fig. 1. Visual stimulation in the visual and visuo-vestibular sessions consisted of three superimposed layers, oriented perpendicular to the subject's antero-posterior axis from back to front (see Methods). Layer 0 comprised the background of the virtual scene. Layer 1, positioned immediately in front of Layer 0, contained the visual target, a grey-textured disk with a diameter of 1.68° of visual angle. Located in front of Layer 1, Layer 2 included two elements: a uniformly colored surface with a central circular aperture; and a black cross tilted by 45° at its center. In Layer 1 and Layer 2, transparent portions of the scenario are represented with a squared texture. The resulting 3D virtual scenario was perceived by the subject through the VR-headset system (see also Supplementary Video 1 and Supplementary Video 3). In the vestibular session, visual stimulation consisted solely of the background (Layer 0; see also Supplementary Video 2). The W_{RF} (axes x, y, z, represented in red in the figure) has its origin in the flying base motion centroid position when the flying base is in its home settled position. Motion centroid is the centroid of the joints below the flying base of the motion platform. The z-axis is orthogonal to the ground and oriented upwards. The x-axis is oriented along the posterior-anterior direction of the chair.

Each experimental session started with the same initial calibration scene, consisting of a plain, uniform scene (the intensity of each color in the RGB color code was set to 20). Participants were asked to keep their head in a relaxed position, facing forward along the perceived antero-posterior direction and looking straight-ahead. The mean position of the eyes acquired over 1 s during the calibration phase was used to define the position of the working area (see below for a description of the working area). This area was placed 42 cm in front of the participant, along their sightline, and was roughly perpendicular to the world reference ground (plane x-y, see W_{RF} in Fig. 1 for further details), with its surface exposed to the subject. Afterwards, the location of the working area was held constant in world reference, independently of the subject's actual position and orientation. By contrast, the point of view was updated according to the line of sight of the participant estimated through the convergence of the gaze directions of the two eyes recorded by the VR-headset at 90 Hz throughout the experimental session.

The aforementioned working area consisted of a group of virtual objects organized in three superimposed layers (Layer0, Layer 1 and Layer 2 from bottom to top. See Fig. 1). They were illuminated by an orthogonal directional white light (RGB color intensity set to 255 for each color). Layer 0, the background layer, was a uniformly grey surface (see below for the RGB color intensity). In front of this, Layer 1 contained the visual target in a transparent environment. The target was a disk with a diameter of 1.68° of visual angle, featuring a grey texture with high contrast against the Layer 0 (Michelson contrast 74%). This grey texture consisted of four concentric circles, each displaying a distinct grayscale color level (RGB color intensity set to 136, 89, 204, 102 for each color, from outer to inner circles, respectively), similarly to the high-contrast condition of Moscatelli et al.⁴⁴. For illustrative purposes, the transparent environment of Layer 1 is depicted with a squared texture in Fig. 1. Further anteriorly, Layer 2 included two elements: a black (RGB color intensity set to 0 for each color) fixation cross tilted by 45° measuring 0.42° of visual angle in both width and height, and a uniformly colored vertical surface (RGB color intensity set to 25 for each color) with a circular transparent aperture (in Fig. 1, the transparency was represented with the same squared texture as for Layer 1) centered on the fixation cross. The aperture had a diameter of 14.25° of visual angle.

The experiment consisted of three sessions, each involving vertical movements of different elements of the virtual scenario and/or of the chair. (i) In the visual session (VI), all Layers of the visual scene were visible, the chair remained static throughout the session, as did the Layer 0 (background; RGB color intensity set to 30 for each color) and the Layer 2. In its initial position, the target in Layer 1 was either vertically above or vertically below the fixation cross of Layer 2. The target became visible through the aperture when the Layer 1 moved vertically (see Supplementary Video 1). When the target, during Layer 1 motion, was aligned with the fixation cross of Layer 2, it was partially hidden by the cross (Layer 2 was stacked on top of Layer 1). (ii) In the vestibular session (VE), only the Layer 0 (background) was visible in the virtual scene (Layer 1 and Layer 2 were hidden from view) while the chair moved vertically (see Supplementary Video 2). To match the overall luminosity of the virtual scenario between sessions, in VE the Layer 0 was slightly darker than in the other sessions (RGB color intensity set to 27 for each color), (iii) In the visuo-vestibular session (VV), the chair moved vertically as in the VE session. All three layers of the visual scene were visible as in VI. Unlike VI, the visual target in Layer 1 was stationary, immediately above or below the aperture of Layer 2, however Layer 2 position was linked to that of the chair movement. As the chair moved, so did the aperture and the fixation cross of Layer 2, making the target temporarily visible during the stimulation (see Supplementary Video 3). In the VV session, the positions of the chair and Layer 2 (aperture and fixation cross) were synchronized by adjusting the timing between the platform motion and VR-headset updates. To assess the lag between the two systems, we measured the onset of platform motion by means of a MEMS triaxial accelerometer (MPU-6050, InvenSense, US), and the onset of visual motion by means of a photodiode (BPW21 Siemens, Germany). We found that platform motion lagged behind visual motion by ~ 40 ms. To compensate for this lag, we shifted the first update of the position of Layer 2 by 40 ms in each trial, so that the onset of visual motion was synchronized with the start of platform motion. The position of Layer 2 was then updated at 90 Hz in a feedforward manner following the same motion profile as the chair.

Stimuli

The path of the stimuli was identical across conditions and sessions (VI, VE and VV), but the direction was either downward or upward. The length of the path (Δp) of either the visual target or the chair was 11.73 cm. For VI, this corresponded to 15.90° of visual angle when the stimulus was at 42 cm from the observer (i.e., the diameter of the aperture plus two radii of the visual target). The stimuli followed a motion profile sinusoidal in acceleration defined by:

$$p(t) = p_0 \pm \frac{\Delta p}{T} \left(t - \frac{T}{2\pi} \sin \frac{2\pi t}{T} \right), \quad (1)$$

where p represents the vertical instantaneous position of the stimulus, t the instantaneous time elapsed from the beginning of motion, p_0 the vertical initial position, and T the total duration of the stimulus. This motion profile was chosen based on previous research on vestibular motion perception⁵².

In each trial, we presented two consecutive stimuli moving vertically in opposite directions: a reference stimulus (R) and a comparison stimulus (C), the order of presentation of the two being randomly permuted, either R first (R/C) or C first (C/R). R had a fixed motion duration of $T = 1.400$ s, while C had 9 possible motion durations (ranging from $T = 0.840$ s to $T = 1.960$ s, in 0.140 s steps, corresponding to peak velocities of 27.93 cm/s for the fastest C and 11.97 cm/s for the slowest C). Thus, all stimuli were at a relatively low frequency (between about 0.5 Hz and 1.2 Hz). In VE sessions, these frequencies corresponded to those thought to maximize vestibular over somesthetic contributions during self-motion perception^{53,54}.

In Fig. 2, we show the stimulus trajectories for all R/C conditions in which the C stimulus was directed downwards. In all panels, from left to right, vertical dotted lines indicate the beginning and the end of the R stimulus and the start of the C stimulus. The position of the chair in world reference coordinates is shown in Fig. 2a, d, g for VI, VE and VV, respectively. All possible durations of the C stimuli are color-coded lines from blue to red, from 0.840 s to 1.960 s, with the 1.400 s condition in black in the middle. In VI, the chair is shown motionless (Fig. 2a). Figure 2b, e, h depicts the visual target position (i.e., Target) in world reference coordinates for VI, VE and VV, respectively. Since the visual scene did not include Layer 1 in VE, the visual target position is absent in Fig. 2e. The visual target was static in VV (Fig. 2i). Finally, Fig. 2c, f, i represents the visual stimuli resulting from the integration of visual and vestibular stimulation for VI, VE and VV, respectively. Notably, the visual stimuli were opposite between equivalent VI and VV conditions, while they were absent in VE conditions.

Motion platform control

Before the calibration phase, the MOOG motion platform was put on *engage*, at the center of its workspace. During the VI session, this position was maintained. In the VE and VV sessions, the motion of the chair was controlled in degrees of freedom at 1000 Hz by means of custom programs written in LabVIEW (2023, National Instruments, USA). All the details about the MOOG control are provided in La Scaleia et al.⁵¹.

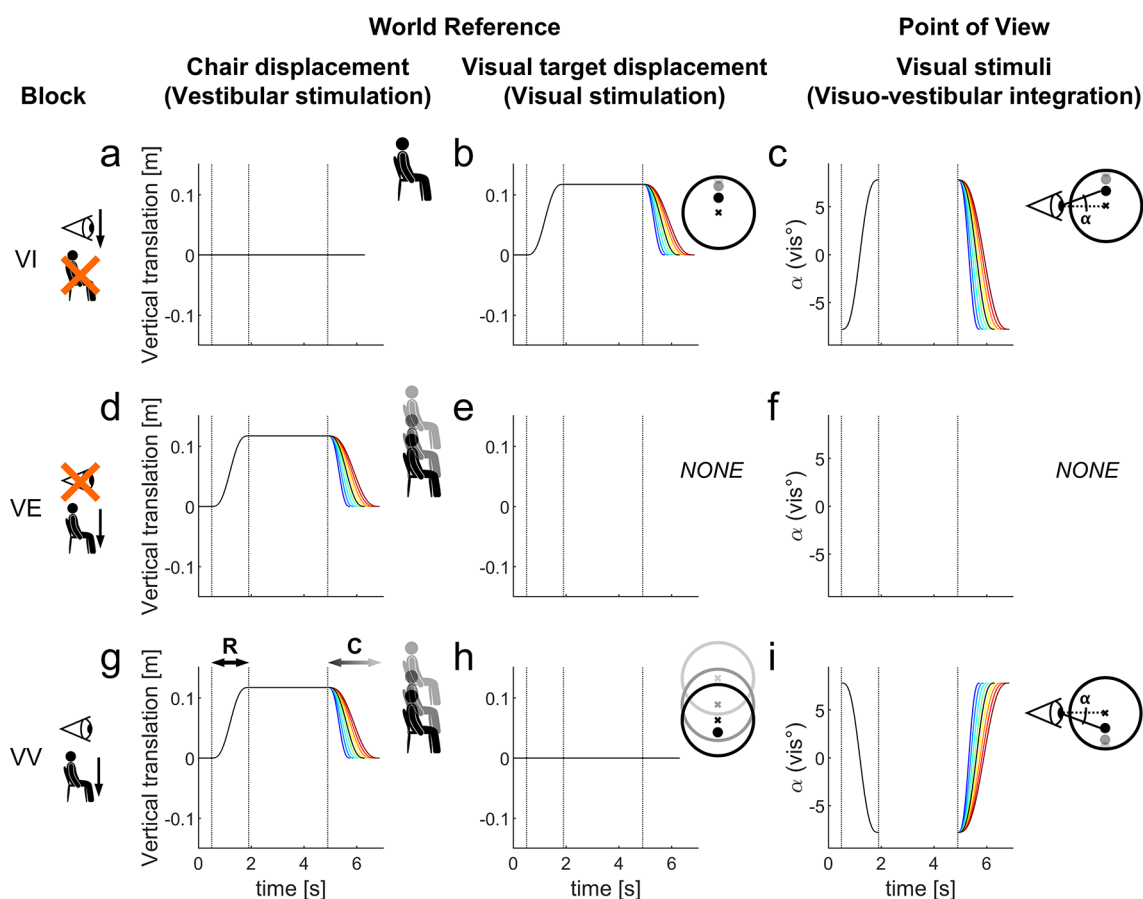


Fig. 2. Vestibular and visual stimuli presented during the execution of the task in R/C conditions for all possible C durations (blue-dark red lines, in black the 1.400 s duration) in the blocks where C stimuli moved downwards (C direction Down); as depicted by the vertical arrows in block icons. VI (a–c), VE (d–f) and VV (g–i) sessions are hereby represented for their chair (a, d, g) and the visual target relative positions (b, e, h) in world reference. From the participant point of view, the visual stimuli visible through the VR-headset, resulting from visual (when present) and vestibular (when present) stimulations, are represented in panels c, f and i for VI, VE and VV sessions, respectively. In this representation the working area is 42 cm in front of the participant, as defined during the calibration phase (see text). In the sessions where the visual (VV) or vestibular (VI) stimulations were static, only a horizontal black line is depicted (as constant were the positions of the visual stimulus, panel h, or the chair, panel a, respectively). In VE no visual stimuli appeared on screen. Consequently, no positions are represented in e and f panels. In each panel from left to right, the three vertical dotted lines represent the beginning and the end of R and the beginning of C stimuli. C/R conditions were identical, changing only the order of presentation of C and R stimuli (not depicted). In C direction Up blocks, the C stimuli moved upwards, while the R stimuli moved downwards (not depicted).

Protocol

Each session included two blocks identified by C direction: one block with C stimuli directed downwards and one block with C stimuli directed upwards. In VI, during the presentation of the two stimuli (R and C) in each trial, the visual target moved along the vertical diameter of the aperture, parallel to the dark-grey surface, following a trajectory defined by Eq. 1. During the VE session, the chair moved vertically according to the same Eq. 1. Finally, in the VV session, where both the visual stimulation and the vestibular stimulation were present, the chair moved as in the VE sessions, while the visual target was static. Therefore, in VV-Down (VV-Up), the chair moved downwards (upwards) for C stimuli. Moreover, in VV-Down (VV-Up), not only was the vestibular stimulus the same as that of the VE-Down (VE-Up) but, since the chair motion disclosed the static visual target, the resulting visual stimulation was comparable with that of VI-Up (VI-Down). See Fig. 2 and the supplementary videos for further details.

Procedure

Throughout each session, white noise was reproduced through the headphones at a volume high enough (75 dB) to ensure acoustic isolation of the participants. The beginning of each trial was signaled by a go-sound (81 dB pure tone at 500 Hz, 125 ms duration). Five hundred milliseconds after the onset of the sound, participants experienced the first stimulus (R or C). Then, the second stimulus (C or R, depending on whether the first stimulus was R or C, respectively) was presented after an interstimulus interval (ISI) of 3 s. This ISI should mitigate vestibular motion after-effects⁵⁵ in VE and VV sessions, and should not involve major memory decays of the first stimulus in any of the sessions, including VI^{42,56–58}. At the end of the second stimulus, a question appeared for 3 s on the screen at the center of the working area. The participants were asked to judge which one of the two stimuli had the shorter duration. The two possible answers were shown with a box on the left and an identical box on the right with “1” and “2” captioned on it, respectively. Responses were given by pressing a button on the Quest Pro left or right controller when the preferred answer was the first or the second stimulus. Visual feedback on the button that was pressed appeared on the screen by highlighting the corresponding box (“1” or “2”), and it stayed there until the question disappeared. In case of no response within the 3 s, no box was lit. No feedback about the correctness of the response was ever given. A new trial started 1 s after the disappearance of the question. In each trial of VI and VV sessions, participants were asked to keep fixation on the fixation cross. In each trial of VE sessions, participants were asked to fixate an imaginary target placed along the sightline and moving together with the subject.

The permutation of the levels of the two factors being manipulated – namely the 2 presentation orders of the stimuli within each trial (R/C or C/R) and the 9 possible durations of C stimuli – led to a total of 18 conditions, repeated 5 times each in a pseudorandom order. Thus, each experimental block consisted of 90 trials (18 conditions x 5 repetitions) for a duration of around 20 min. The two blocks of each session were separated by a 5–30 min break according to participants’ needs, while the three sessions were performed on different days (~6 days apart). Both the sessions and the C direction order of presentation were counterbalanced across participants.

Data analysis

To assess the head stability, we measured in each experimental block the shift of the head in 3D relative to the calibration reference position, and the absolute value of the maximum shift of the head relative to the chair. To assess the maintenance of fixation, in each experimental block we computed average and standard deviation of the eyes shift (AES and SES, respectively) from the reference position (i.e. the fixation cross in VI and VV sessions, or the imaginary target along the sightline estimated at the beginning of each trial in VE) on the horizontal and vertical axis.

We analyzed the psychometric results of the discrimination task by means of a two-level algorithm and a generalized linear mixed model (GLMM)⁵⁹. For the two-level algorithm, first we fitted the valid responses of each participant separately for each block with a psychometric function (or general linear model):

$$\Phi^{-1}[P(Y = 1)] = \alpha_0 + \alpha_1 t, \quad (2)$$

where $P(Y = 1)$ is the probability of reporting that the comparison stimulus had a longer duration than the reference stimulus, Φ^{-1} is the probit link function, α_0 and α_1 are the intercept and slope of the general linear model, respectively, and t is C motion duration. In the model, we disregarded the order in which the two stimuli were presented (C/R or R/C) and combined the responses accordingly. The point of subjective equality (PSE) and the just noticeable difference (JND) were computed from the Eq. (2) as in Moscatelli et al.⁴⁴ for every block of each participant. Briefly, PSE estimates the accuracy of the response and is equal to $-\alpha_0/\alpha_1$, while the JND reflects the precision of the response and corresponds to the inverse of the slope $JND = 1/\alpha_1$.

In the Results (Fig. 3a), we refer the computed PSE values to both an absolute and a relative time scale. The absolute scale makes reference to the time duration of the stimuli we employed. The relative scale makes reference to the fact that the psychometric function can only provide information about how the comparison stimulus (e.g. downward motion) is perceived relative to the reference stimulus (e.g. upward motion). Thus, a shift of the PSE is equivalently due to underestimation of the duration of downward movement, overestimation of the duration of upward movement, or some combination of the two.

Finally, we also defined a GLMM to take into account the variability of the parameters estimates between subjects⁵⁹. We fitted the data from different participants and blocks with a GLMM of the form:

$$\Phi^{-1}[P(Y = 1)] = \beta_0 + u_0 + (\beta_1 + u_1)t + \beta_2 s + \beta_3 d + \beta_4 ts + \beta_5 sd, \quad (3)$$

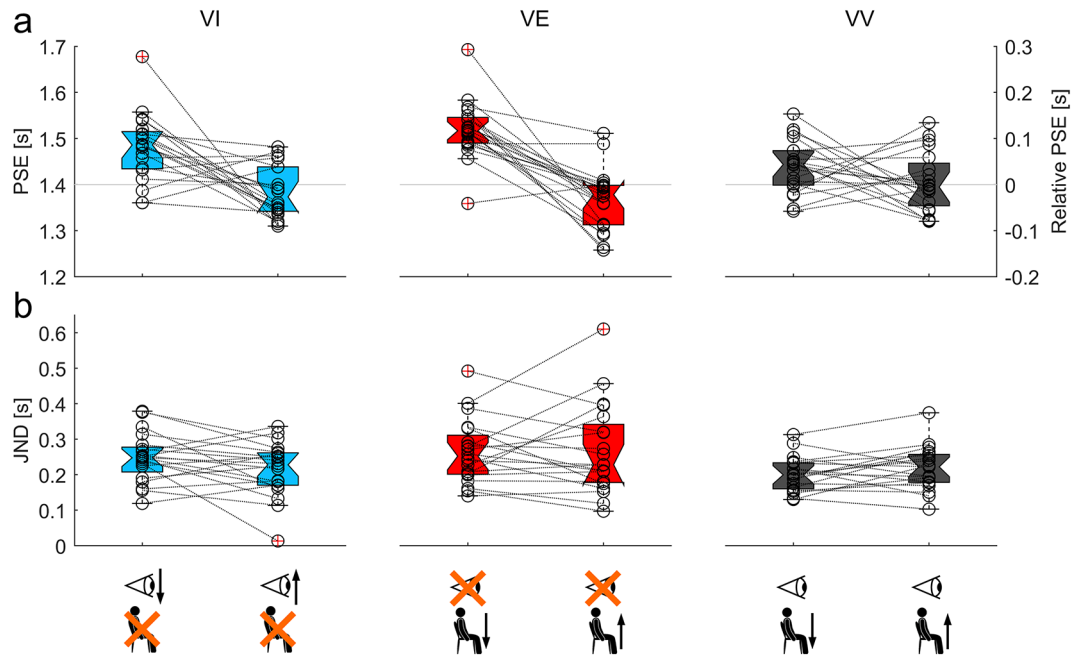


Fig. 3. Boxplots for PSEs (a) and JNDs (b) obtained from all subjects in the two C directions of VI, VE and VV sessions. In panel a, PSE values are referred to an absolute time scale on the left y-axis, and to a relative time scale on the right y-axis. In light blue, red and dark grey data from VI, VE and VV sessions, respectively. Block icons in the same format as Fig. 2. VV blocks are depicted according to their vestibular stimulation. Notches indicate 95%CI for the median. Data marked with red crosses indicate outliers ($> \pm 2.7$ SD). In the a) panel, the continuous horizontal line represents the 1.400 s duration (i.e. the duration of R stimuli).

where the differences from the psychometric function are related to t , s , d accounting for the predictor variables, motion duration (t), session (s) and C direction (d). The parameters β and u represent fixed-effect and random-effects, respectively. We selected the model based on the lowest Bayesian Information Criterion (BIC), similarly to the procedures described in Moscatelli et al.⁵⁹. Estimates on PSEs and JNDs were computed with a 95% Confidence Interval (95%CI) by means of the bootstrap method (2000 bootstrap sampled datasets).

In the Results section, we define *observed* PSE and JND values those derived from Eq. 2, and *estimated* PSE and JND values those derived from Eq. 3.

We used custom programs in MATLAB (R2023a, Mathworks, US) for data analysis.

Statistics

Shapiro-Wilk test was used to assess the normality of the data distributions (alpha level=0.05). For normally distributed data, we report mean values and 95%CI of the selected parameter over all participants ($n=20$). Next, we defined a repeated measures ANOVA (RM-ANOVA) model, with session (3 levels: VI, VV and VE), direction of C stimuli (2 levels: Down and Up) and their interaction as within subject factors (the Greenhouse-Geisser correction was applied when the data did not meet the sphericity assumption). We performed a planned comparison by means of a pairwise comparison (t-tests, Bonferroni corrected). In particular, we excluded the comparison between VV and VI sessions within the same direction of C stimuli given that the visual stimulations in VV blocks were in opposite directions than those in corresponding VI blocks (see Fig. 2 panels c and i). By contrast, we compared VE-Down (VE-Up) to VV-Down (VV-Up) and VI-Down (VV-Up) since the stimulations were in the same direction (see Fig. 2 panels d and g and Fig. 2 panels d and b for VE to VV and VE to VI, respectively). We performed additional statistical (t-test) analyses to directly compare VV blocks to the VE and VI blocks with similar vestibular and visual stimulations, respectively (i.e. VV-Down to VE-Down and VI-Up on the one side, VV-Up to VE-Up and VI-Down on the other side).

For not-normally distributed data, we report the median value and 95%CI and non-parametric statistics (Friedman test and Wilcoxon signed-rank test for planned comparisons, Bonferroni corrected).

We used custom programs in R (4.3.3, R Core Team) for statistical analyses.

Results

We first report the results about the compliance of the participants to the instruction of keeping the gaze fixed during all trials. Next, we report the psychophysical results of the discrimination of motion duration for the sessions with a visual target (VI), self-motion without an overt visual target (VE), and visuo-vestibular stimuli (VV).

Head stability

Participants generally kept the head in a roughly constant position during the trials. At the start of the trial, the median shift of the head in 3D relative to the calibration reference position was -0.185 cm (95%CI $[-0.582, 0.135]$, $n=60$ (i.e. 3 sessions * 20 participants)), 0.050 cm (95%CI $[-0.217, 0.180]$, $n=60$), -0.756 cm (95%CI $[-0.934, -0.585]$, $n=60$) in x, y and z, respectively; and 0.162° (95%CI $[-0.175, 0.368]$, $n=60$), 1.520° (95%CI $[0.825, 2.281]$, $n=60$), -0.211° (95%CI $[-0.599, 0.116]$, $n=60$) in roll, pitch and yaw, respectively. At the start of the trial, no differences were observed in either displacements or rotations of the head between directions of C stimuli (Friedman test, all $p>0.179$) and session (Friedman test, $p>0.086$), except for roll (Friedman test, $p=0.043$) whose changes were significantly smaller in VI than in VV (Wilcoxon test, $p=0.022$, difference $<1^\circ$) and in VI than in VE (Wilcoxon test, $p=0.036$, difference $<1^\circ$).

During each trial, the median absolute value of the maximum shift of the head relative to the chair position was 0.200 cm (95%CI $[0.174, 0.216]$, $n=60$), 0.177 cm (95%CI $[0.163, 0.192]$, $n=60$), 0.313 cm (95%CI $[0.294, 0.329]$, $n=60$) in x, y and z, respectively; and 0.286° (95%CI $[0.261, 0.319]$, $n=60$), 0.557° (95%CI $[0.513, 0.629]$, $n=60$), 0.264° (95%CI $[0.244, 0.297]$, $n=60$) in roll, pitch and yaw, respectively. The maximum head shift during trial presentation showed no significant differences in either displacements or rotations between directions of C stimuli (Friedman test, all $p>0.370$). On the other side, it depended on the session (Friedman test, all $p<0.005$, except for roll and yaw, $p>0.115$), even though these differences were smaller than 2.5 mm and 0.1° for displacement and rotation, respectively. In particular, greater displacements and pitch rotations were observed in the sessions where the chair moved than in the visual session VI (Wilcoxon test, x direction: VI-VV $p<0.001$ and VI-VE $p=0.132$ after correction; y and z directions: all $p<0.001$ after correction; pitch: all $p<0.004$ after correction). Moreover, no significant differences were observed between VV and VE for displacements along the three axes and pitch rotations (Wilcoxon test, all $p=1.000$ after correction).

Fixation

Participants also maintained a good fixation during the trials. Eye position data showed that the median AES from the reference position was 0.550° of visual angle (95%CI $[0.464, 0.683]$, $n=60$) and 0.873° of visual angle (95%CI $[0.724, 0.963]$, $n=60$) on the horizontal and vertical axis, respectively. The median SES from the reference position was 0.329° of visual angle (95%CI $[0.300, 0.379]$, $n=60$) horizontally and 0.579° of visual angle (95%CI $[0.493, 0.631]$, $n=60$) vertically. Vertical AES and SES were significantly larger than horizontal AES and SES (Wilcoxon signed-rank test, $p<0.002$).

No significant differences were observed in AES or SES between directions of C stimuli (Friedman test, all $p>0.074$) on both vertical and horizontal axis. Vertical AES also did not significantly differ across sessions ($p=0.387$); however, vertical SES were significantly different across sessions (Friedman test, $p=0.004$), though post-hoc tests revealed no significant pairwise differences (Wilcoxon tests, all $p>0.051$ after correction, maximum difference $<0.28^\circ$ of visual angle). Both horizontal AES and SES were significantly different across sessions (Friedman test, all $p<0.005$); however, post-hoc tests for horizontal AES revealed no significant differences (Wilcoxon tests, $p>0.079$ after correction, maximum difference $<0.32^\circ$ of visual angle). On the other side, horizontal SES exhibit significantly lower values in sessions with the fixation cross (i.e. VI and VV) than in VE (Wilcoxon test, all $p<0.001$ after correction, maximum difference $<0.35^\circ$ of visual angle). Noteworthy, previous studies have shown that horizontal eye positions standard deviations were smaller when fixating on a visible target compared to an imaginary target⁶⁰.

PSE values

As explained in the Methods, we computed the psychometric functions (Eq. 2) for each condition and individual. They are plotted as Supplementary Fig. 1, Supplementary Fig. 2 and Supplementary Fig. 3 for the VI, VE and VV condition, respectively. From each psychometric function, we derived the *observed* PSE (point of subjective equality). We found that these PSEs were normally distributed in all blocks (Shapiro-Wilk tests, all $p>0.050$). The individual values and the box plots of the *observed* PSEs are depicted in Fig. 3a. The mean values and 95%CI over all participants are reported in Table 1.

Figure 3a plots PSE values referred to both an absolute time scale (left y-axis) and a relative time scale (right y-axis). In the visual (VI) and vestibular (VE) sessions, we found a consistent trend across subjects for the values of PSE in the two motion directions (up or down) of the comparison (C) stimuli. In particular, we found that

		Observed PSE [s]			Estimated PSE [s]		
Session	C direction	Mean	-95%CI	+ 95%CI	Mean	-95%CI	+ 95%CI
VI	Down	1.481	1.449	1.514	1.477	1.457	1.499
VE		1.520	1.492	1.547	1.512	1.489	1.536
VV		1.439	1.414	1.464	1.437	1.418	1.456
VI	Up	1.385	1.360	1.409	1.384	1.366	1.403
VE		1.367	1.339	1.396	1.370	1.349	1.391
VV		1.406	1.378	1.434	1.405	1.388	1.424

Table 1. Observed and estimated values (mean \pm 95%CI) for PSE. Observed values are based on the individual psychometric functions defined on participants’ responses. Estimated values are the estimates of PSE estimated by the GLMM.

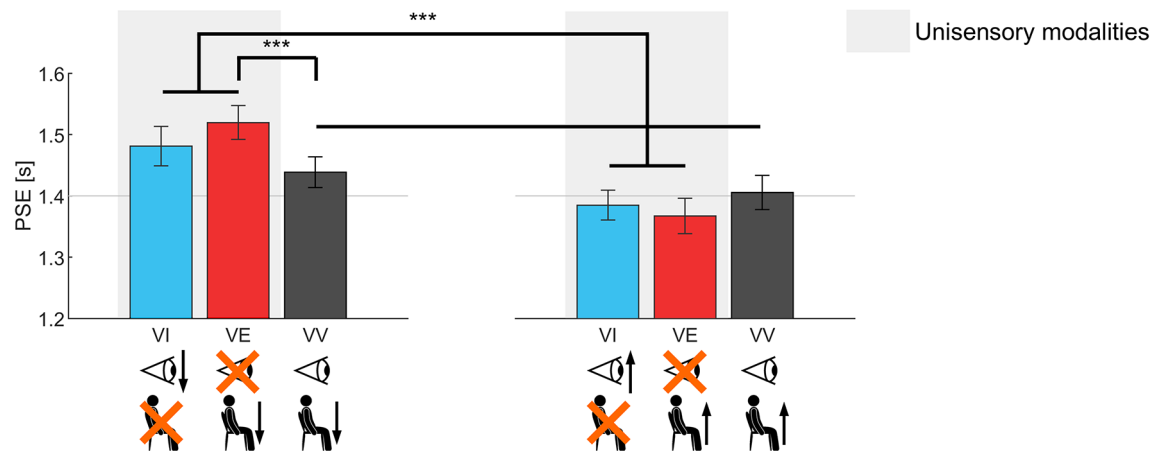


Fig. 4. Bar plot for mean PSE values (error vertical line representing 95%CI) in C Down (left) and C Up (right), same icons and color code as Fig. 3. RM-ANOVA for PSEs highlighted a significant effect of C direction and of the interaction between session and C direction. Planned pairwise comparisons highlighted that the latter effect is due to higher PSEs in VI-Down and VE-Down than in VI-Up and VE-Up, respectively, and to higher PSEs in VE-Down than in VV-Down. VV blocks are depicted according to the C direction of their vestibular stimulation. (**, $p < 0.010$. ***, $p < 0.001$).

factors	df1	df2	PSE		JND	
			F	p	F	p
s	2	38	1.658	0.204	5.789	0.006
d	1	19	53.387	<0.001	0.188	0.670
s * d	2	38	6.263	0.006	1.915	0.165

Table 2. RM-ANOVA results for PSE and JND in the three sessions “s” (VI, VE and VV) and the two C directions “d” (Comparison -Down and comparison -Up).

the mean PSEs over all participants were significantly ($p < 0.050$) larger than 1.400 s (i.e. the duration of R, the reference stimulus) when the C stimuli were directed downwards in both VI-Down and VE-Down blocks (right tailed one sample t-tests, all $p < 0.001$) (see Fig. 3a). Conversely, the mean PSEs tended to be smaller than 1.400 s when the C stimuli were directed upwards in both VI-Up (left tailed one sample t-test, $p = 0.121$) and VE-Up (left tailed one sample t-test, $p = 0.020$). In other words, downward motion of the comparison stimuli was perceived as of the same duration as the upward motion of the reference stimulus only when it lasted longer than the reference, and vice versa for comparison stimuli moving upwards. At the individual level, the PSE for VI-Down was greater than the PSE for VI-Up in 16/20 participants, and the PSE for VE-Down was greater than the PSE for VE-Up in 19/20 participants.

The results were less consistent for the visuo-vestibular (VV) blocks (Fig. 3a). While the mean PSE over all participants was significantly (right tailed one sample t-test, $p < 0.004$) larger than the reference duration when the C stimuli were directed downwards (VV-Down), the mean PSE was not significantly different (left tailed one sample t-test, $p = 0.655$) from the reference when the C stimuli were directed upwards (VV-Up). Moreover, the trend of the PSE values as a function of the direction of motion of C stimuli was quite variable across participants. Thus, the PSE for VV-Down was greater than the PSE for VV-Up in 12/20 participants, and it was lower than the PSE for VV-Up in the remaining 8 participants.

Confirming the previous observations, the RM-ANOVA model for PSE showed a strong dependency of the observed PSEs on the C motion direction ($F(1,19) = 53.387$, $p < 0.001$ Greenhouse-Geisser corrected), and on the interaction between sessions (VI, VE, VV) and C motion direction (Up, Down) ($F(2,38) = 6.263$, $p = 0.006$, Greenhouse-Geisser corrected. See Fig. 4; Table 2). Planned pairwise comparisons highlighted a significant difference between the two directions of C stimuli in both VI and VE ($p \leq 0.001$, Bonferroni corrected), with larger PSEs in VI-Down and VE-Down. No significant difference emerged between the two directions of C stimuli in VV, VV-Down and VV-Up ($p = 0.160$ after Bonferroni correction).

Remarkably, the PSEs were not significantly different in VI and VE when the C motion direction was the same ($p = 0.688$ for the comparison VI-Down vs. VE-Down, and $p = 0.184$ for the comparison VI-Up vs. VE-Up, Bonferroni corrected). However, the PSE in VV-Down was significantly shorter than in VE-Down ($p < 0.001$ after Bonferroni correction). The PSE in VV-Up was longer than in VE-Up, but the difference was not statistically significant ($p = 0.160$, after correction).

Since we found different PSE values when we applied a vestibular stimulation without overt visual stimuli in VE blocks and when we applied a vestibular stimulation in the presence of a static visual target in the VV blocks,

we directly compared VV blocks with the VE and VI blocks with similar vestibular and visual stimulations, respectively (Fig. 5a and b). We found that the mean PSE value was significantly smaller in VI-Up than in VV-Down (paired t-test, $p=0.006$ after correction), despite the visual target shifted upwards relative to the subjects' fixation point in both cases. In addition, as previously noticed, the mean PSE value was significantly smaller in VV-Down than in VE-Down (paired t-test, $p<0.001$ after correction), despite the subjects were displaced downwards in both cases (Fig. 5a). Symmetrically, the mean PSE value was smaller in VV-Up than in VI-Down (paired t-test, $p=0.003$ after correction), as it was the mean PSE in VE-Up relative to that in VV-Up (however, this difference was not significant, paired t-test $p=0.160$ after correction) (Fig. 5b).

To account for the variability of the results between subjects, we also performed a GLMM (Eq. 3). The model parameters are reported in Table 3. The PSE values *estimated* by the GLMM are reported in Table 1, and are quite comparable to the *observed* PSE values. The psychometric functions derived from the GLMM are plotted for VV-Down (Fig. 5c) and VV-Up (Fig. 5d) so as to compare VV blocks with the VE and VI blocks with similar vestibular and visual stimulations, respectively.

JND values

In addition to the *observed* PSE, the individual psychometric functions (Eq. 2) yielded the value of *observed* JND (the just noticeable difference), which is inversely related to the precision of discrimination. We found that these JNDs were normally distributed in all blocks (Shapiro-Wilk tests, all $p>0.050$). The individual values and the box plots of the *observed* JNDs are shown in Fig. 3b. The mean values and 95% CI over all participants are reported in Table 4.

For all sessions (VI, VE, VV), the precision was similar between the Up and Down directions of C (paired t-tests, all $p>0.076$) (Fig. 3b; Table 4). Moreover, the precision in VI and VE tended to be lower (i.e. higher JNDs) than in VV. Accordingly, RM-ANOVA of the *observed* JNDs showed a significant difference across sessions ($F(2,38)=5.789$, $p=0.009$, Greenhouse-Geisser corrected. See Table 2), mainly due to significantly higher JND

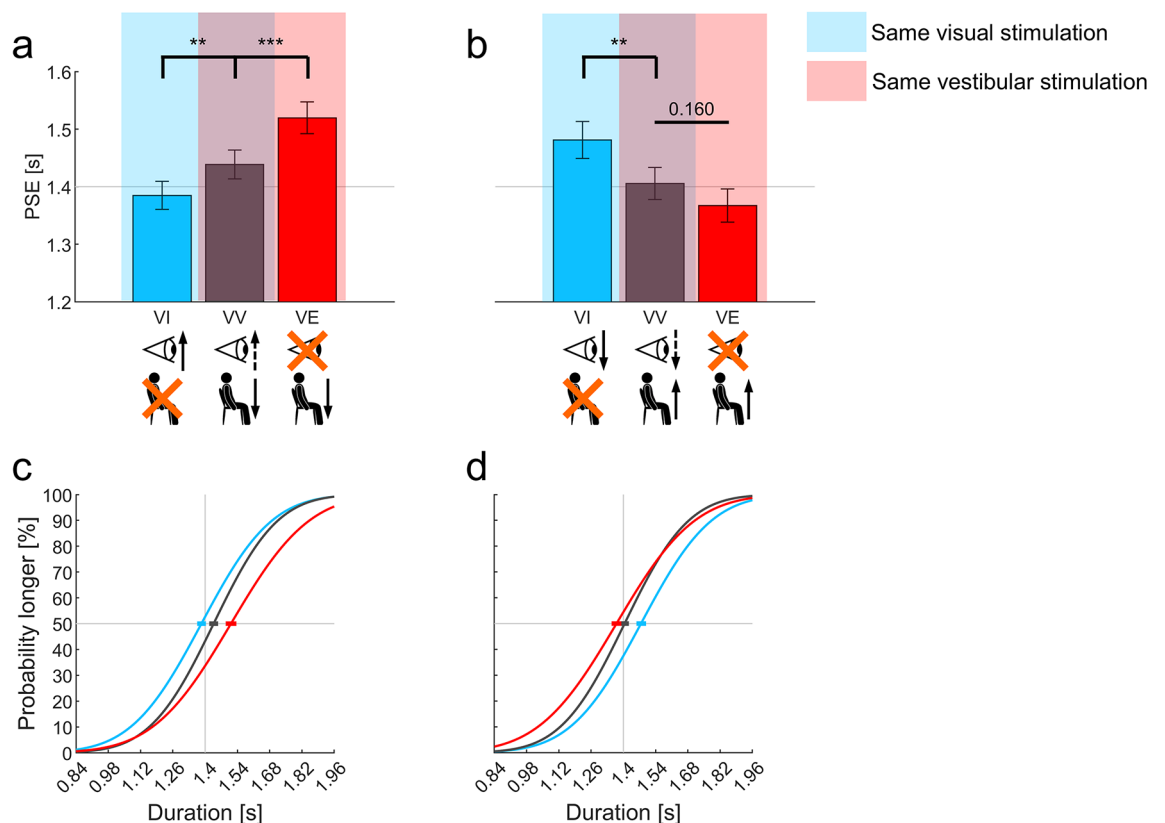


Fig. 5. Bar plot for mean PSE values (error vertical line representing 95%CI) rearranged to match VV blocks with VI and VE blocks that correspond to equivalent visual and vestibular stimulations, respectively (panel a: VI-Up, VV-Down and VE-Down blocks; panel b: VI-Down, VV-Up and VE-Up blocks). Planned t-tests highlighted how PSEs from VV blocks fall in between values from unisensory conditions with correspondent visual and vestibular stimulations (**, $p<0.010$. ***, $p<0.001$). Continuous arrows represent the C direction in the same format as Fig. 2. Dashed arrows represent the visual C direction resulting from chair motion. Panels c and d represent the psychometric functions defined from *estimated* PSE and JND (error horizontal line representing 95%CI on the PSEs) with the same rearrangement of panels a and b, respectively. Same color code as Fig. 3.

Parameter	Estimate	Standard Error	z value	Pr(> z)
(Intercept)	-5.683	0.272	-20.891	<0.001
t	3.758	0.190	19.714	<0.001
s VI	-0.473	0.228	-2.080	0.038
s VV	-0.908	0.237	-3.828	<0.001
d Up	0.533	0.056	9.531	<0.001
t * s VI	0.411	0.149	2.764	0.006
t * s VV	0.829	0.157	5.268	<0.001
s VI * d Up	-0.147	0.080	-1.835	0.067
s VV * d Up	-0.388	0.082	-4.748	<0.001

Table 3. GLMM fixed effects parameters. *t*, motion duration; *s*, session; *d*, direction of C stimuli.

		Observed JND [s]			Estimated JND [s]		
Session	C direction	Mean	-95%CI	+ 95%CI	Mean	-95%CI	+ 95%CI
VI	Down	0.248	0.218	0.278	0.239	0.218	0.260
VE		0.265	0.225	0.304	0.265	0.242	0.288
VV		0.201	0.180	0.223	0.217	0.200	0.235
VI	Up	0.215	0.181	0.248	0.239	0.218	0.260
VE		0.268	0.213	0.324	0.265	0.242	0.288
VV		0.219	0.192	0.246	0.217	0.200	0.235

Table 4. As for Table 1 for JND. Please note that estimated JNDs are identical across sessions of corresponding sensory modalities (see Methods and Table 3 for further details on the characterization of the GLMM).

values observed in VE than in VV ($p=0.021$, Bonferroni corrected). JNDs were higher in VE than in VV in 11/20 and 18/20 subjects in the Down and Up conditions, respectively.

The results obtained with the GLMM (Eq. 3) confirmed those reported above. The values of *estimated* JNDs were similar to the values of *observed* JNDs (see Table 4).

Discussion

In order to investigate whether the internal model of gravity affects visual, vestibular and visuo-vestibular perception of vertical motion duration, we asked participants to compare the durations of two consecutive stimuli moving in opposite vertical directions. The stimuli consisted of visual targets moving relative to the participant (VI), passive whole-body movements (VE) or their combination (VV), presented in three separate sessions. To reveal the presence of perceptual biases, we computed the point of subjective equality (PSE) of the psychometric functions with two different methods, and obtained very similar results.

Consistent with the prediction of the internal model of gravity, we found that downwardly directed motions of either the visual target (VI) or the participant (VE) were judged as lasting less (greater PSE) than upward motions of the same duration, and vice-versa for the opposite directions of motion. Moreover, the average PSEs over all participants for the corresponding visual and vestibular blocks (i.e. with the comparison stimuli moving up or down in both VI and VE blocks) were not significantly different, indicating that a similar perceptual bias related to gravity affects both sensory modalities. Our protocols involving both upward and downward motion in the same block (one direction for the comparison stimulus and the other one for the reference stimulus) did not allow separate estimates of precision (JND) for downward and upward motions.

The present results with visual stimuli are consistent with previous findings by Moscatelli et al. (2019) who also found a downward bias⁴⁴. In the latter study, the observers were asked to compare the speed of downward versus upward motions, while motion duration was made unreliable as a cue by randomizing path length. In the present study, the participants were asked to compare motion durations, but since stimulus speed and acceleration covaried with duration, in line of principle they could have used these kinematic cues to estimate duration. We will return to this potential confound below. The present results also agree with the previous findings by Senot et al.²⁷ that observers trigger hand movements to intercept a virtual target earlier when the target comes from above than when it comes from below, despite both target motions have the same durations and irrespective of their law of motion (1g, 0g, -1g). Therefore, the present and the previous studies reveal a bias consistent with the prior that downward motion is accelerated by gravity while upward motion is decelerated by gravity.

The present results with vestibular stimuli revealing an up-down bias identical to the visual bias are novel. To our knowledge, previous studies focused on the discrimination of up-down directions in the dark, rather than motion duration (or speed) as in the present protocol. Some studies reported a higher precision (lower discrimination thresholds) for downward direction than for upward direction^{48,61–63}, while other studies did not find a significant difference^{64–66}. None of the above studies reported significant biases (asymmetries) between downward and upward directions in the upright posture (as in our participants), but their main focus was on the

precision of discrimination. However, Kobel et al. (2021)⁴⁸ found a main effect of motion relative to gravity due to a larger positive bias in earth-vertical motions than earth-horizontal motions for 2-Hz displacements. Since no significant bias was found for 1-Hz displacements where vestibular cues dominate over somesthetic cues^{53,54}, the bias at 2-Hz could have resulted from up-down asymmetries in tactile cues^{48,62}. However, the general conclusion of Kobel's study is in agreement with the hypothesis that internal models of gravity influence the perception of vertical translations, since they found that earth-vertical thresholds, where the translation stimulus must be disambiguated from the colinear gravitational acceleration, were significantly higher than earth-horizontal thresholds, where the translation is independent of (i.e., perpendicular to) gravitational acceleration⁴⁸.

While we cannot rule out the influence of somesthetic cues in the responses to up-down displacements of the participants, we notice that the present stimuli had frequencies between about 0.5 Hz and 1.2 Hz, thus near the range where vestibular cues are known to dominate over somesthetic cues^{53,54}. Bruschetta et al. (2021)⁶⁷ have specifically dissected the role of somesthetic versus vestibular cues in the perception of longitudinal and lateral motions in the dark. They found that somesthetic cues play a significant role for strong accelerations, while our stimuli were smooth ramp and hold waveforms. Moreover, it is known that the perception of earth-vertical translations –such as the present ones– is especially compromised by complete bilateral vestibular ablation, as shown by very high thresholds of vertical direction discrimination in these patients^{53,54,68}. Nevertheless, even if the perceptual responses during VE (and VV) sessions included some non-vestibular sensory (e.g., somesthetic and visceral) contributions in addition to vestibular contributions, this would not change the conclusion that the judgment of duration of vertical whole-body motion is biased by the prior assumption about the effects of gravity.

As mentioned above, in the present experiments stimulus speed and acceleration covaried with duration, and the participants could have used these kinematic cues to estimate duration. Accordingly, the observed bias in the estimates of time duration might derive from misperceived speed or acceleration of the stimuli. In fact, time per se might not be measured directly by the nervous system, but it might be estimated by integrating appropriate informations over discrete intervals using internally generated and/or externally triggered signals⁶. If this is the case, any kinematic variable monitored by the sensory apparatus, including but not limited to speed and acceleration, might be used by subjects to judge motion durations.

We do not know where in the brain the up-down visual and (primarily) vestibular biases arise. Directional anisotropies in neural responses to either visual^{69,70} or vestibular⁷¹ stimuli have been described, mainly related to the directional preference for stimuli oriented along the cardinal axes relative to oblique stimuli. However, we are not aware of neural responses that can be reconciled with the downward bias that we described. For instance, in the monkey the responses of primary vestibular afferents preferentially excited by upward translations are similar to the responses of afferents preferentially excited by downward translations⁷². Also the responses of Medial Superior Temporal (MST) area neurons to upward and downward body translations are similar⁷¹. By the same token, multiple measures of Middle Temporal (MT) visual area neuronal responses do not provide evidence of a directional anisotropy to visual motion⁷³.

Downward biases may result from processing at neural stages downstream of MT/MST, in particular at the level of visual-vestibular brain regions, such as the parieto-insular vestibular cortex where the internal model of gravity effects has been shown to be encoded by means of functional magnetic resonance imaging (fMRI)^{20,74–77} and transcranial magnetic stimulation (TMS) studies^{78,79}. Thus, neural populations in the temporo-parietal junction, posterior insula, retro-insula and OP2 in the parietal operculum are selectively engaged by vertical visual motion of objects and self-motion coherent with gravity, as well as by vestibular stimuli^{21,74,75,80–82}. Patients with lesions of the temporo-parietal junction and peri-insular regions show specific deficits in the perceptual estimates of passive motion durations in the dark⁴⁷, as well as deficits in the processing of visual targets accelerated by gravity⁸³.

On the other hand, we did not expect systematic biases in the visual-vestibular sessions (VV), given that the vestibular stimulus was in the opposite direction of the visual stimulus, with the result that the visual and vestibular bias should cancel out. In fact, on average there was no significant difference in the bias (PSE) between the block with downward motion of the participant and that with upward motion. However, this lack of significant difference of the average values did not seem to depend on the individual values of the visual and vestibular biases cancelling out. It rather depended on a large inter-subject variability of PSE values: in about half of the participants, the PSE for downward self-motion was greater than that for upward self-motion, but in the other half of the participants the opposite was true. This result suggests that each subject weighed relatively more one cue (visual or vestibular) than the other one, since visual motion and vestibular motion were in opposite directions. Inter-subject variability might depend on the different way each participant resolved the potential ambiguity inherent in our VV protocol. While the participants underwent whole-body vertical motion, the visual target was static throughout and it only appeared to move in the opposite direction of the body due to body motion.

Nevertheless, some evidence for visual-vestibular fusion in VV sessions is provided by the significantly higher average precision (lower JND) than in the vestibular sessions (VE).

Conclusion

We confirmed and extended previous findings that downward visual targets are perceived as lasting less (or being faster) than upward targets. We further showed that the same bias occurs with whole-body motion. When visual and vestibular cues were combined together, on average the downward bias disappeared. Overall, the results suggest that visual and vestibular modalities of motion perception share the same prior assumption about the effects of gravity, putatively encoded in the parieto-insular vestibular cortex.

Data availability

Data analyzed in this study are included in the published article and its online supplementary file. Additional data are available from the corresponding authors upon reasonable request.

Received: 11 December 2024; Accepted: 14 March 2025

Published online: 26 March 2025

References

- Hansen, B. C. & Essock, E. A. A horizontal bias in human visual processing of orientation and its correspondence to the structural components of natural scenes. *J. Vis.* **4**, 5 (2004).
- Carriot, J., Jamali, M., Chacron, M. J. & Cullen, K. E. Statistics of the Vestibular Input Experienced during Natural Self-Motion: Implications for Neural Processing. *J. Neurosci.* **34**, 8347–8357 (2014).
- Buhusi, C. V. & Meck, W. H. What makes us tick? Functional and neural mechanisms of interval timing. *Nat. Rev. Neurosci.* **6**, 755–765 (2005).
- Ivry, R. B. & Schlerf, J. E. Dedicated and intrinsic models of time perception. *Trends Cogn. Sci.* **12**, 273–280 (2008).
- Paton, J. J. & Buonomano, D. V. The Neural Basis of Timing: Distributed Mechanisms for Diverse Functions. *Neuron* **98**, 687–705 (2018).
- Lacquaniti, F. et al. Gravity in the Brain as a Reference for Space and Time Perception. *Multisensory Res.* **28**, 397–426 (2015).
- Werkhoven, P., Snippe, H. P. & Alexander, T. Visual processing of optic acceleration. *Vis. Res.* **32**, 2313–2329 (1992).
- Fernández, C. & Goldberg, J. M. Physiology of peripheral neurons innervating otolith organs of the squirrel monkey. I. Response to static tilts and to long-duration centrifugal force. *J. Neurophysiol.* **39**, 970–984 (1976).
- Zago, M., McIntyre, J., Senot, P. & Lacquaniti, F. Internal models and prediction of visual gravitational motion. *Vis. Res.* **48**, 1532–1538 (2008).
- Sanborn, A. N., Mansinghka, V. K. & Griffiths, T. L. Reconciling intuitive physics and Newtonian mechanics for colliding objects. *Psychol. Rev.* **120**, 411–437 (2013).
- Merfeld, D. M., Clark, T. K., Lu, Y. M. & Karmali, F. Dynamics of individual perceptual decisions. *J. Neurophysiol.* **115**, 39–59 (2016).
- Jörges, B. & López-Moliner, J. Gravity as a Strong Prior: Implications for Perception and Action. *Front. Hum. Neurosci.* **11**, (2017).
- Hubbard, T. L. Representational gravity: Empirical findings and theoretical implications. *Psychon. Bull. Rev.* **27**, 36–55 (2020).
- Delle Monache, S. et al. Interception of vertically approaching objects: temporal recruitment of the internal model of gravity and contribution of optical information. *Front. Physiol.* (2023).
- Huang, T. & Liu, J. A stochastic world model on gravity for stability inference. *eLife* **12**, (2024).
- Lacquaniti, F. & Maioli, C. The role of preparation in tuning anticipatory and reflex responses during catching. *J. Neurosci.* **9**, 134–148 (1989).
- Merfeld, D. M., Young, L. R., Oman, C. M. & Shelhamer, M. J. A multidimensional model of the effect of gravity on the spatial orientation of the monkey. *J. Vestib. Res.* **3**, 141–161 (1993).
- Hubbard, T. L. Environmental invariants in the representation of motion: Implied dynamics and representational momentum, gravity, friction, and centripetal force. *Psychon. Bull. Rev.* **2**, 322–338 (1995).
- Zago, M., McIntyre, J., Senot, P. & Lacquaniti, F. Visuo-motor coordination and internal models for object interception. *Exp. Brain Res.* **192**, 571–604 (2009).
- Lacquaniti, F. et al. Visual gravitational motion and the vestibular system in humans. *Front. Integr. Neurosci.* **7**, 101 (2013).
- Delle Monache, S. et al. Watching the Effects of Gravity. Vestibular Cortex and the Neural Representation of Visual Gravity. *Front. Integr. Neurosci.* **15**, 1–17 (2021).
- Lacquaniti, F. & Maioli, C. Adaptation to suppression of visual information during catching. *J. Neurosci.* **9**, 149–159 (1989).
- La Scaleia, B., Zago, M. & Lacquaniti, F. Hand interception of occluded motion in humans: a test of model-based vs. on-line control. *J. Neurophysiol.* **114**, (2015).
- McIntyre, J., Zago, M., Berthoz, A. & Lacquaniti, F. Does the brain model Newton's laws? *Nat. Neurosci.* **4**, 693–694 (2001).
- Zago, M. et al. Internal Models of Target Motion: Expected Dynamics Overrides Measured Kinematics in Timing Manual Interceptions. *J. Neurophysiol.* **91**, 1620–1634 (2004).
- Zago, M., Iosa, M., Maffei, V. & Lacquaniti, F. Extrapolation of vertical target motion through a brief visual occlusion. *Exp. Brain Res.* **201**, 365–384 (2010).
- Senot, P., Zago, M., Lacquaniti, F. & McIntyre, J. Anticipating the effects of gravity when intercepting moving objects: Differentiating up and down based on nonvisual cues. *J. Neurophysiol.* **94**, 4471–4480 (2005).
- Mijatović, A., La Scaleia, B., Mercuri, N., Lacquaniti, F. & Zago, M. Familiar trajectories facilitate the interpretation of physical forces when intercepting a moving target. *Exp. Brain Res.* **232**, 3803–3811 (2014).
- La Scaleia, B., Lacquaniti, F. & Zago, M. Body orientation contributes to modelling the effects of gravity for target interception in humans. *J. Physiol.* **597**, 2021–2043 (2019).
- Senot, P. et al. When up is down in 0 g: How gravity sensing affects the timing of interceptive actions. *J. Neurosci.* **32**, 1969–1973 (2012).
- Bosco, G. et al. Filling gaps in visual motion for target capture. *Front. Integr. Neurosci.* **9**, 1–17 (2015).
- Delle Monache, S., Lacquaniti, F. & Bosco, G. Ocular tracking of occluded ballistic trajectories: Effects of visual context and of target law of motion. *J. Vis.* **19**, 13 (2019).
- Jörges, B. & López-Moliner, J. Earth-gravity congruent motion facilitates ocular control for pursuit of parabolic trajectories. *Sci. Rep.* **9**, 14094 (2019).
- Ke, S. R., Lam, J., Pai, D. K. & Sperling, M. Directional asymmetries in human smooth pursuit eye movements. *Invest. Ophthalmol. Vis. Sci.* **54**, 4409–4421 (2013).
- Angelaki, D. E., McHenry, M. Q., Dickman, J. D., Newlands, S. D. & Hess, B. J. Computation of inertial motion: neural strategies to resolve ambiguous otolith information. *J. Neurosci.* **19**, 316–327 (1999).
- Merfeld, D. M., Zupan, L. & Peterka, R. J. Humans use internal models to estimate gravity and linear acceleration. *Nature* **398**, 615–618 (1999).
- Merfeld, D. M., Park, S., Gianna-Poulin, C., Black, F. O. & Wood, S. Vestibular perception and action employ qualitatively different mechanisms. I. Frequency response of VOR and perceptual responses during translation and tilt. *J. Neurophysiol.* **94**, 186–198 (2005).
- Zago, M. Perceptual and Motor Biases in Reference to Gravity. in *Spatial Biases in Perception and Cognition* (ed Hubbard, T. L.) 156–166 (Cambridge University Press, Cambridge, doi:<https://doi.org/10.1017/9781316651247.011>. (2018).
- McCloskey, M., Caramazza, A. & Green, B. Curvilinear motion in the absence of external forces: Naive beliefs about the motion of objects. *Science* **210**, 1139–1141 (1980).
- Tresilian, J. R. Perceptual and cognitive processes in time-to-contact estimation: Analysis of prediction-motion and relative judgment tasks. *Percept. Psychophys.* **57**, 231–245 (1995).

41. Zago, M. & Lacquaniti, F. Cognitive, perceptual and action-oriented representations of falling objects. *Neuropsychologia* **43**, 178–188 (2005).
42. Moscatelli, A. & Lacquaniti, F. The weight of time: Gravitational force enhances discrimination of visual motion duration. *J. Vis.* **11**, (2011).
43. Torok, A., Gallagher, M., Lasbareilles, C. & Ferrè, E. R. Getting ready for Mars: How the brain perceives new simulated gravitational environments. *Q. J. Exp. Psychol.* **72**, 2342–2349 (2019).
44. Moscatelli, A., La Scaleia, B., Zago, M. & Lacquaniti, F. Motion direction, luminance contrast, and speed perception: An unexpected meeting. *J. Vis.* **19**, 16 (2019).
45. Navarro Morales, D. C., Kuldavletova, O., Quarck, G., Denise, P. & Clément, G. Time perception in astronauts on board the International Space Station. *Npj Microgravity*. **9**, 6 (2023).
46. Denise, P., Harris, L. R., Clément, G. & Editorial Role of the vestibular system in the perception of time and space. *Front. Integr. Neurosci.* **16**, (2022).
47. Kaski, D. et al. Temporoparietal encoding of space and time during vestibular-guided orientation. *Brain* **139**, 392–403 (2016).
48. Kobel, M. J., Wagner, A. R. & Merfeld, D. M. Impact of gravity on the perception of linear motion. *J. Neurophysiol.* **126**, 875–887 (2021).
49. Wei, S., Bloemers, D. & Rovira, A. A Preliminary Study of the Eye Tracker in the Meta Quest Pro. in *Proceedings of the ACM International Conference on Interactive Media Experiences* 216–221 (Association for Computing Machinery, New York, NY, USA, 2023). (2023). <https://doi.org/10.1145/3573381.3596467>
50. Aziz, S., Lohr, D. J., Friedman, L. & Komogortsev, O. Evaluation of Eye Tracking Signal Quality for Virtual Reality Applications: A Case Study in the Meta Quest Pro. in *Proceedings of the Symposium on Eye Tracking Research and Applications* 1–8 (Association for Computing Machinery, New York, NY, USA, 2024). (2024). <https://doi.org/10.1145/3649902.3653347>
51. La Scaleia, B., Brunetti, C., Lacquaniti, F. & Zago, M. Head-centric computing for vestibular stimulation under head-free conditions. *Front. Bioeng. Biotechnol.* **11**, 1296901 (2023).
52. Lacquaniti, F., La Scaleia, B. & Zago, M. Noise and vestibular perception of passive self-motion. *Front. Neurol.* **14**, 1159242 (2023).
53. Valko, Y., Lewis, R. F., Priesol, A. J. & Merfeld, D. M. Vestibular labyrinth contributions to human whole-body motion discrimination. *J. Neurosci.* **32**, 13537–13542 (2012).
54. Kobel, M. J., Wagner, A. R. & Merfeld, D. M. Vestibular contributions to linear motion perception. *Exp. Brain Res.* **242**, 385–402 (2024).
55. Crane, B. T. Roll aftereffects: influence of tilt and inter-stimulus interval. *Exp. Brain Res.* **223**, 89–98 (2012).
56. Bigelow, J. & Poremba, A. Achilles' ear? Inferior human short-term and recognition memory in the auditory modality. *PLoS ONE*. **9**, e89914 (2014).
57. Freyd, J. J. & Johnson, J. Q. Probing the time course of representational momentum. *J. Exp. Psychol. Learn. Mem. Cogn.* **13**, 259–268 (1987).
58. De Sá Teixeira, N. A. The visual representations of motion and of gravity are functionally independent: Evidence of a differential effect of smooth pursuit eye movements. *Exp. Brain Res.* **234**, 2491–2504 (2016).
59. Moscatelli, A., Mezzetti, M. & Lacquaniti, F. Modeling psychophysical data at the population-level: The generalized linear mixed model. *J. Vis.* **12**, 26–26 (2012).
60. Morisita, M. & Yagi, T. The stability of human eye orientation during visual fixation and imagined fixation in three dimensions. *Auris Nasus Larynx*. **28**, 301–304 (2001).
61. Gurnee, H. Thresholds of vertical movement of the body. *J. Exp. Psychol.* **17**, 270–285 (1934).
62. Nesti, A., Barnett-Cowan, M., MacNeilage, P. R. & Bühlhoff, H. H. Human sensitivity to vertical self-motion. *Exp. Brain Res.* **232**, 303–314 (2014).
63. Roditi, R. E. & Crane, B. T. Directional asymmetries and age effects in human self-motion perception. *J. Assoc. Res. Otolaryngol.* **13**, 381–401 (2012).
64. MacNeilage, P. R., Banks, M. S., DeAngelis, G. C. & Angelaki, D. E. Vestibular heading discrimination and sensitivity to linear acceleration in head and world coordinates. *J. Neurosci.* **30**, 9084–9094 (2010).
65. Crane, B. T. Human visual and vestibular heading perception in the vertical planes. *J. Assoc. Res. Otolaryngol.* **15**, 87–102 (2014).
66. Hummel, N., Cuturi, L. F., MacNeilage, P. R. & Flanagan, V. L. The effect of supine body position on human heading perception. *J. Vis.* **16**, 19 (2016).
67. Bruschetta, M. et al. Assessing the contribution of active somatosensory stimulation to self-acceleration perception in dynamic driving simulators. *PLOS ONE*. **16**, e0259015 (2021).
68. Walsh, E. G. The Perception of Rhythmically Repeated Linear Motion in the Vertical Plane. *Q. J. Exp. Physiol. Cogn. Med. Sci.* **49**, 58–65 (1964).
69. Hubel, D. H. & Wiesel, T. N. Receptive fields of single neurones in the cat's striate cortex. *J. Physiol.* **148**, 574–591 (1959).
70. Sabbah, S. et al. A retinal code for motion along the gravitational and body axes. *Nature* **546**, 492–497 (2017).
71. Gu, Y., Watkins, P. V., Angelaki, D. E. & DeAngelis, G. C. Visual and nonvisual contributions to three-dimensional heading selectivity in the medial superior temporal area. *J. Neurosci.* **26**, 73–85 (2006).
72. Jamali, M., Sadeghi, S. G. & Cullen, K. E. Response of vestibular nerve afferents innervating utricle and saccule during passive and active translations. *J. Neurophysiol.* **101**, 141–149 (2009).
73. Churchland, A. K., Gardner, J. L., Chou, I., Priebe, N. J. & Lisberger, S. G. Directional anisotropies reveal a functional segregation of visual motion processing for perception and action. *Neuron* **37**, 1001–1011 (2003).
74. Indovina, I. et al. Simulated self-motion in a visual gravity field: Sensitivity to vertical and horizontal heading in the human brain. *NeuroImage* **71**, 114–124 (2013).
75. Indovina, I. et al. Representation of visual gravitational motion in the human vestibular cortex. *Science* **308**, 416–419 (2005).
76. Indovina, I., Maffei, V. & Lacquaniti, F. Anticipating the effects of visual gravity during simulated self-motion: Estimates of time-to-passage along vertical and horizontal paths. *Exp. Brain Res.* **229**, 579–586 (2013).
77. Indovina, I. et al. Structural connectome and connectivity lateralization of the multimodal vestibular cortical network. *NeuroImage* **222**, 117247 (2020).
78. Bosco, G., Carrozzo, M. & Lacquaniti, F. Contributions of the human temporoparietal junction and MT/V5+ to the timing of interception revealed by transcranial magnetic stimulation. *J. Neurosci.* **28**, 12071–12084 (2008).
79. Delle Monache, S., Lacquaniti, F. & Bosco, G. Differential contributions to the interception of occluded ballistic trajectories by the temporoparietal junction, area hMT/V5+, and the intraparietal cortex. *J. Neurophysiol.* **118**, 1809–1823 (2017).
80. Frank, S. M. & Greenlee, M. W. The parieto-insular vestibular cortex in humans: More than a single area? *J. Neurophysiol.* **120**, 1438–1450 (2018).
81. Ibitoye, R. T. et al. The human vestibular cortex: functional anatomy of OP2, its connectivity and the effect of vestibular disease. *Cereb. Cortex*. **33**, 567–582 (2022).
82. Alexandre, N. et al. The role of cortical areas hMT/V5+ and TPJ on the magnitude of representational momentum and representational gravity: a transcranial magnetic stimulation study. *Exp. Brain Res.* **237**(12), 3375–3390. <https://doi.org/10.1007/s00221-019-05683-z> (2019).
83. Maffei, V. et al. Processing of visual gravitational motion in the peri-sylvian cortex: Evidence from brain-damaged patients. *Cortex* **78**, 55–69 (2016).

Acknowledgements

This study was supported by grants from the Italian Space Agency (I/006/06/0), INAIL (BRIC 2022 LABORIOUS), Italian University Ministry (PRIN 2020EM9A8X, 2020RB4N9, 2022T9YJXT, 2022YXLNR7, #NEXTGENERATIONEU NGEU National Recovery and Resilience Plan NRRP, project MNESYS PE0000006 – A Multiscale integrated approach to the study of the nervous system in health and disease DN. 1553 11.10.2022) and Space It Up project funded by the Italian Space Agency and the Ministry of University and Research - Contract No. 2024-5-E.0 - CUP No. I53D24000060005. The authors thank Greta Dimasi for help with the setup.

Author contributions

SDM and BLS conceived the study, researched and analyzed data, and wrote the first draft of the manuscript. AFA contributed to the characterization of the setup. FL and MZ contributed to discussions and successive drafts of the manuscript. MZ contributed resources. All authors approved the final version of the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-94512-1>.

Correspondence and requests for materials should be addressed to S.D.M. or B.L.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025