Motor Signals Mediate Stationarity Perception

Savannah Halow[1,2,*], **James Liu**[2], **Eelke Folmer**[2], **and Paul MacNeilage**[1]

[1]University of Nevada, Reno, Psychology, Reno, 89557, USA
[2]University of Nevada, Reno, Department of Computer Science and Engineering, Reno, 89557, USA
[*]savvyhalow@nevada.unr.edu

ABSTRACT

Head movement relative to the stationary environment gives rise to congruent vestibular and visual optic flow signals. The resulting perception of a stationary visual environment, referred to herein as stationarity perception, depends on mechanisms that compare visual and vestibular signals to evaluate their congruence. Here we investigate the functioning of these mechanisms and their dependence on fixation behavior as well as on the active versus passive nature of the head movement. Stationarity perception was measured by modifying the gain on visual motion relative to head movement on individual trials and asking subjects to report whether the gain was too low or too high. Low and high gains result in perception of the environment moving with or against head movement, respectively. Fitting a psychometric function to the resulting data yields two key parameters to characterize performance; the mean and standard deviation (SD) of the cumulative Gaussian fit. The mean is a measure of accuracy and indicates the single visual gain value that is perceived to match head movement. The SD is a measure of precision and indicates the range of gains that are compatible with perception of a stationary visual environment. Experiments were conducted using a head-mounted display capable of rendering visual scene motion contingent on head motion, with fixation behavior monitored by an embedded eye tracker. The experimental design included combinations of active or passive head movement together with head-fixed or scene-fixed fixation. During active conditions, subjects rotated their heads in yaw ∼15 deg/s over ∼1 sec. Each subject's movements were recorded and played back via rotating chair during the passive condition. During head-fixed and scene-fixed fixation the target moved with the head or scene, respectively. Both precision (quantified by SD) and accuracy (quantified by mean) were better during active than passive head movement, likely due to increased precision on the head movement estimate arising from motor prediction and neck proprioception. Performance was also better during scene-fixed than head-fixed fixation, perhaps due to decreased velocity of retinal image motion and increased precision on the retinal image motion estimate.

Introduction

Perception of a stationary visual environment, i.e. stationarity perception, is often taken for granted but it depends on neural mechanisms that compare visual and non-visual (e.g. vestibular) self-motion estimates and evaluate their congruence. Incongruence can result in the clinical symptom known as vertigo, which is the false perception that either the observer, the environment, or both, are in motion[1,2]. This incongruence can be caused, for example, by vestibular deficits that alter vestibular self-motion signals[2,3], or by immersive visual environments[?,?,3] that drive visual self-motion estimates that do not match the physical motion of the observer. In these cases, stationarity perception can fail, leading to vertigo and associated symptoms including imbalance and motion sickness[4]. It is therefore important to understand how stationarity perception is mediated by motor and sensory signals.

Here we investigate how stationarity perception is mediated by oculomotor and neck-motor signals in particular. Prior psychophysical work has found differences in perception during active and passive head movements.[5–8] However, no prior study has directly compared stationarity perception under conditions in which the movement was generated actively versus passively. This comparison is important because we know that vestibular signals are processed differently at the earliest stages of vestibular processing, depending on whether they were actively or passively generated[?]. Additionally, prior research has established that visual motion perception is mediated by the contributions of multiple extra-retinal cues including neck proprioception, efference copy, and vestibular canal stimulation.[7,9] Because neck-motor signals have been shown to impact both visual and vestibular function, we predicted they would also impact statinoarity perception, which depends on the comparison of visual and vestibular signals.

Of equal importance is the investigation of the role of oculomotor signals; because vestibular signals are sensed in head coordinates and retinal image motion is sensed in eye coordinates, knowledge of eye-in-head motion is necessary to allow visual and vestibular comparison in a common reference frame. During natural behaviors, humans often fixate features of the stationary environment[10] and this is facilitated by the vestibulo-ocular reflex (VOR)[11]. In this case, eye movement is meant to cancel visual self-motion signals at the fovea and visual-vestibular congruence can, in principle, be evaluated by comparing the vestibular and oculomotor signals. The question the nervous system must evaluate is whether the object being fixated is stationary. However, if a moving observer pursues an independently moving object, i.e., one that is not fixed to the stationary environment, evaluation of congruence requires comparing the vestibular signal to the background visual motion, or equivalently with the oculomotor pursuit signal that is superimposed on the VOR. In this case, the question the nervous system must evaluate is whether the background of the object being fixated is stationary. Thus, we predicted that oculomotor behavior would impact stationary perception in the current study, consistent with prior reports[12].

Similar studies have been conducted previously, but these studies have not explicitly addressed the role of neck-motor signals and their interaction with oculomotor signals. Beginning in the 1960's Hans Wallach and colleagues conducted several experiments to quantify the accuracy and precision of stationarity perception[13–15]. Yaw head movement was mechanically tracked with a helmet and this movement was used to drive rotation of the visual scene with a variable gain ratio via either a movable mirror or shadow cage. Participants adjusted the gain to achieve a gain perceived as stationary, allowing measurement of both the single gain value judged to be stationary as well as the range of gains across which the environment was perceived as stationary, which was referred to as the "range of immobility". More recently, several studies have conducted similar research using virtual reality rather than analog methods[12,16–19] Gain values for stationarity perception, as well as the range of gains perceived as stationary, vary considerably across studies due to differences in methodology (e.g. head-tracking method, rendering latency, etc.) as well as differences in the presented motions (e.g. linear versus angular, motion profile, etc.). It is worth noting that the studies described above have referred to the problem faced by the nervous system as either perception of a stable environment during self-motion[13–16] or detection of visual-vestibular conflict[12]. Here, we introduce the term stationarity perception as an alternative because we believe it is crucial for the organism to infer whether the environment is stationary or not. This inference determines whether visual motion cues provide a reliable indication of self-motion, and therefore whether they should be used to drive and calibrate essential perceptual, postural, and oculomotor responses.

Results

To investigate stationarity perception, we conducted experiments in which subjects experienced head turns while wearing a head-mounted display. We manipulated the gain of visual scene motion relative to head motion on each trial and asked participants to judge whether the gain was too fast or too slow or equivalently whether the scene appeared to drift with or against their head motion in the world coordinate system (Figure 1). Gain was modified across trials according to an adaptive procedure and a cumulative Gaussian psychometric function[20] was fit to all data from a given condition (Figure 3). The mean and standard deviation of the cumulative Gaussian fit are the dependent measures of interest because they capture the accuracy and precision of visual-vestibular conflict judgments, respectively. The mean is the gain value that elicits approximately fifty percent "with" and "against" responses, and it is therefore consistent with perception of a stationary environment. We refer to this value as the point of subjective equality (PSE)[21]. The standard deviation of the cumulative Gaussian fit indicates the change in gain relative to the PSE that is "just-noticeable." We refer to this value as the just-noticeable difference (JND)[21], and it can be understood as the increase/decrease in gain that leads to 84%/16%[21] of responses being against/with, respectively.

In order to investigate how stationarity perception depends on neck motor and oculomotor signals, we measured and compared performance across both active and passive head movements with both head-fixed and scene-fixed fixation (Figure 2). During Active conditions, participants made trained yaw head rotations of approximately 15 degrees over approximately 1 sec. These trajectories were recorded and played back during the Passive condition for each subject using a rotating chair such that head movement through space and thus vestibular stimulation was as similar as possible between these conditions while neck efference and proprioception differed. During the Scene-fixed fixation conditions (Figure 2, top), the fixation point remained fixed to the scene such that the participant had to counter-rotate their eyes in order to maintain fixation. During the Head-fixed fixation condition (Figure 2, bottom), the fixation point remained fixed relative to the head such that no eye-in-head movement was elicited.

Results demonstrate that both the accuracy (Figure 4, left) and precision (Figure 4, right) of stationarity judgments depended strongly on neck motor and oculomotor signals. The highest accuracy was observed during active head movement with scene-fixed fixation. In this condition, which arguably aligns best with everyday head and eye movement behavior, the PSE was very close to zero, meaning that visual speed that objectively matched head speed (visual/head speed ratio of 1) was most likely to be perceived as stationary. In all other conditions, the PSE was lower. The PSE was significantly affected by head movement type (F=13.38, p=0.002), and fixation type (F=6.55,p=0.02), but there was no interaction (F=0.08, p=0.78).

Interestingly, the best precision was also observed during active head movement with scene-fixed fixation. The JND in this condition was 0.255, meaning that on average a 25 percent increase or decrease in the visual speed relative to head speed was needed in order for it to be reliably recognized as too fast or too slow. JNDs were higher in the other conditions meaning that larger changes in visual speed were needed in order to detect conflict. The effect of head movement type on JND was significant (F=6.25, p=0.02), as was the effect of fixation type (F=7.15, p=0.02), with better sensitivity to conflict during scene-fixed than head-fixed fixation. The interaction between head movement and fixation type was not significant (F=2.27, p=0.15).

In summary, the best accuracy and precision was observed during active head movement with scene-fixed fixation, and in general, for conditions in which performance was less accurate, it was also less precise. To investigate the relationship between precision and accuracy, we examined the correlation between PSE and JND across all subjects and conditions but we found that this correlation was close to, but not significant (rho=-0.20, p=0.08).

Discussion

Stationarity perception depends on mechanisms that evaluate the congruence of visual and vestibular sensory signals. Here we investigated these mechanisms by modifying the speed of the visual stimulus proportional to head speed (visual gain) in a virtual reality headset and asking human subjects to judge whether the speed was too slow or too fast. This allowed quantifying the range of visual gains most compatible with perception of a stationary environment (JND, precision), as well as the single visual gain value that was judged most congruent (PSE, accuracy). We found that stationarity perception was mediated by both neck-motor and oculomotor signals. Specifically, during active head-on-body movements with scene-fixed fixation, arguably the most natural and common situation, veridical visual gains (i.e. gain ~ 1, log(gain) ~ 0) led to perception of stationarity. When head movements were delivered passively and when subjects fixated head-fixed fixation points, visual gains perceived as stationary were significantly reduced; slower, non-veridical visual speeds were needed in order to perceive stationarity.

Stationarity perception during active and passive head movements - the role of neck motor signals

No prior studies have compared stationarity perception during active relative to passive head movement. Most existing studies of stationarity perception during yaw head rotation have used active head movements only, and typically report gains slightly, but generally not significantly, greater than 1 (or equivalently, log(gain)>0, Figure 4). This is true for experiments conducted with both real[13–15] and virtual (Harris studies) environments. Gains greater than 1 have also been reported as optimal for stationarity perception during passive movement[12]. In contrast, we find that gains near 1 are optimal during active head movements, but that gains are significantly reduced during passive movement (Figure 4). In general, deviations of the current gain measurements relative to prior studies can likely be attributed to methodological differences in the visual stimulus, display technology, movement profile, and/or tracking method.

Nevertheless, the novel finding that optimal visual gain is reduced during passive head movement, and that it is less than 1, is somewhat surprising and apparently contrary to everyday experience. When experiencing passive rotation in the real world with a veridical visual gain of 1, we do not typically perceive that the visual world is rotating against our own motion. This is likely because the deviation of the optimal from veridical gain is small and therefore not noticeable, on average, due to noise or uncertainty on both visual and non-visual self-motion estimates. Nevertheless, it is important to consider the underlying cause; why are non-veridical gains less than 1 perceived as matching during passive rotation?

We hypothesize that this finding is due to an underestimation of head motion based on non-visual (e.g. vestibular) signals during passive compared to active head movement. An alternative explanation is that passive movement leads to an increase in the visual speed estimate rather than a decrease in the non-visual head speed estimate. However, prior research (reviewed below) suggests that the former explanation is more likely.

Differences in perception of active versus passive head movements have been documented previously using spatial updating tasks in which subjects are rotated in darkness then asked to indicate the (updated) position of a remembered earth-fixed visual or auditory target that was presented before the rotation. Results of these studies suggest that perceived head displacement (which is inferred based on updating performance) is typically underestimated during passive relative to active head movements[6, 22].

However, it seems equally possible that head speed (rather than only displacement) may be underestimated during passive head rotation. The current study probes speed estimation more directly than updating studies because the task implies ongoing evaluation of the match between estimated visual and head speed, not necessarily ending position or displacement. Furthermore, the visual environment in the present study did not contain obvious and salient landmarks that could directly support estimation of visual displacement. These results extend prior perceptual research comparing the impact of active and passive head movements on spatial updating, and demonstrate that underestimation of passive head movement has consequences for stationarity perception as well.

But what are the neural underpinnings of these effects? The peripheral vestibular system is stimulated by angular acceleration, but afferent responses are generally proportional to angular velocity, and they are equivalent during both active and passive head movements. However, during active head movements, neck motor signals are generated to rotate the head on the body, and efference copies of these motor signals innervate the central vestibular system very early on at the level of the vestibular nuclei[23]. This supports differential vestibular processing of active versus passive head movements; certain types of neurons show responses that are greatly attenuated during active head movements. This differential processing may, in part, underlie observed behavioral differences, such as the reduction in the gain of the VOR during passive relative to active head movements[?]. Here we extend these observations to the domain of perception; the reduction in the gain of visual speed perceived as matching during passive relative to active head movement is likely to be a perceptual manifestation of the neural processes underlying differences in vestibular-driven neurophysiological and oculomotor responses.

In addition to differences in gain, we also observed differences in variability of the stationarity judgements (JND) between active and passive head movements; variability was generally higher during passive than during active head movements. Again, this observation agrees with findings showing that the variability on spatial updating performance is greater during passive than active head movements[6,22], an effect that is typically attributed to the loss of information provided by the efference copy signal during passive movement. Maximum-likelihood estimation models predict that noise on the combined non-visual estimate of head speed will increase when this additional source of information is not available[22].

Stationarity perception during scene-fixed and head-fixed fixation - the role of oculomotor signals

During natural active head movements, humans most often fixate scene-fixed features of the environment[10]. Under these conditions, we find that veridical visual gains close to 1 are most compatible with stationarity perception. When observers instead maintain fixation on head-fixed targets, visual gains for stationarity perception are significantly reduced (Figure 4). This novel finding could be due to underestimation of head speed, overestimation of visual speed, or both. In fact, prior research has documented effects of oculomotor signals on both perceived head motion[24,25] and perceived visual motion[25,26].

Under the assumption of scene-fixed fixation, oculomotor signals can provide direct information about self-motion; greater eye movement implies greater self-motion. In one previous study, subjects moved passively in darkness tended to perceive greater self-motion on trials in which they moved their eyes more, suggesting that amplitude of (assumed scene-fixed) fixation movements influenced perceived self-motion[24]. If such an effect is present in the current study, it would lead to reduced self-motion estimates in the head-fixed fixation condition, and thus reduced visual gain perceived as matching, consistent with our results.

Effects of eye movements on perceived visual speed have also been documented previously, but most often with stationary observers. For example, objects are perceived to move more slowly when pursued with the eyes than when the eyes remain fixed and the object moves past, an effect known as the Aubert-Fleischl (AF) phenomenon[26]. An analogous effect (the vestibular AF) has been reported in moving subjects asked to maintain fixation on a head-fixed point. Head-fixed fixation during self-motion depends on a pursuit signal to cancel the vestibulo-ocular reflex. Perceived visual speed during the classical and vestibular AF are similar and correlated within a given observer, suggesting that they may both be mediated by oculomotor pursuit signals[25].

Similarly, pursuit mechanisms must be acting during head-fixed fixation in the current study. However, the stationarity task in the current study differs from the task in AF studies because it requires the subject to estimate the speed of the background motion (rather than the foreground object) and its congruence with estimated head speed. The significant reduction in visual gain suggests that pursuit behavior during head-fixed fixation may have caused subjects to underestimate the speed of the visual background motion relative to the fixation target. This is roughly equivalent to the underestimation of the speed of the pursued object relative to the environment during the classic AF, except that those measurements are typically taken in stationary observers without background motion. In summary, the reduction in visual gain during head-fixed fixation could be due to the impact of oculomotor signals on estimation of self-motion, estimation of visual motion, or both. Results of the current study do not allow us to differentiate between these alternative explanations.

Regarding the neural underpinnings of these effects, they likely depend on the specific oculomotor circuits that are engaged. For example, eye movement during scene-fixed fixation is likely governed by the visually-enhanced vestibulo-ocular reflex which serves to maintain high-acuity vision of the environment during self-motion[11]. The visually-enhanced VOR is present in most vertebrates and is controlled by circuits that include the vestibular and oculomotor nuclei[11]. Even though the VOR is largely controlled by brainstem circuits, there is evidence that VOR responses can be mediated by cortical mechanisms as well, under certain circumstances[27]. During head-fixed fixation, on the other hand, eye movements are governed by the pursuit system, which serves to maintain high-acuity vision of the pursued object. Pursuit behavior is only present in animals that have a fovea, and it is controlled predominantly by cortical neural circuits[?]. In summary, scene-fixed and head-fixed fixation are distinct behavioral tasks, one governed by self-motion and the other governed by object motion, and they entail distinct oculomotor behaviors and circuits. The novel results presented here demonstrate that differences in motor behavior and underlying neurophysiology have a significant impact on stationarity perception via differential effects on estimation of self-motion, visual motion, or both.

These differences also impact the variability on stationarity judgements (JNDs), with greater variability during head-fixed than during scene-fixed fixation. This increased variability has been reported previously[12,19]. The increase may be due to differences in the variability on the visual speed estimate. Background retinal image motion is minimized during scene-fixed fixation, but increased inline with head speed during head-fixed fixation. Noise proportional to retinal image speed could explain increased noise during head-fixed fixation, however visual-only measurements conducted previously suggest that noise on the visual estimate is comparable between head-fixed and scene-fixed conditions[12]. An alternative explanation (discussed in greater detail in the next section) is that the efficiency of crossmodal comparison mechanisms may be impacted by motor behavior.

Stationarity perception, crossmodal comparison, and causal inference

Stationarity perception is well-suited to be modeled in a causal inference framework. Visual and vestibular signals will be congruent when they are both caused exclusively by movement relative to the stationary environment. So perception of a stationary environment may be modeled as dependent on the posterior probability that these signals have a common cause. One previous study has proposed a causal inference model that evaluates congruence of visual and vestibular self-motion signals, but the model does not address stationarity perception. Instead, the model aims to predict explicit judgments about the congruency of the heading stimulus presented via the visual and vestibular modalities, i.e. are the stimuli congruent or not[?]? In the current study, we instead asked subjects to explicitly judge the direction of conflict between visual and vestibular stimuli, i.e. was the visual stimulus motion too fast or too slow relative to the non-visual motion? This explicit crossmodal discrimination is distinct from explicit judgments about congruency. Nevertheless, it is possible that these two tasks elicit different read-outs from the same underlying probabilistic representation.

To explore this topic in more depth, we briefly review previously proposed models of crossmodal discrimination. The simplest is the standard signal detection model in which variability on the comparison judgment (the JND on stationarity perception) is equal to the sum of the variabilities on the signals being compared (i.e. visual and non-visual estimates); this assumes that the signals are conditionally independent[12]. However, cross modal discrimination can also be modeled in a probabilistic framework that depends not only on signal variabilities but also on the prior probability that signals match. This probability is referred to as the coupling prior[28,29], and it is computationally equivalent to the prior probability of a common cause in causal inference models[?,30]. In the present context, it may also be referred to as a stationarity prior.

In a previous study, we suggest that crossmodal discrimination responses can be predicted from such a model based on the posterior probability of a common cause represented in a two-dimensional stimulus space. In this model, visual and vestibular marginal distributions of the posterior are compared with one another (see Fig. 4 of[12]). If this prior is weak, these marginal distributions accurately reflect the information in the stimulus, and performance is equivalent to the standard signal detection model. If the prior is strong, this implies a strong prior belief that the environment is stationary, and that signals match. The marginal distributions will be biased so as to more closely resemble one another, leading to impaired crossmodal discrimination and increased variability on stationarity judgements (JNDs).

These alternatives are laid out in a previous study[12], in which we observed better crossmodal comparison during scene-fixed compared to head-fixed conditions, similar to the present results. We suggest that this difference may be due to changes in the strength of the coupling (or stationarity) prior contingent on oculomotor behavior with a weaker prior during scene-fixed fixation and a stronger prior during head-fixed fixation. Oculomotor behavior is particularly likely to impact crossmodal comparison because different behaviors serve different goals and engage different neural circuits. Scene-fixed fixation serves to stabilize the visual environment and engages VOR circuits while head-fixed fixation serves to stabilize the fixation point, which mimics an object moving in the world and therefore engages pursuit circuits. We speculate that better crossmodal discrimination (i.e. reduced JNDs) during scene-fixed fixation may reflect the engagement of conflict detection circuits that serve to maintain calibration of the vestibulo-ocular reflex. These circuits may be suppressed during head-fixed fixation due to the engagement of oculomotor pursuit circuits needed to pursue the head-fixed target. During pursuit, visual and non-visual signals may be fused due to a strong stationarity prior (i.e. forced fusion;[31]) leading to compromised crossmodal comparison and conflict detection[12], and therefore an increased range of gains perceived as stationary.

Conclusion

The nature of head and eye movements impacts the precision and accuracy of stationarity perception. This is due to differences in how the nervous system encodes, processes and compares sensory and motor signals contingent on the nature of the head and eye movement. These findings were made possible through the use of a psychophyscial task that allows assessment of stationarity perception. Further research is needed to identify the neural correlates of the psychophyscial performance measured here in order to better understand how the nervous system evaluates environmental stationarity.

Methods

Participants

20 healthy subjects (7 female) with normal or corrected-to-normal vision, varying in age from 21 to 42 (average: 27), participated in this study. Participants had no known muscular, motor, or vestibular impairments or disorders. All but two participants were naïve to the aims of the study, and no subjects were excluded from the study. Participants provided informed consent, and procedures were approved by the Institutional Review Board of the University of Nevada, Reno.

Equipment

We use an HTC Vive Pro HMD to present stimuli. It has a diagonal FOV of 110 degrees, refresh rate of 90Hz, and combined resolution of 2880x1600 pixels. It allows the user to adjust the interpupillary distance (IPD) and focal distance of the HMD to their comfort. The HMD is powered by a Dell XPS 8930 computer with Intel® Core ™ i7-8700 CPU, 16 GB Ram, GeForce GTX 1070. Our stimuli are programmed in Unity v2017.20f3 and integrated with the HMD using the Steam VR plug-in. We use an OptiTrack motion capture system with four Prime 13 cameras operating at 240 Hz to track head movement via an IR reflective rigid body attached to the HMD. Fixation is confirmed using an embedded Pupil Labs HMD add-on eye tracker and integrated with our application via the Pupil Labs Unity plug-in. Passive motion is generated via a custom-built rotating chair[32] with four-point racing harness. In the passive condition, head movement trajectories recorded from the most recent active movement condition were "played back" resulting in passive stimuli that replicate the active stimuli in terms of both the individual trajectories as well as the order in which those trajectories were delivered. Every effort was made to replicate trajectories exactly, but displacement was sometimes marginally reduced in the passive relative to the active trials. Note that this offset, if present, did not impact the key manipulation of this study, which was the gain of the visual stimulus relative to the physical head rotation on each trial. Responses were given with a Logitech Keyboard using the arrow keys. Total system latency was determined to be 27 milliseconds using a latency test described by Neirhorster, Li, and Lappe 2017[33].

Procedure

Participants are trained to make active head movements of approximately 15 degrees over 1 second during 151 practice trials. During this time, they receive verbal feedback from the researcher on their performance in identifying "with" or "against" trials to confirm they understand the task. To start a trial, participants align a 0.5-by-0.5 meter reticle at the center of their FOV in the headset to an identical reticle in the center of the visual scene (1, Top). This ensures that participants start each trial with their head in the center of the scene and fixated on the fixation point. The reticles are at a distance of 10 meters from the participant in virtual space.

Following alignment, a green dot appears to the left or right of the participant and disappears after 0.3 seconds, indicating the direction of rotation (1, Top). While maintaining fixation on the reticle, participants perform the trained head movement in the direction which the cue had appeared and the trial ends when head movement stops (e.g. velocity of the head dropped below 0.1 m/s). The procedure differs slightly in the passive condition in that participants start each trial by pressing the space bar on the keyboard instead of aligning the reticles and the motorized chair rotates the participant. They are still prompted via the cue to the direction in which they would be turning. Participants' heads were secured to the chair via a buckled strap to ensure no head movement occurred during the passive condition. Rotation (left or right) and direction of perturbation (with or against) is randomly determined each trial.

Eye movement differs between the head-fixed and scene-fixed conditions; During the head-fixed conditions, the fixation reticle remains central to the participant's FOV. Therefore, there was minimal eye movement relative to the head, minimizing eye movement but maximizing optic flow. During the scene-fixed condition, the fixation reticle remains central relative to this visual scene. The subject's eyes move opposite of head motion, maximizing eye movement but minimizing optic flow. Fixation is confirmed using the Pupil Labs embedded eye tracker. Calibration of the eye tracker is performed before each set of 151 trials using the 7 point calibration procedure provided with the Pupil Labs plug-in software.

Gains are altered for each trial on two interleaved staircases[12,20,34], with new staircases each condition. Gains start at 0.7 and 1.3 based on prior research which found gains between 0.8 and 1.4 were perceived as stable[16,35]. Gains are adjusted by 0.4 increment adjustments until the second reversal, at which adjustments become 0.2 for each change.

Data collection was spaced out into 4 blocks over 4 days for each participant. Participants were tested in 2 condition types, 151 trials each, for a total of 302 trials each block. Participants were given a 30 second break every 50 trials during testing, and an additional break between conditions. This amounts to 302 trials for each condition pair. A minimum 24 hour break was given between testing blocks. The pairing of conditions was randomized to control for practice effects. Simulator Sickness Questionnaires were collected before and after each set of 151 trials. The experiment was conducted in a dark room to avoid any additional cues to motion. Due to equipment failure, 5 participants had to repeat one block, and 10 trials were not collected from one participant.

Statistical analyses were done using MATLAB R2017b. Psychometric functions were made using the Palamedes toolbox[21]. We performed a two-way ANOVA to compare fixation (Head-fixed versus Scene-fixed) and Head Movement conditions (Active versus Passive).

References

1. Baloh, R. W. Vertigo. *The Lancet* **352**, 1841–1846 (1998).
2. Bronstein, A. M. Vision and vertigo: Some visual aspects of vestibular disorders. *J. Neurol.* **251**, 381–387, DOI: 10.1007/s00415-004-0410-7 (2004).

3. Bertolini, G. & Straumann, D. Moving in a moving world: A review on vestibular motion sickness. *Front. Neurol.* **7**, 1–11, DOI: 10.3389/fneur.2016.00014 (2016).
4. Reason, J. & Brand, J. J. *Motion Sickness* (Academic Press, 1975).
5. Telford, L., Howard, I. P. & Ohmi, M. Heading judgments during active and passive self-motion. *Exp. Brain Res.* **104**, 502–510, DOI: 10.1007/BF00231984 (1995).
6. Mergner, T., Huber, W. & Becker, W. Vestibular-neck interaction and transformation of sensory coordinates. *J. Vestib. Res. Equilib. Orientation* **7**, 347–367, DOI: 10.1016/s0957-4271(96)00176-0 (1997).
7. Crowell, J. A., Banks, M. S., Shenoy, K. V. & Andersen, R. A. Visual self-motion perception during head turns. *Nat. Neurosci.* **1**, 732–737, DOI: 10.1038/3732 (1998).
8. Barnett-Cowan, M. & Harris, L. R. Temporal processing of active and passive head movement. *Exp. Brain Res.* **214**, 27–35, DOI: 10.1007/s00221-011-2802-0 (2011).
9. Biguer, B., Donaldson, I., Hein, A. & Jeannrod, M. Neck Muscle Vibration Modifies the Representation of Visual Motion and Direction in Man. *Brain* **111** (1988).
10. Einhäuser, W. *et al.* Human eye-head co-ordination in natural exploration. *Network: Comput. Neural Syst.* **18**, 267–297, DOI: 10.1080/09548980701671094 (2007).
11. Leigh, R. J. & Zee, D. S. *The Neurology of Eye Movements* (Oxford University Press, 2015).
12. Garzorz, I. & MacNeilage, P. R. Visual-Vestibular Conflict Detection Depends on Fixation. *Curr. Biol.* **27**, 2856–2861, DOI: 10.1016/j.cub.2017.08.011 (2017).
13. Wallach, H. & Kravitz, J. H. The measurement of the constancy of visual direction and of its adaptation. *Psychon. Sci.* **2**, 217–218, DOI: 10.3758/BF03343414 (1965).
14. Wallach, H. Learned stimulation in space and motion perception. *Am. Psychol.* **40**, 399–404, DOI: 10.1037/0003-066X.40.4.399 (1985).
15. Wallach, H. Perceiving A Stable Environment When One Moves. *Annu. Rev. Psychol.* **38**, 1–27, DOI: 10.1146/annurev.psych.38.1.1 (1987).
16. Jaekl, P. M., Jenkin, M. R. & Harris, L. R. Perceiving a stable world during active rotational and translational head movements. *Exp. Brain Res.* **163**, 388–399, DOI: 10.1007/s00221-004-2191-8 (2005).
17. Macneilage, P. R. Psychophysical Investigations o f Visual-vestibular Interactions in Human Spatial Orientation. Tech. Rep. (2007).
18. Correia Grácio, B. J., Bos, J. E., Van Paassen, M. M. & Mulder, M. Perceptual scaling of visual and inertial cues: Effects of field of view, image size, depth cues, and degree of freedom. *Exp. Brain Res.* **232**, 637–646, DOI: 10.1007/s00221-013-3772-1 (2013).
19. Moroz, M., Garzorz, I., Folmer, E. & MacNeilage, P. Sensitivity to visual speed modulation in head-mounted displays depends on fixation. *Displays* **58**, 12–19, DOI: 10.1016/j.displa.2018.09.001 (2019).
20. Leek, M. R. Origins of Adaptive Psychophysical Procedures. *Percept. & Psychophys.* **63**, 1279–1292 (2001).
21. Kingdom, F. & Prins, N. *Psychophysics: A Practical Introduction* (2010).
22. Genzel, D., Firzlaff, U., Wiegrebe, L. & MacNeilage, P. R. Dependence of auditory spatial updating on vestibular, proprioceptive, and efference copy signals. *J. Neurophysiol.* **116**, 765–775, DOI: 10.1152/jn.00052.2016 (2016).
23. Cullen, K. E. Sensory signals during active versus passive movement. *Curr. Opin. Neurobiol.* **14**, 698–706, DOI: 10.1016/j.conb.2004.10.002 (2004).
24. Clemens, I. A. H., Selen, L. P. J., Pomante, A., MacNeilage, P. R. & Medendorp, W. P. Eye Movements in Darkness Modulate Self-Motion Perception. *eneuro* **4**, 0211–16, DOI: 10.1523/ENEURO.0211-16.2016 (2017).
25. Garzorz, I., Freeman, T., Ernst, M. & MacNeilage, P. R. Insufficient compensation for self-motion during perception of object speed: The vestibular Aubert-Fleischl phenomenon. *J. Vis.* **18**, 9, DOI: 10.1167/18.13.9 (2018).
26. Freeman, T. & Banks, M. Perceived head-centric speed is affected by both extra-retinal and retinal errors. *Vis. Res.* **38**, 941–945, DOI: 10.1016/S0042-6989(97)00395-7 (1998).
27. Arshad, Q., Nigmatullina, Y. & Bronstein, A. M. Handedness-related cortical modulation of the vestibular-ocular reflex. *J. Neurosci.* **33**, 3221–3227, DOI: 10.1523/JNEUROSCI.2054-12.2013 (2013).

28. Bresciani, J. P., Dammeier, F. & Ernst, M. O. Vision and touch are automatically integrated for the perception of sequences of events. *J. Vis.* **6**, 554–564, DOI: 10.1167/6.5.2 (2006).

29. Ernst, M. O. & Luca, M. D. CHAPTER 12 Multisensory Perception: From Integration to Remapping. Tech. Rep. (2011).

30. Körding, K. P. *et al.* Causal inference in multisensory perception. *PLoS ONE* **2**, DOI: 10.1371/journal.pone.0000943 (2007).

31. Prsa, M., Gale, S. & Blanke, O. Self-motion leads to mandatory cue fusion across sensory modalities. *J. Neurophysiol.* **108**, 2282–2291, DOI: 10.1152/jn.00439.2012 (2012).

32. Mallery, R. M., Olomu, O. U., Uchanski, R. M., Militchin, V. A. & Hullar, T. E. Human discrimination of rotational velocities. *Exp. Brain Res.* **204**, 11–20, DOI: 10.1007/s00221-010-2288-1 (2010).

33. Niehorster, D. C., Li, L. & Lappe, M. The accuracy and precision of position and orientation tracking in the HTC vive virtual reality system for scientific research. *i-Perception* **8**, 1–23, DOI: 10.1177/2041669517708205 (2017).

34. Merfeld, D. M. Theory and Practical Considerations. *Exp. Brain Res.* **210**, 389–405, DOI: 10.1007/s00221-011-2557-7.Signal (2011).

35. Jürgens, R. & Becker, W. Perception of angular displacement without landmarks: Evidence for Bayesian fusion of vestibular, optokinetic, podokinesthetic, and cognitive information. *Exp. Brain Res.* **174**, 528–543, DOI: 10.1007/s00221-006-0486-7 (2006).

Acknowledgements

This research was supported by NIH P20GM103650.

Author contributions statement

SH and JL programmed and conducted the experiment, in addition to analysing the results. PM conceived the experiment and aided in data analysis. All authors reviewed the manuscript.

Additional information

Figures

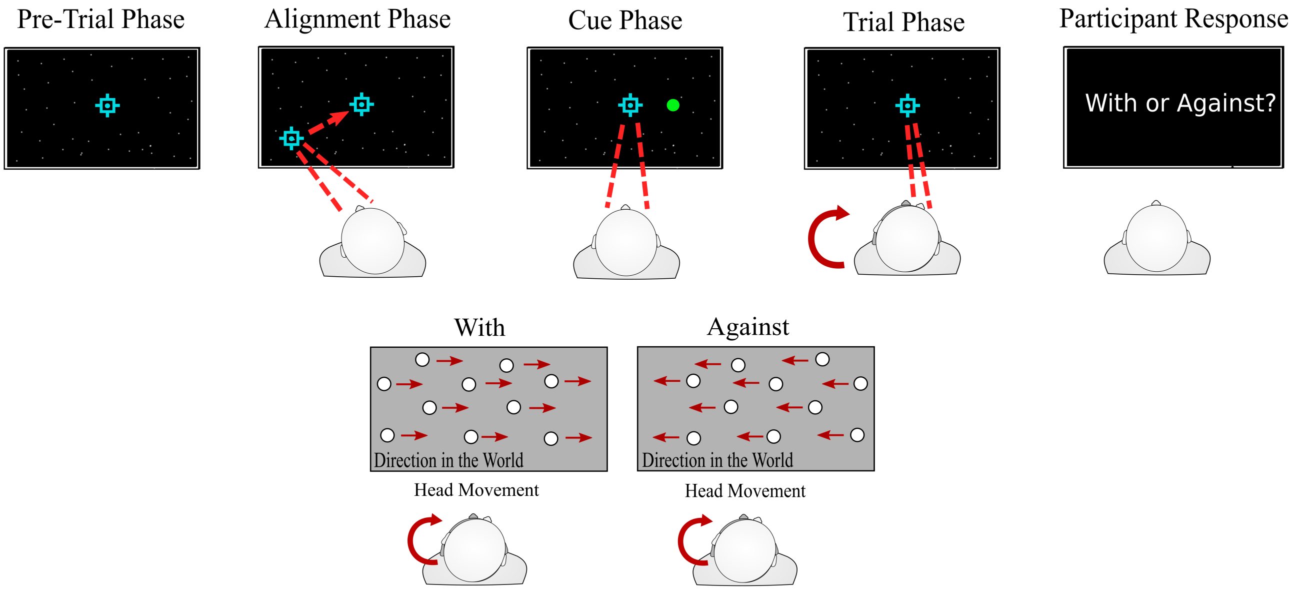

Figure 1. Experimental trial. The time course of a single trial is illustrated (top row). In the alignment phase, the participant moves their head to align a head-fixed target with a central scene-fixed target to ensure that the participant begins each trail with the head in the desired position. When alignment is detected, a cue dot appears briefly 15 deg in the periphery to the left or right side. In the trial phase, the participant must maintain fixation on either the head-fixed or scene-fixed target (depending on the condition) and rotate their head to point to where the cue dot had appeared. Resulting movements are approximately 15 deg over 1 second. During this trial phase, the visual gain on random-dot scene motion is manipulated. Once head movement is complete, the participant responds whether they perceived the scene to be drifting with or against the direction of their own head movement in the world reference frame. The participant provides an answer by pressing a key on the keyboard. For clarity, the bottom row illustrates scene motion that is with and against the direction of head motion in the world reference frame.

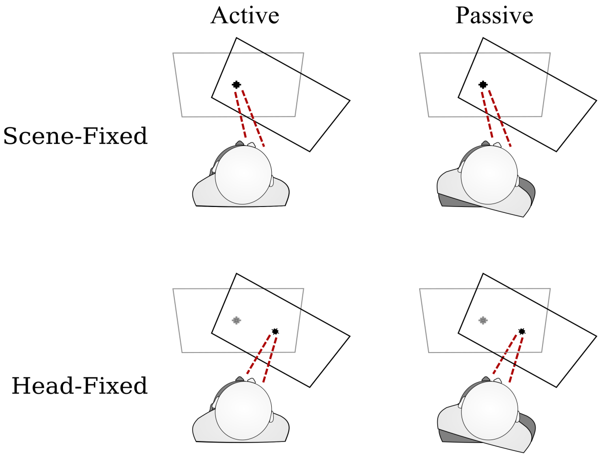

Figure 2. Experimental design - Active/Passive and Scene-Fixed/Head-Fixed conditions. The left column demonstrates the active condition, in which subjects perform yaw head rotations. The right column demonstrates the passive condition, in which a subject's whole body is rotated using a rotating chair. The top row of the diagram demonstrates the scene-fixed condition. Subjects move their head left or right, while fixating a point that remains stationary in the visual scene such that it must move on the head-mounted display. The bottom row depicts the head-fixed condition, where subjects fixate a point that moves with them, staying central relative to the head and head-mounted-display as they rotate.

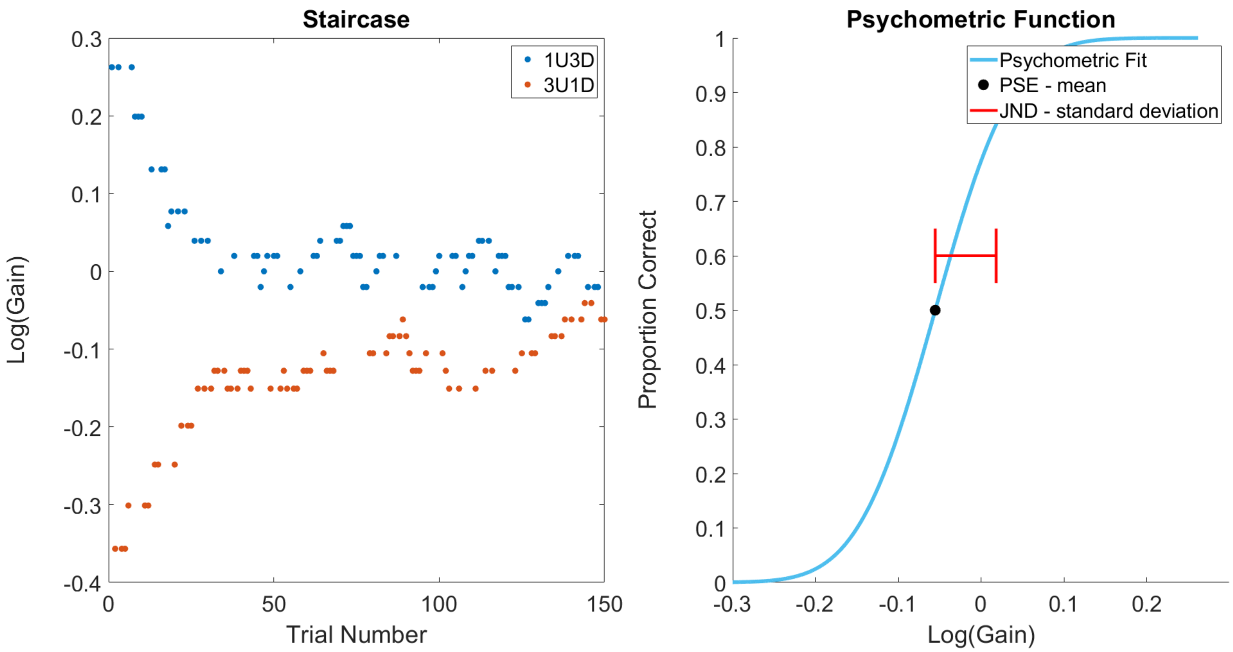

Figure 3. Example staircase and psychometric fit. Staircase procedure (left): over consecutive trials in a given block, the visual gain was manipulated according interleaved 3-down-1-up (3D1U, orange) and 3-up-1-down (3U1D, blue) staircases. Step size was uniform in units of log(gain): 0.04 before the first staircase reversal, and 0.02 after. Note that gain of 1 (visual motion equal and opposite head motion) results in a value of log(gain)=0. Psychometric fit (right): a cumulative Gaussian is fit to the resulting data (blue). The resulting mean parameter (black dot) is taken as the point of subjective equality (PSE), that is the single visual gain value perceived as stationary leading to 50 percent "with"/"against" responses. The resulting standard deviation parameter (red interval) is taken as the just-noticeable difference (JND), that is the increase/decrease in gain that leads to 84/16 percent "against" responses. It provides a measure of precision

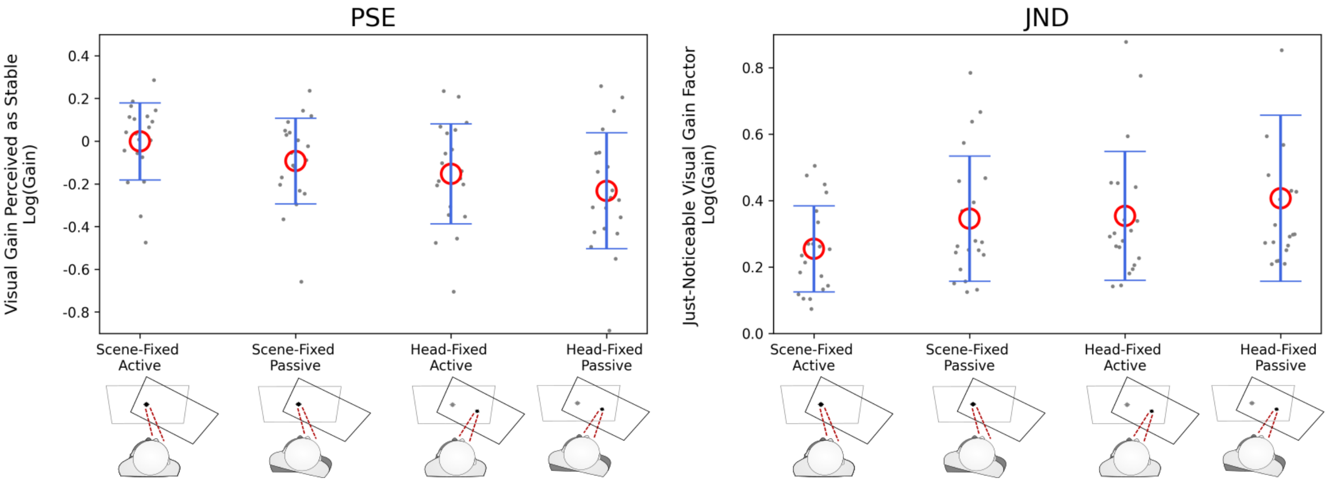

Figure 4. Effect of head and eye movements on stationarity perception. The point of subjective equality (PSE) (left panel) is a measure of accuracy, indicating the visual gain perceived as stationary in each participant and condition. The just-noticeable difference (JND) (right panel) is a measure of precision, indicating the range of gain values that support stationarity perception in each condition. Blue and red points are data from Scene-fixed and Head-fixed conditions, respectively. Dark and light points are data from Active and Passive conditions, respectively.