An allocentric human odometer for perceiving distances on the ground plane

Liu Zhou[1], Wei Wei[1, 2], Teng Leng Ooi[2]*, & Zijiang J. He[1]*

[1]Department of Psychological and Brain Sciences, University of Louisville, Louisville, Kentucky 40292, USA.

[2]College of Optometry, The Ohio State University, Columbus, Ohio 43210, USA.

*e-mail: ooi.22@osu.edu; zjhe@louisville.edu

Abstract

We reliably judge locations of static objects when we walk despite the retinal images of these objects moving with every step we take. Here, we showed our brains solve this optical illusion by adopting an allocentric spatial reference frame. We measured perceived target location after the observer walked a short distance from the home base. Supporting the allocentric coding scheme, we found the intrinsic bias[1, 2], which acts as a spatial reference frame for perceiving location of a dimly lit target in the dark, remained grounded at the home base rather than traveled along with the observer. The path-integration mechanism responsible for this can utilize both active and passive (vestibular) translational motion signals, but only along the horizontal direction. This anisotropic path-integration finding in human visual space perception is reminiscent of the anisotropic spatial memory finding in desert ants[3], pointing to nature's wondrous and logically simple design for terrestrial creatures.

Teaser title: A human odometer fitting of our terrestrial niche

Introduction

When viewing natural scenes, we readily appreciate the visual space expanse that "fills" the scene being enveloped by the sky above and terrain below our feet (figure 1a). Notably, within this phenomenological visual space, our visual system employs the prevalent ground surface, where creatures and objects frequently interact, as a reference frame for coding spatial locations[4-6]. Empirical findings have revealed the ground-based spatial coding scheme accurately localizes objects when the horizontal ground surface is continuous and carries rich depth cues[6-14]. Attesting to the significant role of the ground surface, it has been found that object localization becomes inaccurate when the ground surface is disrupted by a gap or an occluding box[6, 13]. Similarly, object localization is inaccurate in the dark when the ground is not visible[1, 2, 15]. In fact, research conducted in total darkness to prevent the visual system from obtaining depth information from the ground reveals the visual system defaults to using an implicit curved surface representation that we refer to as the *intrinsic bias* to localize objects[1, 2, 16]. Figure 1b illustrates the intrinsic bias's curved representation in the dark (dashed white curve) that leads to an observer perceiving a dimly lit target (green ring) at the intersection (green disc) between its projection line from the eye and the intrinsic bias. Furthermore, when a parallel row of texture elements on the ground serves as the depth cue (figure 1c), object localization becomes more accurate. It is as if the objects are now located along a curved reference (solid white curve), which is less slanted than the intrinsic bias. This suggests the intrinsic bias, acting as the prototype spatial reference of the ground surface when visual cues are not visible, integrates with the visible depth cue to form a new ground surface representation[17].

Nevertheless, current understanding of the ground-based spatial coding scheme is limited to situations where the tested observer stood still. Namely, just as in figure 1b, now redrawn in figure 2a and referred to as the *baseline-stationary* condition, the observer is instructed to stand still at one location during the test, which anchors his/her intrinsic bias on the ground under his/her feet. In this way, object location in visual space can be computed relative to the intrinsic bias. However, because retinal images of the stationary environment move when one moves (e.g., walk), it is unclear what additional computation the ground-based spatial coding scheme requires to construct a stable perceptual space. We proposed that during translational movements, such as walking, the visual system anchors the intrinsic bias at one location on the ground surface using a

path integration mechanism[18-25]. By doing so, the visual system can obtain a stable ground surface representation. Accordingly, the ground-based reference frame can use an allocentric, i.e., world-centered, coordinate system to code locations in visual space. This allocentric hypothesis predicts that during self-motion, the intrinsic bias is fixed to the ground location before the motion begins (home base) and remains at the same ground location during self-motion. This hypothesis stands in contrast to an alternative egocentric coordinate system hypothesis, which predicts the intrinsic bias moves along with the observer's body. To distinguish between these two hypotheses, we began by measuring the intrinsic bias when the observer walked in the dark. Consider the scenario where an observer walks from his/her home base (blue cross added in figure for illustrative purpose) in the dark and stops at a new location (figure 2b; red cross added for illustration). According to the allocentric hypothesis, the intrinsic bias (blue) would remain at the home base. Thus, when he/she stops and is presented with a dimly lit test target (illustrated by the green ring), the observer would perceive the target to be nearer and higher (figure 2b) than if he/she had not walked from the home base (figure 2a; we shall dub this the *baseline-stationary* condition). In contrast, if the intrinsic bias had moved along with the observer to the new location (illustrated as the grey curve in figure 2c), as predicted by the alternative egocentric hypothesis, the observer would perceive the target at the same location as that in the *baseline-stationary* condition (figure 2a).

We have premised the allocentric hypothesis on the ground-based spatial coding scheme employing the path-integration (or spatial updating) system to maintain the intrinsic bias at the home base (blue cross in figure 2b). Previous studies have found that when performing spatial memory tasks,–in conditions where visual cues are not visible in the dark, the path integration mechanism computes the traveled distance based on (non-visual) idiothetic distance information, such as from the vestibular, proprioceptive, and efference copy of motor control signals[18, 21, 23, 25]. Thus, we extended this study to characterize the behaviors of the path integration mechanism in a visual spatial perception task, to investigate if the path integration mechanism behaved in a similar manner as in the spatial memory task.

Results

Experiment 1: Testing Allocentric vs. Egocentric Hypothesis

We first verified the prediction of the allocentric hypothesis (figure 2) by testing a *baseline-stationary* condition (figure 3a) and a *walking* condition (figure 3b). In both conditions, the observer readied for the experiment by sitting on a chair in the waiting area, which was illuminated by a dimly lit LED light from the ceiling. To ready for a test trial, he/she waited for an audio tone. Upon hearing the tone, he/she stood up and aligned his/her feet at the start point (home base) that faced a black curtain in the direction of the testing area. About 30 sec later, the experimenter turned off the LED light in the waiting area and the observer drew the curtain open in the dark to begin the trial.

In the *baseline-stationary* condition, after further waiting at the home base (blue cross) for either 12 or 60 sec, the observer saw a briefly presented (1 sec) dimly lit target. He/she was instructed to judge the target location and to respond by walking blindly to the remembered target location and gesturing its height after reaching the walked destination[1, 2, 7-10]. (While walking, the observer's right hand glided along a horizontal guidance rope, drawn as the yellow line in figure 3a). Effectively, this blind walking-gesturing task reveals the perceived target location (x: walked distance; y: gestured height).

In the *walking* condition (figure 3b), the observer walked blindly from the home base (blue cross) to a new location (red cross). While walking, the observer's right hand glided along a guidance rope and stopped at the new location when he/she made contact with a soft plastic wrap on the rope (1.5m from the home base). After waiting at the new location for either 12 or 60 sec, he/she saw a briefly presented (1 sec) dimly lit target. He/she then performed the blind walking-gesturing task to indicate the perceived target location.

Figure 3c shows the average results of the two conditions. With the 12 sec waiting period, judged locations in the *walking* condition (filled green triangles) were significantly nearer than the *baseline-stationary* (filled red circles) condition ($P<0.0001$; please refer to the supplement for details of all statistical analyses). The two sets of data points are fitted by the same shaped intrinsic bias curve with about 1.35 m horizontal separation, which is close to the initially walked distance from the home base (1.5 m; in the *walking* condition). This result thus confirms the prediction of the allocentric hypothesis (figure 2b). With the 60 sec waiting period, judged locations (open triangles) in the *walking* condition had a smaller, though statistically significant, separation from

the *baseline-stationary* condition (open circles) ($P<0.01$). This suggests if the waiting period was sufficiently long, suggesting the visual memory of the home base decays over time[5, 19], the visual system automatically resets the intrinsic bias to the observer's (new/current) location, i.e., making it the updated home base.

The results above demonstrate the action of the path integration mechanism. It tracks the observer's location with respect to the external world during locomotion and plays a critical role in maintaining an allocentric reference frame. In the *walking* condition, the path-integration mechanism computes the traveled distance relative to the home base. Experiments 2-4 below explored the characteristics of the path-integration mechanism.

Experiment 2: Path Integration Affected by Cognitive Load

Since in the *walking* condition (figure 3b) observers were simply attentive to the task at hand but otherwise neutral and not subjected to excessive cognitive demands, one might assume that the path-integration mechanism operates automatically. To test this assumption, we investigated whether the path-integration mechanism requires some attentional resources to function normally[26, 27]. We thus tested a new, *divided attention walking* condition, where the observer continuously performed a cognitive task, counting number backward, while walking from the home base to the new location. We predicted this would reduce the attention resources available for path-integration. For comparison, we also tested the *baseline-stationary* condition as in Experiment 1 (figure 3a).

The average results in figure 3d reveal a small, though significant, difference between the *divided attention walking* condition and the *baseline-stationary* conditions ($P=0.022$). Noticeably, the distance underestimation in the *divided attention walking* condition (figure 3d) was much smaller than that found in Experiment 1 (figure 3c). This suggests path-integration is less effective when observers were engaged in another mental task that distracted their spatial attention.

Experiment 3: Path-integration from Vestibular Input

We addressed two issues regarding the idiothetic distance cues used by the path-integration mechanism. First, does it operate only when the observer's movement is self-initiated? Second, does it operate when observer moves in the backward direction? To answer these questions, we

129 moved the observers passively to stimulate their vestibular system[24, 28-29]. That is, instead of
130 instructing the observer to walk in the dark from the home base, he/she stood upright on a rolling
131 platform that was moved by the experimenter (insets in figure 4). Doing so negated the
132 contributions of the proprioception and motor efference copy information to path-integration while
133 keeping the vestibular information intact.

134 We first tested a new *vestibular-forward* condition in which the experimenter pushed the
135 observer on the platform forward by 1.5 m to the new location. Once at the new location, the
136 observer stepped down from the platform and stood on the floor to wait for the dimly lit test target
137 to be presented. He/she responded with the blind walking-gesturing task. For comparison, we
138 also tested the *walking* (figure 3b) and *baseline-stationary* (figure 3a) conditions as in Experiment
139 1. Figure 4a reveals the judged target locations in the *vestibular-forward condition* (blue squares)
140 were similar to the *walking* condition (green triangles; P=0.718), and were significantly nearer
141 than the *baseline-stationary* condition (red circles; P<0.0001). This indicates the path-integration
142 mechanism can function during passive movements that stimulate the vestibular system.

143 To answer the second question, we then tested a complementary condition wherein the
144 experimenter pulled the rolling platform supporting the observer backward by 1.5 m from the home
145 base (*vestibular-backward* condition). As expected, the backward displacement caused judged
146 target locations to be farther than the *baseline-stationary* condition (P<0.0001; figure 4b). Taken
147 together, the experiment reveals the path-integration mechanism can sufficiently utilize the
148 vestibular cue in both forward and backward moving directions.

149 **Experiment 4: The Allocentric Principle Is Not Task Specific**

150 Previous studies of space perception with stationary observers not undergoing self-motion
151 measured with different response tasks have shown a concordance in finding. Specifically, the
152 perceptual effects found were similar between an action-based task (blind walking-gesturing) and
153 a perception-based task (e.g., verbal reports and perceptual matching) [1, 6, 12, 30]. To confirm that
154 this is also true for observers undergoing self-motion, we repeated the *vestibular-forward* and the
155 *baseline-stationary* conditions in Experiment 3 and employed the verbal report task, wherein the
156 observer verbally reported the perceived eye-to-target distances in feet or meters. We found that
157 consistent with the results from the blind walking-gesturing task in Experiment 3 (figure 4a),

158 verbally reported distances were significantly shorter in the *vestibular-forward* than *the baseline-*
159 *stationary* condition ($P<0.0001$, figure 4c).

160 **Experiment 5: Anisotropic Path-integration – A Bias for the Ground Surface**

161 An assumption underlying the allocentric hypothesis is that adoption of the ground-based
162 spatial coding scheme is fitting for our terrestrial existence, where everyday activities, including
163 navigation, are performed with respect to the horizontal ground surface[4]. This dependence on the
164 ground surface, i.e., an ecological constraint, predicts when the observer travels on a sloping
165 surface in the dark, the visual system would not be able to simultaneously maintain the intrinsic
166 bias at the home base in the vertical and horizontal directions. To investigate this prediction of
167 anisotropic path-integration, we tested a new, *stepladder* condition, where the observer descended
168 a stepladder in the dark (figure 5a). As the self-motion here consisted of both horizontal forward
169 and vertical downward vectors, we predicted the intrinsic bias would be spatially updated in the
170 horizontal but not in the vertical direction. This prediction is illustrated in figure 5a where the
171 horizontal coordinate of the intrinsic bias (blue curve) remains at the home base location on the
172 floor while the vertical coordinate travels along with the observer. Thus, upon stepping down from
173 the stepladder and standing on the floor, the observer underestimates the horizontal distance of a
174 dimly lit target. We expected that the target underestimation will be similar to that of a *horizontal-*
175 *walking* condition (figure 5b). In contrast, if the intrinsic bias is spatially updated in both the
176 horizontal and vertical dimensions, the intrinsic bias will remain at the top of the staircase (orange
177 curve, figure 5a). The perceived target location will then be dramatically different from the
178 predicted horizontal-only-updating strategy (figure 5b).

179 Figure 5c depicts the average judged target locations (n=9) from the *stepladder* (blue
180 triangles), *horizontal-walking* (green square) and *baseline-stationary* (red circles) conditions (the
181 *baseline-stationary* condition's setup was the same as in figure 3a). A comparison between the
182 *horizontal-walking* and the *baseline-stationary* conditions reveals the judged horizontal distances
183 were significantly shorter in the *horizontal-walking* condition ($P<0.001$). The two sets of data are
184 fitted by the same intrinsic bias profile with a horizontal separation of 1.0 m, which is close to the
185 walked distance (1.06 m) in the *horizontal-walking* condition.

The judged horizontal distances in the *stepladder* condition were significantly shorter than that in the *baseline-stationary* condition ($P<0.001$). This confirms the prediction that while descending the stepladder, the visual system spatially updated the horizontal but not the vertical vector of the intrinsic bias (blue intrinsic bias, figure 5a). Further supporting this, we found the data from the *stepladder* and the *horizontal-walking* conditions overlap substantially.

Taken together, our experiment revealed when stepping down from a stepladder, the horizontal coordinate of the intrinsic bias is kept at the home base on the floor while the vertical coordinate moves downward with the body. This suggests human's path-integration mechanism, responsible for allocentric spatial coding, operates much more efficiently along the horizontal ground plane than in the vertical direction.

Discussion

We constantly move relative to the environment, which causes our retinal images of the environment to also be in constant motion. Yet, we reliably localize objects as we navigate our environment. This raises the question of how our visual system creates a stable visual space for us to operate and plan future actions, such as walking, despite the inadequate retinal inputs. While it was known the visual system relies predominantly on the ground surface to accurately localize objects, it was unclear what coordinate system it uses for coding visual space. In this study, we proposed the visual system uses an allocentric, or world-centered, coordinate system for the ground-based reference. In this way, the visual system could alleviate the challenge of constant retinal motion during self-motion by streamlining online location computation to only the moving observer and the object of interest. The locations of surrounding objects are anchored to a fixed position on the ground negating excessive computation, thereby reducing coding redundancy. Our main finding provides support for the allocentric hypothesis. We showed during self-motion, the intrinsic bias that acts as the spatial reference of the ground surface when visual cues are not visible, is fixed to the ground location before the motion begins (home base) and remains at the same ground location during self-motion.

Previous studies mainly tested location judgments of observers who stood stationary at one location on the ground while viewing a stationary target (similar to our *baseline-stationary* condition). Because the observers need not generate significant self-motion in those studies, it was not possible to investigate the allocentric hypothesis. In principle, without body motion, the allocentric coordinate system does not have an advantage over the egocentric coordinate system for processing efficiency. In fact, previous findings with stationary observers led to the widely accepted notion that absolute distance judgment relies on the egocentric coordinate system.

The visual system relies on the path-integration mechanism to ground the intrinsic bias at the home base. The path-integration mechanism can function when the observer's motion is passive, suggesting the traveled distance can be based on the vestibular signals alone. Also, the path-integration mechanism fails to operate efficiently when the observer's attention is distracted to an unrelated cognitive demand during walking. This suggests some degree of attentional effort is required to reliably path integrate an observer's body location relative to the environment. Lastly, we showed the path-integration mechanism is more efficient in the horizontal than the vertical direction (anisotropic with a ground-bias). This observation further highlights one of the most important design principles underlying our sensory and motor systems, namely, fitting our terrestrial niche. We inhabit a land-based environment and most of our daily activities are structured relative to the ground. The current finding on the vertical-horizontal anisotropy of the path integration mechanism for visual space perception is consistent with previous studies on spatial memory, which revealed the path-integration systems of land-dwelling animals are more efficient in estimating horizontally traveled distance[3, 31-34]. In particular, our observation is reminiscent of the behavioral studies of the desert ants, Cataglyphis fortis. Wohlgemuth et al (2001) found desert ants used path integration to return from foraging excursions on a shortcut way to their nests (also see Ronacher, 2020). Specifically, it was revealed when the ants were crawling over a slope surface with uphill and downhill trajectory, their path integration system computed only the horizontal component of the traveled distance. We note, however, further research is needed to investigate whether the same path-integration mechanism is employed for a spatial memory task (previous studies by others) and a spatial perception task (current study).

Overall, our study suggests the visual system, by anchoring its spatial reference frame to the ground during navigation can create a world-centered visual space. One advantage of adopting

this space coding strategy is that when we travel, the spatial representations of most surrounding objects, which are static relative to the ground, will remain constant (static) despite the retinal image motion. Accordingly, our visual system only needs to dynamically update the spatial representations of our bodies and the few objects that are moving relative to an allocentric origin on the ground (within a world-centered context). A consequential byproduct of operating within the framework of an allocentric, world-centered visual space is the perceived stability of our visual environment during locomotion.

Materials and Methods

Observers. Fourteen observers (age=24.47 ± 1.45 years old; eye height= 1.60 ± 0.02 m; 7 males and 7 females) participated in Experiments 1-4 and nine observers (age=22.78 ± 1.22 years old; eye height= 1.58 ± 0.04 cm; 6 males and 3 males) participated in Experiment 5. They were naïve to the purpose of the study and signed the informed consent form at the start of the study. All observers had normal, or corrected-to-normal, visual acuity (at least 20/20) and a stereoscopic resolution of a 20 arc sec or better. They viewed the visual scene binocularly. A within-subject experimental design was used. The study protocol was approved by the Institutional Review Board and followed the tenets of the Declaration of Helsinki.

General stimulus and testing environment.

All experiments 1-4 were performed in a dark room whose layout and dimensions were unknown to the observers. One end of the room, just before the testing area, served as the waiting area (~3 m^2) for the observer. The waiting area had a chair facing the wall for the observer to sit in between trials so that his/her back faced the testing area. Two white LED lights on the ceiling provided ambient illumination while the observer waited in between trials. The testing and waiting areas were separated by a black curtain. A long guidance rope (0.8 m above the floor) was tied to both ends of the room and served to guide the observer while walking blindly. A plastic wrap was tied to the guidance rope on the part of the rope located in the waiting area near the curtain to mark the start point (home base). To ready for a trial, the observer walked to this start point and faced the test area while holding onto the plastic wrap on the rope and called out "ready". The curtain was then drawn open for the trial to begin.

The dimly lit test target used in all experiments was a diffused green LED (0.16 cd m^{-2}). The LED was placed in the center of a ping-pong ball that was encased in a small opaque box. An adjustable iris-diaphragm aperture was placed at the front of the ping-pong ball to keep its visual angular size constant at 0.22° when measured at the eye level. During testing, the target was displayed with a 5 Hz flicker for 1 sec. Music was played aloud during the entire experimental session to mask extraneous auditory information during the experiments.

Observer's response tasks. The main task used in all experiments was the *blind walking-gesturing task* [35, 36.] For each trial, the observer stood by the guidance rope in the dark and judged the location of the 1 sec flickering target. After which he/she then put on the blindfold and called out "ready to walk". This signaled the experimenter to quickly remove the target and shook the guidance rope to indicate it was safe to walk. The observer walked while sliding his/her right hand along the guidance rope until he/she reached the remembered target location. Once there, he/she indicated the remembered target height with his/her left hand and called out "done". The experimenter turned on the flashlight, marked the observer's feet location, measured the gestured height, and asked the observer to turn around and walk back to the start point using the guidance rope. When the observer arrived at the start point, the experimenter turned on the ceiling LED lights in the waiting area. The observer then removed the blindfold, sat down, and waited for the next trial.

An additional task used, in Experiment 4, was the *verbal reporting task* [35, 37]. Here, the observer stood next to the guidance rope in the dark and viewed the target for 5 sec to judge its absolute distance between the target and his/her eyes. Once the target was turned off, an audio tone was presented to signal the observer to immediately report the estimated distance either in meters or feet. For both this and the blind walking-gesturing task, the observer was provided five practice trials before each test session. No feedback regarding performance was provided to the observer during the practice or test session.

Experiment 1

Design: Two viewing conditions were tested: *walking* and *baseline-stationary*. Each condition was tested separately with a 12 sec and a 60 sec waiting period. The target was placed at one of six locations in both conditions. Four locations were on the floor at 3, 4.25, 5.25, or 6.5 m from the observer, the fifth location was 0.5 m above the floor at 6.5 m from the observer, and the sixth location was at the observer's eye level and 6.5 m from the observer. Testing of the 4.25, 5.25, and 6.5 m targets on the floor were repeated three times while testing of the remaining targets was repeated twice. A total of 60 trials were run over two days. The order of stimulus presentation was randomized.

Procedure: While the observer sat at the waiting area before each trial, he/she was informed of the upcoming test condition (*baseline-stationary* or *walking*) and of the waiting period (12 sec or 60 sec). After that, an audio tone was presented to signal to the observer to walk to the start point (home base) and face the black curtain in the direction of the testing area. About 30 sec later, the experimenter turned off the ambient LED lights in the waiting area and the observer drew the curtain open in the dark.

For the *baseline-stationary* condition, the observer stood in the dark at the start point (home base) over the predetermined waiting duration (12 or 60 sec). He/she was instructed to stand upright with minimal head motion during the waiting period, and to expect hearing a pure tone at the end of the waiting period. Roughly, two sec after hearing the tone, the test target was presented at one of the six predetermined locations. The observer's task was to judge its location and perform the blind walking-gesturing task. For the *walking* trial, the observer stood at the start point until a white noise (instead of pure tone) was heard. He/she then walked forward until his/her right hand touched a plastic wrap tied on the guidance rope at the new location (1.5m from the start point). He/she then stopped walking, called out "ready" and waited there for either 12 sec or 60 sec before the test target was presented. The remaining procedural steps were the same as in the *baseline-stationary* condition.

Experiment 2

Design and Procedure: Both the *baseline-stationary* and *walking* conditions with the 12 sec waiting period were tested but with one modification. The modification was that during the 12 second waiting period at the start point (home base) (for *baseline-stationary* condition), or while walking to, and waiting at the new location (for *walking* condition), the observer performed a counting task. The experimenter would provide a random number between 50 and 99 for the observer to count backward. The same six target locations tested in Experiment 1 were tested here. A total of 30 trials were run. The order of stimulus presentation was randomized.

Experiment 3

3.1 Vestibular-forward condition

Design: The two conditions from Experiment 1 (*walking* and *baseline-stationary* conditions with the 12 seconds waiting period) and a new, *vestibular-forward* condition were tested. The latter *vestibular-forward* condition was the same as in the *walking* condition, except that the observer now stood on a rolling-platform that was pushed forward by the experimenter for 1.5 m. The rolling-platform had four wheels that was 0.30 m above the floor. The same six target locations tested in Experiment 1 were tested here. There were 60 trials in total with 30 trials tested per day. The order of stimulus presentation was randomized.

Procedure: Two experimenters (A and B) conducted the experiment. Before each trial, experimenter A instructed the observer of the upcoming test condition and to prepare for an audio signal. Once the signal was heard, the observer stepped onto the rolling-platform at the start point (home base) behind the curtain with eyes opened and facing the test area. He/she then called out "ready". A second audio signal was presented after 30 seconds. A pure tone indicated to the observer <u>to stay</u> (*baseline-stationary* condition), while a white noise indicated <u>to walk</u> (*walking* condition) or <u>to be pushed forward</u> (*vestibular-forward* condition). At the same time, experimenter A turned off the LED lights in the waiting area, and the observer drew the curtain open. For the *baseline-stationary* condition trial, the observer stepped down from the rolling-platform, stood at the start point, and called out "ready to view". For the *walking* condition trial, the observer stepped down from the rolling-platform, and walked forward until he/she touched the plastic wrap tied onto the guidance rope at the new location (1.5 m), and then stopped and called out "ready to view". For the *vestibular-forward* condition trial, the observer kept standing on the rolling-platform. Experimenter B verbally informed the observer the trial was starting and pushed the rolling platform forward by 1.5 m to the new location. Upon arrival, Experiment B instructed the observer to step down from the rolling platform and to call out "ready to view". For all three trial types, after the observer called out "ready to view", Experimenter A turned on the test target after a 12 sec waiting period for the observer to judge its location. He/she then responded by performing the blind walking-gesturing task.

<u>*3.2 Vestibular-backward* condition</u>

Design and procedure: The *baseline-stationary* condition was the same as in 3.1 above. The *vestibular-backward* condition was modified from the *vestibular-forward* in 3.1 above, by pulling

the rolling-platform backward for 1.5 m. The same six target locations as in 3.1 were tested. A total 30 of trials were tested in one session. The order of stimulus presentation was randomized.

Experiment 4

Design: The two conditions tested were the *vestibular-forward* and *baseline-stationary* conditions as in 3.1 above. But instead of the six test target locations, only five target locations on the floor were tested. The target location at 6.5 m distance and 0.5m above the ground was not tested. All targets were presented for 5 sec. Twenty-eight trials in total were tested in one session. The order of stimulus presentation was randomized.

Procedure: The procedure was the same as in 3.1 above, except the observers performed the verbal report task. To do so, the observer judged the absolute distance between the target and his/her eyes. Immediately after the target was turned off, an audio signal was presented to prompt the observer to report the estimated distance aloud either in meters or feet.

Experiment 5

Design: Three conditions (*baseline-stationary*, *horizontal-walking* and *stepladder*) were tested in two different blocks. The *horizontal-walking* and *baseline-stationary* conditions were mixed in test block-A, while the *stepladder* and *baseline-stationary* conditions in test block-B. Each block consisted of 20 test trials. The order of test conditions within each block was randomized. Each block was tested in a daily session and ran twice. The testing order of the four blocks was alternated between observers.

Procedure: Before each trial, the experimenter informed the observer of the condition to be tested. The *baseline-stationary* and *horizontal-walking* conditions were conducted as in the main experiment (Experiment 1). For the *stepladder* condition, the observer first ascended the stepladder, and waited for an audio tone that signaled to descend the stepladder with eyes opened and looking at the invisible horizon (in the dark). Upon reaching the foot of the stepladder, he/she held onto the guidance rope and walked forward until he/she felt the plastic wrap on the rope at the new location. The observer then stood still and waited for 12 s before the test target was presented. After judging the target's location, he/she put on the blindfold and called out "ready" to begin the blindfolded walking-gesturing task.

Statistical tests. Data were analyzed using analysis of variance with repeated measures. The Mauchly's test was applied to verify the assumption of sphericity.

References

1. Ooi, T. L., Wu, B. & He, Z. J. Distance determined by the angular declination below the horizon. *Nature* **414**, 197–200 (2001).
2. Ooi, T. L., Wu, B. & He, Z. J. Perceptual space in the dark affected by the intrinsic bias of the visual system. *Perception* **35**, 605–624 (2006).
3. Wohlgemuth, S., Ronacher, B., & Wehner, R. Ant odometry in the third dimension. *Nature* **411**, 795-798 (2001).
4. Gibson, J. J. *The Perception of the Visual World* (Houghton Mifflin, 1950).
5. Sedgwick, H. A. Space perception. in *Handbook of Perception and Human Performance* (eds Boff, K. R. et al.) 21.1–21.57 (New York Wiley, 1986).
6. Sinai, M. J., Ooi, T. L. & He, Z. J. Terrain influences the accurate judgement of distance. *Nature* **395**, 497–500 (1998).
7. Thomson, J. A. Is continuous visual monitoring necessary in visually guided locomotion? *J. Exp. Psychol. Hum. Percept. Perform* **9**, 427–443 (1983).
8. Rieser, J. J., Ashmead, D., Talor, C. & Youngquist, G. Visual perception and the guidance of locomotion without vision to previously seen targets. *Perception* **19**, 675–689 (1990).
9. Loomis, J. M., Da Silva, J. A., Fujita, N. & Fukusima, S. S. Visual space perception and visually directed action. *J. Exp. Psychol. Hum. Percept. Perform* **18**, 906–921 (1992).
10. Loomis, J. M., DaSilva, J., Philbeck, J. W. & Fukusima, S. S. Visual perception of location and distance. *Curr. Dir. Psychol. Sci* **5**, 72–77 (1996).
11. Meng, J. C. & Sedgwick, H. A. Distance perception mediated through nested contact relations among surfaces. *Percept. Psychophys.* **63**, 1–15 (2001).
12. Wu, B., Ooi, T. L. & He, Z. J. Perceiving distance accurately by a directional process of integrating ground information. *Nature* **428**, 73–77 (2004).
13. He, Z. J., Wu, B., Ooi, T. L., Yarbrough, G., & Wu, J. Judging egocentric distance on the ground: Occlusion and surface integration. Perception, 33, 789–806 (2004).
14. Bian, Z., Braunstein, M. L. & Andersen, G. J. The ground dominance effect in the perception of 3-D layout. *Percept. Psychophys* **67**, 802–815 (2005).

15. Philbeck, J. W. & Loomis, J. M. Comparison of two indicators of perceived egocentric distance under full-cue and reduced-cue conditions. *J. Exp. Psychol. Hum. Percept. Perform.* **23**, 72–85 (1997).
16. Gogel, W. C. & Tietz, J. D. A comparison of oculomotor and motion parallax cues of egocentric distance. *Vision Res.* **19**, 1161–1170 (1979).
17. Wu, J., Zhou, L., Shi, P., He, Z. J. & Ooi, T. L. The visible ground surface as a reference frame for scaling binocular depth of a target in midair. J. Exp. Psychol. Hum. Percept. Perform. 41, 111–126 (2015).
18. Mittelstaedt ML, Mittelstaedt H. Homing by path integration in a mammal. Naturwissenschaften. 67, 566–567 (1980).
19. Wehner, R. & Srinivasan, M. V. Searching behavior of desert ants, genusCataglyphis (Formicidae, Hymenoptera). *J. Comp. Physiol.* **142**, 315–318 (1981).
20. Rieser, J. J. Access to knowledge of spatial structure at novel points of observation. *J. Exp. Psychol. Learn. Mem. Cogn.* **15**, 1157-1165 (1989).
21. Etienne AS, Maurer R, Séguinot V. Path integration in mammals and its interaction with visual landmarks. J Exper Biol. 199, 201-209 (1996).
22. McNaughton, B. L. et al. Deciphering the hippocampal polyglot: the hippocampus as a path integration system. *J. Exp. Biol.* **199**, 173-185 (1996).
23. Loomis, J. M., Klatzky, R. L., Golledge, R. G. & Philbeck, J. W. Human navigation by path integration. In *Wayfinding: Cognitive mapping and other spatial processes* (eds Golledge, R. G) 125-151, (Baltimore Johns Hopkins, 1999).
24. Mittelstaedt, M. L. & Mittelstaedt, H. Idiothetic navigation in humans: estimation of path length. *Exp. Brain Res.* **139**, 318–332 (2001).
25. Loomis, J. M., Klatzky, R. L. & Giudice, N. A. Representing 3D space in working memory: Spatial images from vision, hearing, touch, and language., in *Multisensory imagery* (eds Lacey, S. & Lawson, R.) 131–155 (New York Springer, 2013).
26. Amorim, M. A., Glasauer, S., Corpinot, K. & Berthoz, A. Updating an object's orientation and location during nonvisual navigation: A comparison between two processing modes. *Percept. Psychophys.* **59**, 404-418 (1997).
27. Cavanagh, P., Hunt A. R, Afraz. A. & Rolfs, M. Visual stability based on remapping of attention pointers. *Trends Cogn Sci*.**14**, 147–153 (2010).

452 28. Israël, I. & Berthoz, A. Contribution of the otoliths to the calculation of linear
453 displacement. *J. Neurophysiol.* **62**, 247–263 (1989).
454 29. Cohen, H. S. Vestibular disorders and impaired path integration along a linear trajectory.
455 *J. Vestib. Res.* **10**, 7–15 (2000).
456 30. Zhou, L., He, Z. J. & Ooi, T. L. The visual system's intrinsic bias and knowledge of size
457 mediate perceived size and location in the dark. *J. Exp. Psychol. Learn. Mem. Cogn.* **39**,
458 1930–1942 (2013).
459 31. Jovalekic, A., Hayman, R. M., Becares, N., Reid, H., Thomas, G., Wilson, J. J., et al.
460 Horizontal biases in rats' use of three-dimensional space. *Behavioural Brain Research,*
461 **222(2)**, 279-288 (2011).
462 32. Hayman, R. M., Verriotis, M. A., Jovalekic, A., Fenton, A. A., & Jeffery, K. J.
463 Anisotropic encoding of three-dimensional space by place cells and grid cells. *Nature*
464 *Neuroscience*, **14(9)**, 1182-1188 (2011).
465 33. Zwergal, A., Schoberl, F., Xiong, G., Pradhan, C., Covic, A., Werner, P., et al.
466 Anisotropy of human horizontal and vertical navigation in real space: Behavioral and
467 PET correlates. *Cerebral Cortex*, **26(11)**, 4392-4404 (2015).
468 34. Ronacher, B. Path integration in a three-dimensional world: the case of desert ants. *J*
469 *Comp Physiol A,* **206**, 379–387 (2020).
470

Acknowledgements

Funding: The work was supported by a grant from the National Institutes of Health (EY033190) to Z.J.H. and T.L.O. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions: T.L.O. and Z.J.H. conceived of the project. L.Z, T.L.O. and Z.J.H. designed experiments and analyses. L.Z. and W.W. performed the experiments and analyzed the data. L.Z. contributed to the manuscript writing; T.L.O. and Z.J.H. wrote the manuscript.

Competing interests: The authors declare no competing interests.

Data and materials availability: All data are available in the main text or the supplementary materials. The data sets generated are available from the corresponding authors upon reasonable request.

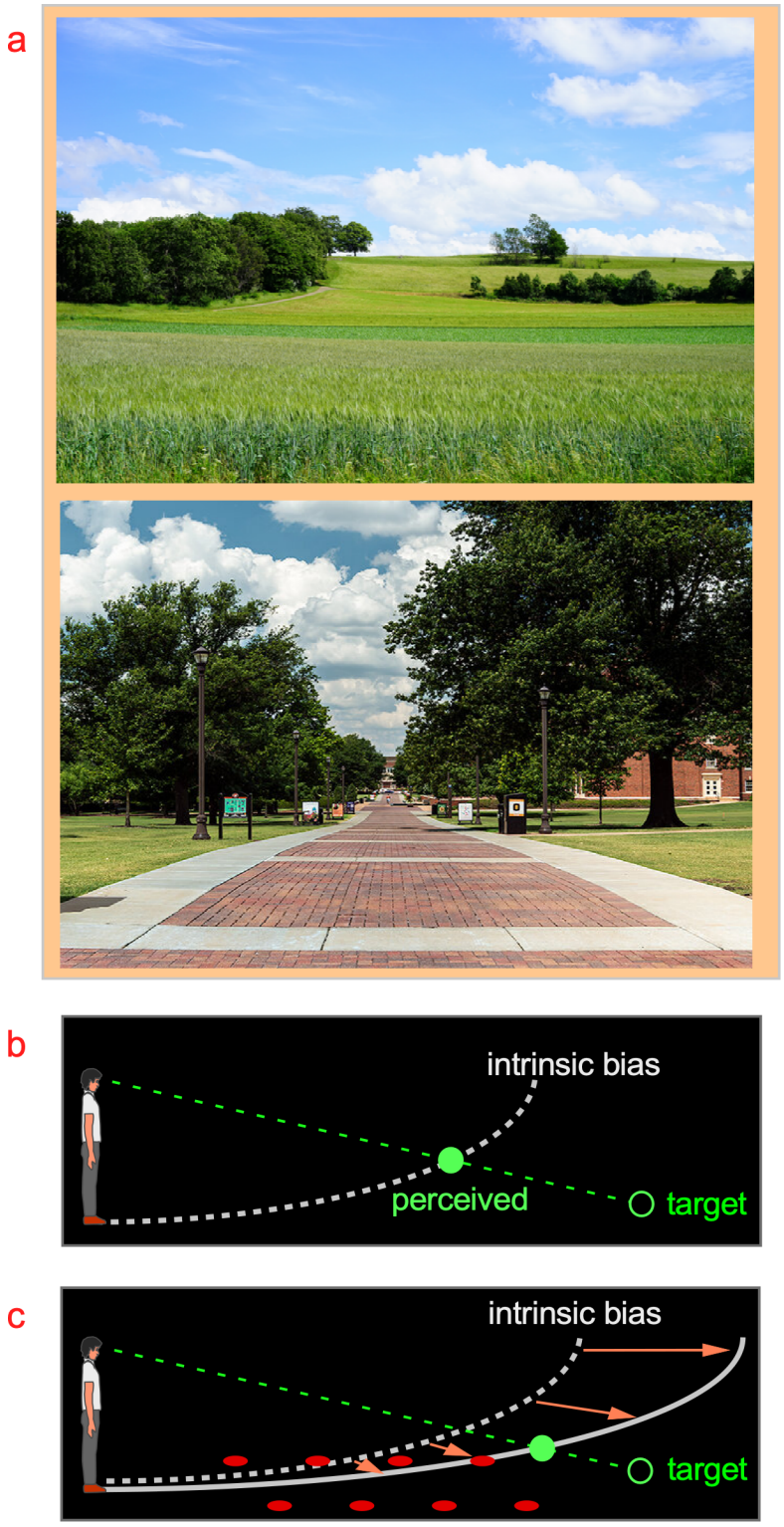
a
b
intrinsic bias
perceived
target
c
intrinsic bias
target

Fig 1

483

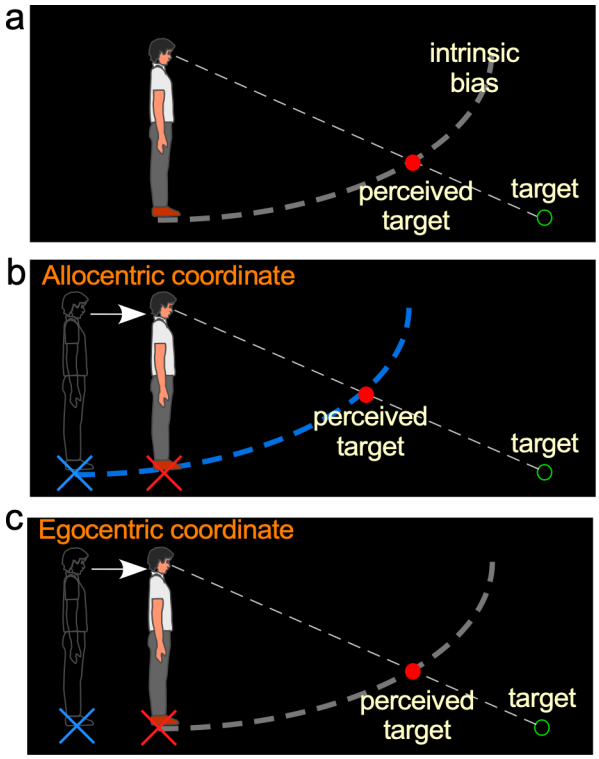
a
intrinsic
bias
perceived
target
target
b
Allocentric coordinate
perceived
target
target
c
Egocentric coordinate
perceived
target
target

Fig 2

484

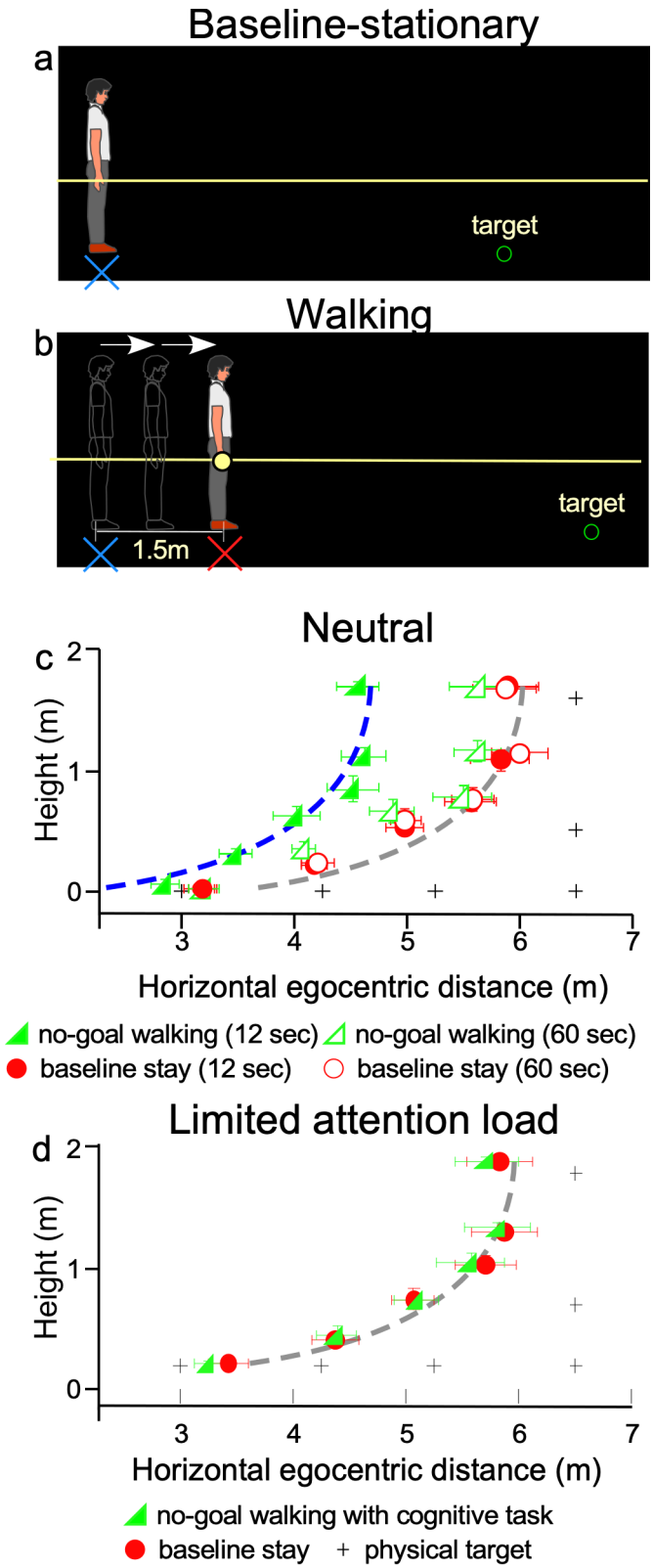
Baseline-stationary
a
target
Walking
b
target
1.5m
Neutral
c
Height (m)
Horizontal egocentric distance (m)
no-goal walking (12 sec)
no-goal walking (60 sec)
baseline stay (12 sec)
baseline stay (60 sec)
Limited attention load
d
Height (m)
Horizontal egocentric distance (m)
no-goal walking with cognitive task
baseline stay
physical target

Fig 3

485

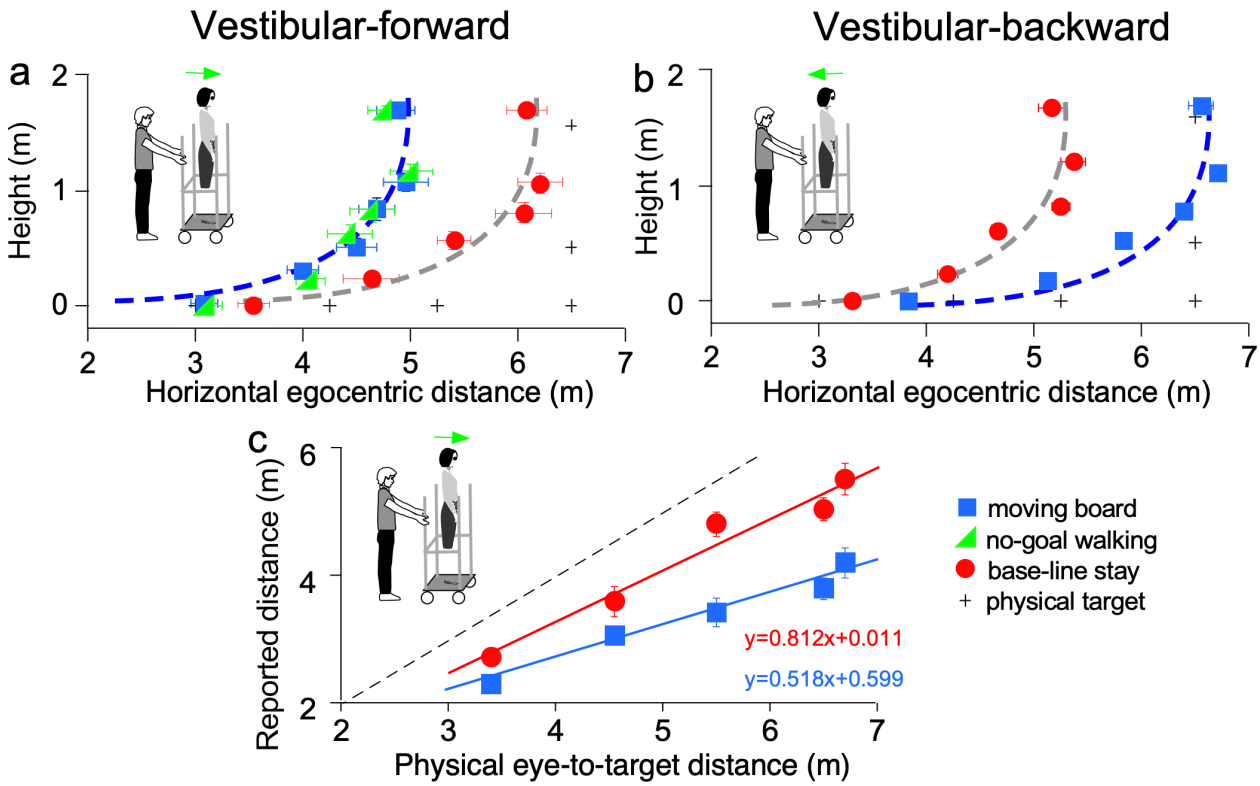
Vestibular-forward
a
Height (m)
Horizontal egocentric distance (m)
Vestibular-backward
b
Height (m)
Horizontal egocentric distance (m)
c
Reported distance (m)
Physical eye-to-target distance (m)
y=0.812x+0.011
y=0.518x+0.599
moving board
no-goal walking
base-line stay
physical target

Fig 4.

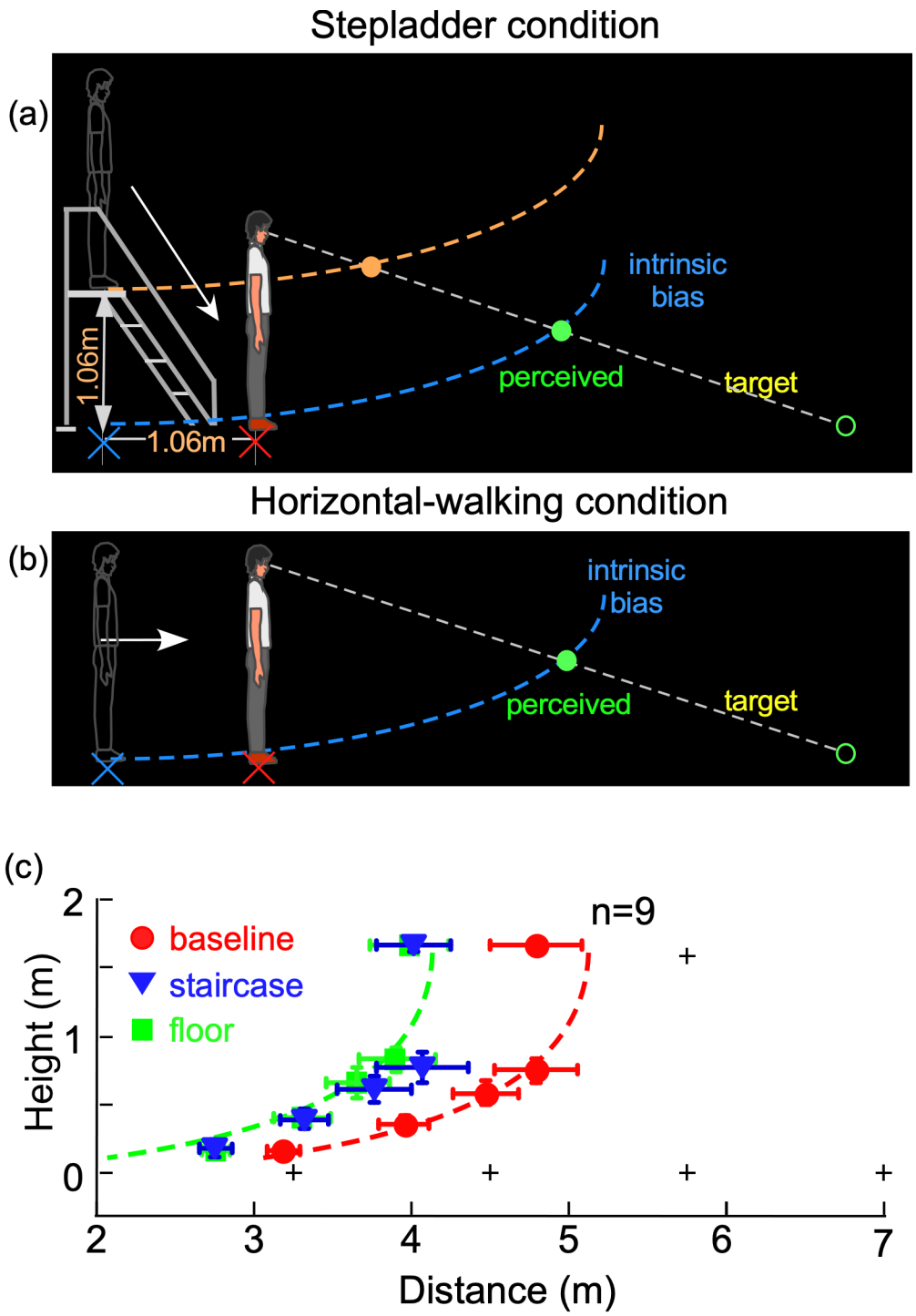
Stepladder condition
(a)
intrinsic
bias
perceived
target
1.06m
1.06m
Horizontal-walking condition
(b)
intrinsic
bias
perceived
target
(c)
n=9
baseline
staircase
floor
Height (m)
Distance (m)
0
1
2
2
3
4
5
6
7

Fig 5.

Figure legends

Figure 1. The ground-based spatial coding scheme. a. Natural scenes are largely enveloped by the prevalent sky above us and the ground on which we stand. Evidence suggests target locations are judged with respect to the ground surface, which is our terrestrial niche. **b.** The visual system uses the intrinsic bias, an implicit curved surface representation in the dark (illustrated as dashed white curve), to locate a dimly-lit target (illustrated as unfilled green disc). The target is perceived (filled green disc) at the intersection between the projection line from the eye to the target and the intrinsic bias. Essentially, the intrinsic bias acts like the visual system's internal model of the ground surface. **c.** When the ground becomes visible, the intrinsic bias integrates with the visible depth cues to form a ground surface representation, which serves as a reference frame to code target location. For example, the parallel rows of texture elements on the ground (illustrated as filled red circles) provide the depth cue for the visual system to construct a ground surface representation (solid white curve) that is less slanted than the intrinsic bias. This leads to a more accurate target localization than in the dark (b).

Figure 2. Hypotheses and predictions of Experiment 1. **a.** *Baseline-stationary* condition. A non-moving, static observer perceives the dimly lit target (illustrated as unfilled green disc) at the intersection (illustrated as filled red disc) between the projection line from the eye to the target and the intrinsic bias (dashed white curve). **b.** The allocentric hypothesis predicts when the observer walks forward from the home base (illustrated as blue cross) to a new location (illustrated as red cross), the visual system relies on the path-integration mechanism to keep the intrinsic bias (illustrated as dashed blue curve) at the home base while monitoring the homing error vector (illustrated as white arrow). From the new location, he/she perceives the target (illustrated as unfilled green circle) at the intersection (illustrated as red disk) between the intrinsic bias anchored at the home base and the projection line from the eye to the target. **c.** The egocentric hypothesis predicts the intrinsic bias tags along with the observer's body (illustrated as gray intrinsic bias) when he/she walks to the new location (illustrated as red cross). From the new location, the target (illustrated as unfilled green circle) is perceived (illustrated as red disk) at the same location as in the *baseline-stationary* condition (in **a**).

Figure 3. Experiment 1. **a.** *Baseline-stationary* condition: The observer stood at the home base (illustrated as blue cross) and judged the location of a briefly presented dimly-lit target. **b**. *Walking* condition: The observer walked blindly from the home base (illustrated as blue cross) for 1.5 m to the new location (illustrated as red cross). After a short delay (12 or 60 sec), the dimly-lit target was presented for him/her to judge the location. **c.** Graph plotting the average (n=8) judged target locations from Experiment 1. The black plus symbols, here and in other graphs, represent the physical target locations. The filled and open green triangle symbols represent the results for the 12 sec and 60 sec waiting periods, respectively, in the *walking* condition. The red circle symbols represent the *baseline-stationary* condition. With the 12 sec waiting period, judged locations (filled green triangles) in the *walking* condition were significantly nearer than in the *baseline-stationary* (filled red circles) condition. With the 60 sec waiting period, judged locations (open triangles) had a much smaller separation from the *baseline-stationary* condition. The blue and gray curves with the same shape are the intrinsic bias fitted to the data by eye. Their horizontal shift is about 1.35m. **d**. Graph plotting the average (n=8) judged target locations from Experiment 2. The green triangle and red circle symbols represent results from the *divided-attention-walking* and *baseline-stationary* conditions, respectively. Judged locations from both conditions were similar and are fitted by the same intrinsic bias curve. Error bars represent the standard errors of the mean.

Figure 4. Results of Experiment 3 (**a, b**) and Experiment 4 (**c**). **a.** *Vestibular-forward* condition: The blue square and green triangle symbols represent the *vestibular-forward* and *walking* conditions, respectively, and they show that judged target locations were similar (n=8). These judged target locations were nearer than those from the *baseline-stationary* condition (red circles). **b.** *Vestibular-backward* condition: Backward translation caused judged target locations (blue squares) to be farther than those from the *baseline-stationary* condition (red circles) (n=8). **c.** Experiment 4. Verbally reported results of *vestibular-forward* condition show shorter reported eye-to-target distances (blue squares) than in the baseline-stationary (red disc) condition (n=8). Error bars represent standard errors of the mean.

Figure 5. Experiment 5: Predictions (**a & b**) and average results (**c**). **a.** *Stepladder* condition where the observer descends from the stepladder in the dark. If the path-integration process only gauges the horizontally traveled distance, the intrinsic bias would be path integrated in the horizontal but

not vertical direction. As such, the horizontal coordinate of the intrinsic bias (blue curve) is kept at the home base location on the floor while the vertical coordinate travels along with the observer. Thus, upon stepping down from the stepladder and standing on the floor, the observer underestimates the horizontal distance of the target. The target underestimation would be similar to that of the *horizontal-walking* condition depicted in **b**. In contrast, if the path-integration process integrates distance in both the horizontal and vertical directions, the intrinsic bias will remain at the top of the staircase (orange curve). The perceived target location will then be dramatically different from that in the *horizontal-walking* condition. **b.** *Horizontal-walking* condition. **c**. The average results of the *baseline-stationary*, *horizontal-walking* and *stepladder* conditions, respectively represented by the red circle, green square and blue triangle symbols (n=9). The judged horizontal distances are significantly shorter in the *horizontal-walking* condition than in the *baseline-stationary* conditions. The two sets of data points are fitted by the same intrinsic bias profile with a horizontal separation of 1.0 m, which is close to the walked distance (1.06 m) from the home base to the new location in the *horizontal-walking* condition. Of significance, the judged locations in the *stepladder* condition (blue triangles) are similar to that in the *horizontal-walking* condition (green squares). This confirms that the path-integration process mainly gauges the horizontal (ground) distance travelled.

Supplementary information is available for this paper.

Supplementary Information

Statistical Analysis of Data

<u>Experiment 1 (figure 3c)</u>

1. Baseline-stationary condition vs. walking condition with a 12 sec waiting period

We applied 2-way ANOVA with repeated measures (2 test conditions x 6 horizontal distances) to the walked horizontal distance data. The analysis reveals:

- Main effect of test condition (*baseline-stationary vs. walking*): $F(1,7)=439.133, p<0.001$
- Main effect of horizontal distance: $F(5, 35)=158.676, p<0.001$
- Interaction effect: $F(5, 35)=13.255, p<0.001$

2. Baseline-stationary condition vs. walking condition with a 60 sec waiting period

We applied 2-way ANOVA with repeated measures (2 test conditions x 6 horizontal distances) to the walked horizontal distance data. The analysis reveals:

- Main effect of test condition (*baseline-stationary vs. walking)*: $F(1,7)=13.406, p<0.01$
- Main effect of horizontal distance: $F(5, 35)=133.358, p<0.001$
- Interaction effect: $F(5, 35)=1.555, p=0.198$

<u>Experiment 2 (figure 3d)</u>

Baseline-stationary condition vs. divided attention walking condition

We applied 2-way ANOVA with repeated measures (2 test conditions x 6 horizontal distances) to the walked horizontal distance data. The analysis reveals:

- Main effect of test condition (*baseline-stationary vs. divided attention walking*): $F(1,7)=8.529, p=0.022$
- Main effect of horizontal distance: $F(5, 35)=63.780, p<0.001$
- Interaction effect: $F(5, 35)=1.072, p=0.392$

<u>Experiment 3 (figure 4)</u>

1. Vestibular-forward condition vs. walking condition (figure 4a)

We applied 2-way ANOVA with repeated measures (2 test conditions x 6 horizontal distances) to the walked horizontal distance data. The analysis reveals:

- Main effect of test condition (*vestibular-forward vs. walking)*: $F(1,7)=0.142, p=0.718$
- Main effect of horizontal distance: $F(5, 35)=99.788, p<0.001$
- Interaction effect: $F(5, 35)=0.714, p=0.617$

2. Vestibular-forward condition vs. baseline-stationary condition (figure 4a)

We applied 2-way ANOVA with repeated measures (2 test conditions x 6 horizontal distances) to the walked horizontal distance data. The analysis reveals:

- Main effect of test condition (*baseline-stationary vs. vestibular-forward)*: $F(1,7)=598.551, p<0.001$
- Main effect of horizontal distance: $F(5, 35)=168.855, p<0.001$
- Interaction effect: $F(5, 35)=19.422, p<0.001$

3. Baseline-stationary condition vs. walking condition (figure 4a)

We applied 2-way ANOVA with repeated measures (2 test conditions x 6 horizontal distances) to the walked horizontal distance data. The analysis reveals:

- Main effect of test condition (*baseline-stationary vs. walking*): $F(1,7)=299.648, p<0.001$
- Main effect of horizontal distance: $F(5, 35)=160.078, p<0.001$
- Interaction effect: $F(5, 35)=12.607, p<0.001$

4. Vestibular-backward condition vs. baseline-stationary condition (figure 4b)

We applied 2-way ANOVA with repeated measures (2 test conditions x 6 horizontal distances) to the walked horizontal distance data. The analysis reveals:

- Main effect of test condition (*vestibular-backward vs. baseline-stationary)*: $F(1,7)=651.863, p<0.001$
- Main effect of horizontal distance: $F(5, 35)=51.558, p<0.001$
- Interaction effect: $F(5, 35)=10.848, p<0.001$

<u>Experiment 4 (figure 4c)</u>

Vestibular-forward condition vs. baseline-stationary condition (verbal reports of eye-to-target distances)

We applied 2-way ANOVA with repeated measures (2 test conditions x 5 horizontal distances) to the verbally reported distance data. The analysis reveals:

- Main effect of test condition (*vestibular-forward vs. baseline-stationary)*: $F(1,7)=217.749, p<0.001$
- Main effect of distance: $F(4, 28)=36.945, p<0.001$
- Interaction effect: $F(4, 28)=8.973, p<0.001$

<u>Experiment 5 (figure 5c)</u>

1. Baseline-stationary condition vs. stepladder condition

<u>Horizontal distances</u>:

We applied 2-way ANOVA with repeated measures (2 test conditions x 5 horizontal distances) to the walked horizontal distance data. The analysis reveals:

- Main effect of test condition (*baseline-stationary vs. stepladder*): $F(1,8)=153.541, p=0.000$
- Main effect of horizontal distance: $F(4, 32)=35.954, p=0.000$
- Interaction effect: $F(2.58, 20.69)= 3.89, p=0.028$

<u>Heights</u>:

We applied 2-way ANOVA with repeated measures (2 test conditions x 5 heights) to the judged height data. The analysis reveals:

- Main effect of test condition (*baseline-stationary vs stepladder*): $F(1,8)=0.112, p=0.747$
- Main effect of height: $F(4, 32)=181.666, p=0.000$
- Interaction effect: $F(4, 32)= 0.288, p=0.883$

2. Baseline-stationary condition vs. horizontal-walking condition

<u>Horizontal distances</u>:

We applied 2-way ANOVA with repeated measures (2 test conditions x 5 horizontal distances) to the walked horizontal distance data. The analysis reveals:

- Main effect of test condition (*baseline-stationary vs. horizontal-walking*): $F(1,8)=104.198, p=0.000$

- Main effect of horizontal distance: $F(4, 32)=45.768, p=0.000$
- Interaction effect: $F(1.892, 15.137)= 8.461, p=0.004$

Heights:

We applied 2-way ANOVA with repeated measures (2 test conditions x 5 heights) to the judged height data. The analysis reveals:

- Main effect of test condition (*baseline-stationary vs. horizontal-walking*): $F(1,8)=7.695, p=0.024$
- Main effect of height distance: $F(4, 32)=236.530, p=0.000$
- Interaction effect: $F(2.636, 21.088)= 2.472, p=0.096$

3. Stepladder condition vs. horizontal-walking condition

Horizontal distances:

We applied 2-way ANOVA with repeated measures (2 test conditions x 5 horizontal distances) to the walked horizontal distance data. The analysis reveals:

- Main effect of test condition (*stepladder vs. horizontal-walking*): $F(1,8)=0.532, p=0.487$
- Main effect of horizontal distance: $F(4, 32)=32.372, p=0.000$
- Interaction effect: $F(2.967, 23.737)=0.874, p=0.467$

Heights:

We applied 2-way ANOVA with repeated measures (2 test conditions x 5 heights) to the judged height data. The analysis reveals:

- Main effect of test condition (*stepladder vs. horizontal-walking*): $F(1,8)=0.864, p=0.380$
- Main effect of height: $F(4, 32)=185.248, p=0.000$
- Interaction effect: $F(2.864, 22.908)=0.829, p=0.487$