nature human behaviour

LETTERS

PUBLISHED: 19 DECEMBER 2016 | VOLUME: 1 | ARTICLE NUMBER: 0004

# Attention modulates perception of visual space

Liu Zhou[1], Chenglong Deng[1], Teng Leng Ooi[2]* and Zijiang J. He[1,3,4]*

**Attention readily facilitates the detection and discrimination of objects, but it is not known whether it helps to form the vast volume of visual space that contains the objects and where actions are implemented. Conventional wisdom suggests not, given the effortless ease with which we perceive three-dimensional (3D) scenes on opening our eyes. Here, we show evidence to the contrary. In Experiment 1, the observer judged the location of a briefly presented target, placed either on the textured ground or ceiling surface. Judged location was more accurate for a target on the ground, provided that the ground was visible and that the observer directed attention to the lower visual field, not the upper field. This reveals that attention facilitates space perception with reference to the ground. Experiment 2 showed that judged location of a target in mid-air, with both ground and ceiling surfaces present, was more accurate when the observer directed their attention to the lower visual field; this indicates that the attention effect extends to visual space above the ground. These findings underscore the role of attention in anchoring visual orientation in space, which is arguably a primal event that enhances one's ability to interact with objects and surface layouts within the visual space. The fact that the effect of attention was contingent on the ground being visible suggests that our terrestrial visual system is best served by its ecological niche.**

Our visual space is bounded by large background surfaces that comprise the ground, sky/ceiling and horizon/wall surfaces. Evolution has attributed a special importance to the ground surface in determining our visual space[1–5], possibly due to our dependence on the ground for support. Thus, the visual system can efficiently represent the ground surface using external depth information and/or its intrinsic spatial knowledge of our ecological niche. Referred to as the intrinsic bias, intrinsic spatial knowledge is best revealed in the dark, where a dimly lit target on the ground is perceived to be located along its projection line from the eyes of the observer, but raised and closer to the observer. Successive measurements of the perceived locations of the dimly lit target at various distances on the ground reveal that the perceived locations fall along an imaginary slanted surface that is the intrinsic bias (Fig. 1a)[5–9]. The interplay between the intrinsic bias and visual cues that represent the ground surface is revealed in a reduced cue environment, where the ground surface is sparsely delineated by an array of dimly lit texture elements (Fig. 1b). This leads to a target on the ground being perceived at a location that is slightly further away than that in the dark, along an implicit slanted surface that is less slanted than the intrinsic bias's surface. Essentially, the visible ground surface helps expand the volume of visual space from relative compression[9–10].

We hypothesized that if attention to the lower visual field affects the representation of the ground surface, it should consequently influence the perceived locations of targets that utilize ground representation as a reference frame. To investigate this, we used Posner's attention paradigm[11] to direct attention to, or away from, the ground surface and conducted space perception testing. We predicted that when the ground surface was weakly delineated by texture elements and the observer directed attention to the lower field (attention lower, Fig. 2a), the perceived target location would be more accurate, compared with when the observer directed attention to the upper field (attention upper, Fig. 2b). We also tested the complementary scenario, where the visible surface was the ceiling rather than the ground. However, we did not predict a strong effect of attention, owing to the lower field bias of attention[12] and/or insufficient visual representation of the ceiling surface in the upper field[13–15].

We tested the predictions in a test block that comprises three different stimulus background conditions: ground texture (Fig. 2c), ceiling texture (Fig. 2d) and no texture (dark). Each test trial began with the observer binocularly fixating a small light-emitting diode (LED) at eye level (1.5 m) in the dark. To ensure stable fixation, the experimenter randomly flickered the LED and required the observer to verbally indicate whenever the flicker was detected. (Please refer to our control experiment in the Supplementary Information, which shows that the ability to detect a flickering target was contingent on the target not deviating from the fixation point by more than 1.5°.) Meanwhile, the observer was also instructed to direct their attention to either the lower (attention lower) or upper (attention upper) field. The fixation LED was presented for 4–6 s before being switched off. This was followed (interstimulus interval (ISI) = 0 s) by the presentation of one of the three background conditions for 0.15 s. Then, a test LED target (1 s) was presented (ISI = 0 s) at a predetermined location in the upper or lower field. The target was coplanar with either the ceiling- or ground-texture surface, while its distance varied among trials. The observer judged the target location and responded by walking blindly to the judged location[5,16–19]. On reaching the destination, the observer gestured the remembered target height. The walked distance ($x$) and gestured height ($y$) defined the judged target location.

Figure 3a plots the physical target locations (+ symbols) and average judged locations (coloured symbols). For targets below eye level, judged locations were closer to the physical locations (more accurate) in the attention-lower compared with the attention-upper paradigm, when the ground-texture background was present (filled and open green triangles, respectively). Notably, the profile of the attention-lower data extends forward and is less slanted than that of the attention-upper data. This increased accuracy was not due to the effect of attention on target detection alone, based on results from the ceiling-texture and dark conditions, where the observer directed their attention to the lower field, but there was no ground texture.

[1]Key Laboratory of Brain Functional Genomics (MOE & STCSM), Institute of Cognitive Neuroscience, School of Psychology and Cognitive Science, East China Normal University, Shanghai 200062, China. [2]College of Optometry, The Ohio State University, Columbus, Ohio 43210, USA. [3]Department of Psychological and Brain Sciences, University of Louisville, Louisville, Kentucky 40292, USA. [4]CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China. *e-mail: ooi.22@osu.edu; zjhe@louisville.edu

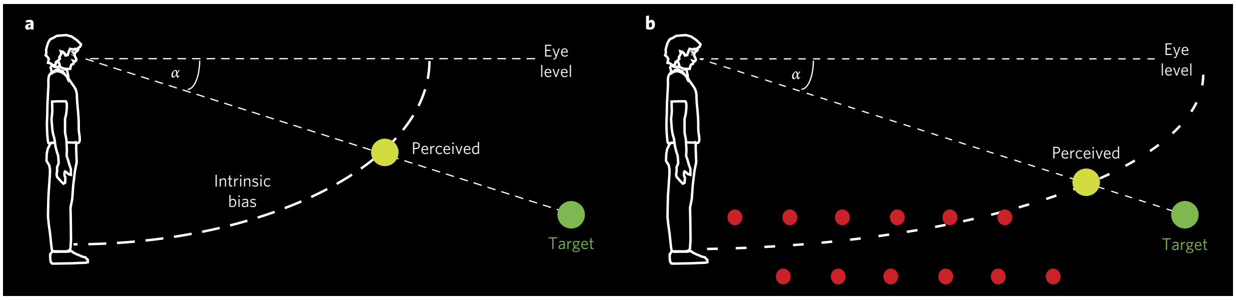


**Figure 1 | Ground-based space perception. a**, Perceptual space in the dark is compressed. Consequently, a dimly lit target (green circle) is perceived (yellow circle) to be located at the intersection between the projection line from the eyes and the intrinsic bias (dashed curve). **b**, Perceptual space expands when the ground is delineated by dimly lit texture elements (red circles). This leads to the target being perceived at the intersection of the projection line from the eyes and the ground surface representation (dashed curve), the slant of which is lesser than the slant of the intrinsic bias depicted in **a**. α refers to the angular declination of the target below the eye level.

In fact, the judged locations of targets below eye level in the ceiling-texture (green diamonds) and dark (red triangles) conditions were similar to those made when the observers directed attention to the upper field (open green triangles) in the presence of ground texture. This similarity indicates that ground texture is not well represented without attention being drawn to it. These findings support the prediction (Figs 2a,b) that attention directed to ground texture leads to more accurate target localization.

For the targets above eye level, judged locations in the presence of ceiling texture were not better when attention was paid to the upper field (filled blue triangles) compared with the lower field (open blue triangles). This suggests that attention has little effect on the representation of ceiling-texture surface. (Increasing the ceiling-texture display duration from 150 to 500 ms did not significantly improve accuracy in the attention-upper paradigm.) Furthermore, the findings are similar to the ground-texture (blue diamonds) and dark (red triangles) conditions when the observer directed their attention to the upper field.

We derived the judged eye-to-target distances using the formula $\sqrt{x^2+(y-h)^2}$, where $h$ is the observer's eye height, and $x$ and $y$ represent the judged location. For targets below eye level (Fig. 3b), judged eye-to-target distance was longer in the attention-lower (filled green triangles) than the attention-upper (open green triangles) paradigm, when ground texture was present (main effect of attention: $F(1,5)=85.43$, $P<0.001$; main effect of distance: $F(2,10)=42.15$, $P<0.001$; interaction: $F(2,10)=13.72$, $P<0.001$). There were no significant differences among the data obtained from three conditions: (1) the ceiling-texture condition (green diamonds) when attention was directed to the lower field; (2) the dark condition (filled red triangles) when attention was directed to the lower field; and (3) the ground-texture condition (open green triangles) when attention was directed to the upper field (main effect of test condition: $F(2,10)=2.08$, $P=0.175$; main effect of distance: $F(2,10)=36.35$, $P<0.001$; interaction: $F(4,20)=1.42$, $P=0.262$). Most notably, no significant differences between the attention-lower dark condition (filled red triangles) and attention-upper ground-texture condition (open green triangles) paradigms suggests that without attention being directed to the texture on the ground, the ground texture itself contributes little to space perception. It is as if the texture on the ground becomes invisible without attention being paid to it. This finding underscores the critical role of visual attention in space perception.

For targets above eye level (Fig. 3c), there was no significant difference between directing attention to the upper (filled blue triangles) or lower (open blue triangles) field, under the ceiling-texture condition (main effect of attention: $F(1,5)=1.57$, $P=0.265$; main effect of distance: $F(2,10)=15.68$, $P<0.005$; interaction: $F(2,10)=0.48$, $P=0.633$). There were significant differences among three conditions: (1) the ground-texture condition (blue diamonds) when attention was directed to the upper field; (2) the dark condition (red triangles) when attention was directed to the upper field; (3) the ceiling-texture condition (open blue triangle) when attention was directed to the lower field (main effect of test condition: $F(2,10)=5.30$, $P<0.05$; main effect of distance: $F(2,10)=16.54$, $P<0.001$; interaction: $F(4,20)=0.80$, $P=0.542$). Pairwise comparisons showed that only the ground-texture and dark conditions were significantly different ($P<0.005$).

Since ground representation is the reference frame for locating a target in mid-air[15,19–21], attention should affect spatial perception above the ground surface, that is, the volume of perceptual space extending from the observer's eyes to the ground surface representation. To investigate this, we modified the stimuli in Experiment 1

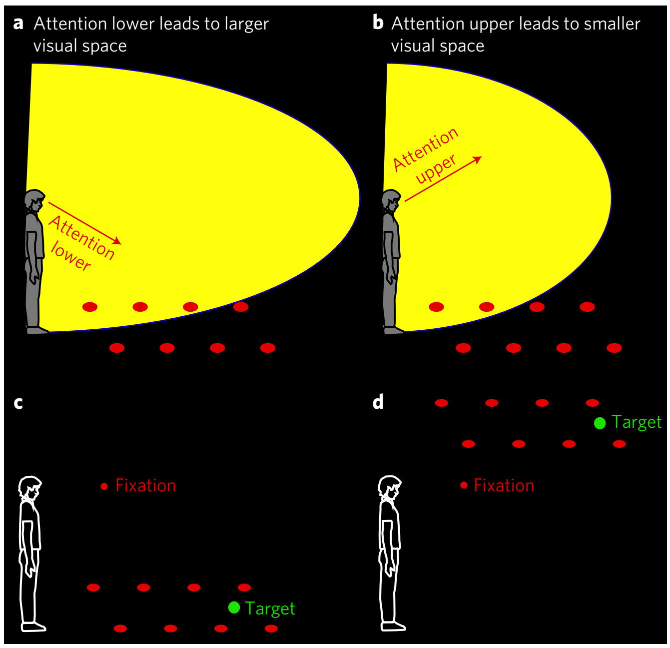


**Figure 2 | Predictions and experimental design. a**, When attention is directed to the lower visual field in the presence of texture information delineating the ground, the visual system represents the ground surface more accurately, leading to a larger expanse of visual perceptual space (yellow area). **b**, If attention is directed to the upper visual field, the perceptual space formed is compressed (yellow area). This leads to the prediction that judged target location will be farther (closer to the target) in **a** than in **b**. **c**,**d**, The ground-texture (**c**) and ceiling-texture (**d**) stimulus conditions used in Experiment 1.

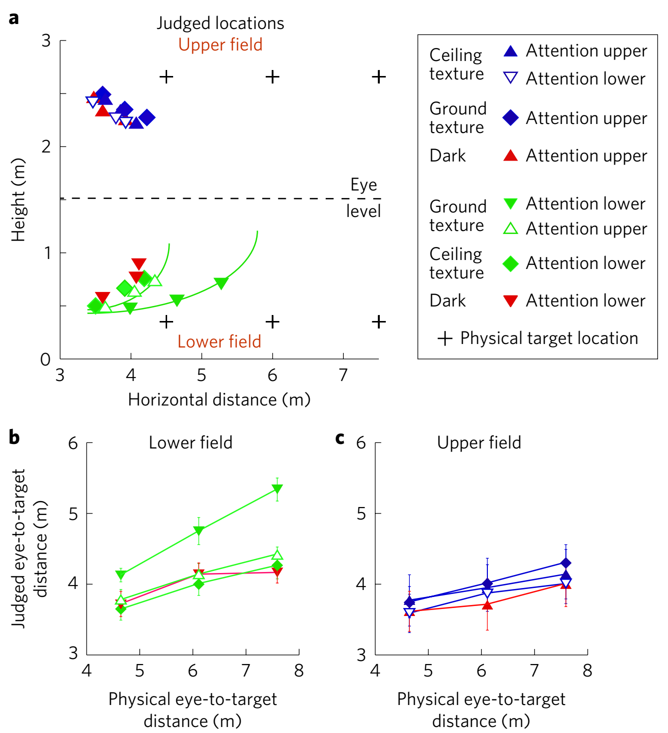


**Figure 3 | Attention effect on judged target location. a**, The physical (black + symbols) and judged (coloured symbols) target locations. Observers judged a target location in the lower field as farther when they directed attention to the ground texture (filled green triangles) compared with the empty space in the upper field (open green triangles). There was no attention effect in the dark condition (red triangles), the ceiling-texture condition (green diamonds) or when the target was in the upper field. **b**,**c**, Averaged eye-to-target distances derived from the judged locations in **a** for targets in the lower (**b**) and upper (**c**) fields. The error bars represent standard error of the mean.

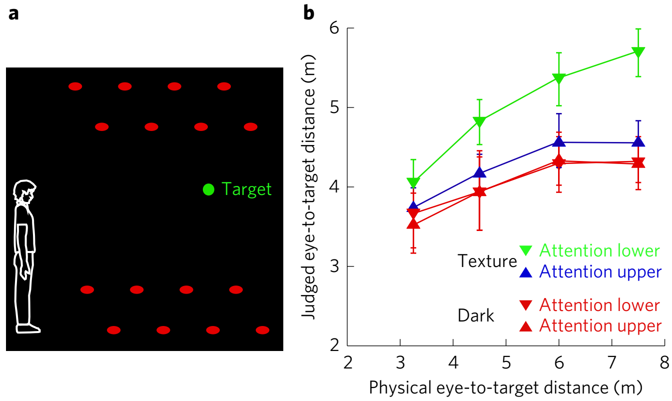


**Figure 4 | Attention effect on judged location of a mid-air target in Experiment 2.** (**a**) Illustration of the ceiling and ground texture displays used. (**b**) The perceived eye-to-target distance is derived from the judged target location. In the presence of the texture displays, perceived distance was further away (closer to the target) when attention was directed to the lower field (green triangles) compared with the upper field (blue triangles). When the texture displays were removed (dark condition), the attention effect was not found (red triangles). The error bars represent standard error of the mean.

by placing the target at eye level instead of having it coplanar with the textured surface and by presenting both the ground-texture and ceiling-texture backgrounds simultaneously (Experiment 2, Fig. 4a). Figure 4b shows the average judged eye-to-target distance. As predicted, observers judged the target as farther away (more accurate) when they directed attention to the lower (green triangles) compared with the upper (blue triangles) field (main effect of attention direction: $F(1,5)=119.44$, $P<0.001$; main effect of distance: $F(3,15)=23.91$, $P<0.001$; interaction: $F(3,15)=23.32$, $P=0.001$). The attention effect is tied to the textured ground surface, because there was no significant difference between directing attention to the lower (red inverted triangles) or upper (red upright triangles) field under the dark condition (main effect of attention direction: $F(1,5)=0.001$, $P=0.973$; main effect of distance: $F(1.43,7.13)=10.38$, $P<0.05$; interaction: $F(3,15)=0.38$, $P=0.77$). Also, judged distance was slightly longer, though not significantly, when attention was directed to the upper field under the texture condition (blue triangles) compared with the dark condition (red upright triangles) (main effect of texture: $F(1,5)=2.27$, $P=0.192$; main effect of distance: $F(3,15)=10.92$, $P<0.001$; interaction: $F(3,15)=0.26$, $P=0.994$). This again suggests the visual system does not represent a textured surface well without attention being directed to it.

Overall, our experiments indicate that attention to the visible ground is vital for forming accurate visual space. Our study is distinct from previous studies that focused on the role of attention in representing objects embedded in structured backgrounds[22–24]; these studies did not address whether attention affected the construction of the 3D visual space wherein the target and structured background were located. Instead, they assumed visual space was pre-attentively established and the role of attention was to select the target within the visual space. This assumption is largely reasonable, since the stimuli in a typical experiment were presented on a two-dimensional (2D) computer screen. The observer would have ample time to form the visual space encompassing the computer screen before the stimulus onset, making consideration about 3D visual space formation moot. Therefore, our study is also distinct in that the attention mechanism responsible for visual space operates over a large spatial scale, rather than on a local target(s). That is, the attention process tested in our study here is more 'ambient'. We previously showed that being able to access a larger surface area leads to more accurate ground surface representation and more reliable space perception[25–26]. Such an ambient attention process is distinct from focal attention, in that focal attention operates on isolated objects and mainly affects the relative spatial relationship between targets[27].

Studies of real-world scenes displayed on a 2D computer screen (not real, 3D space) often require the observers to extract the global scene properties, such as its spatial layout (for example, qualitative depth) and meaning (for example, an urban versus natural scene)[28–31]. For instance, on presentation of a single picture of a real-world scene on a 2D computer screen, it was found that observers could quickly determine the scene's category (for example, qualitative mean depth and navigability)[30] within a stimulus duration as short as 19 ms. This kind of observation might lead one to assume that attention is not required for capturing the global features of a scene. The assumption, however, is inconsistent with the finding that attention is required to search for a 2D picture of a real-world scene with a unique global scene property among other distracting scenes without the unique property[32].

On the other hand, the ambient attention revealed in our study plays a critical role in forming 3D perceptual space to facilitate object localization and direct actions in the real world. We speculate that the role of ambient attention could also be extended to perception of real 3D objects, since perceived shapes and sizes are scaled with perceptual distance. Our findings also allow us to speculate about how humans can reliably and effortlessly perceive the vast visual space, despite the huge amount of information contained in natural scenes. First, having adapted to the terrestrial environment, our brain is 'hard-wired' to efficiently construct the ground representation for use as a scaffold for the entire visual space. Second, we habitually orient to the lower field because our gaze tends to fixate on, or near, the ground surface, where objects of interest are frequently encountered. As visual attention is often directed to, or near, the ground surface, the process of deploying attention to the ground surface becomes more efficient through experience. Indeed, we revealed a stronger attention effect on the ground than on the ceiling texture surface. Both factors contribute to the rapid formation of a reliable visual space, providing the context for us to efficiently pay attention to objects and to direct and guide our actions in the terrestrial environment[22,33–34].

## Methods

**Observers.** Six paid observers with informed consent (age = 23.33 ± 1.46 years old; eye height = 1.50 ± 0.02 m; 3 males and 3 females) who were naïve to the purpose of the study participated in both experiments. (As such, the experimenters were not blinded to subject group allocation.) The observers were either undergraduate or graduate students at the East China Normal University (ECNU) who responded to subject recruitment flyers that were posted on campus. All observers had normal, or corrected-to-normal, visual acuity (at least 20/20) and a stereoscopic resolution of a 20 arcsec or better. They viewed the visual scene binocularly. The observer sample size follows the number of subjects used by similar studies in the field. A within-subject experimental design was used. The study protocol was approved by the ECNU Institutional Review Board.

**Experiment 1.** *Design.* In Experiment 1 the target was positioned coplanar with the ground-or ceiling-textured surface. The three conditions tested were the ground-texture, ceiling-texture and dark conditions. For standardization among observers, the background texture surface delineating the ground was placed 0.35 m above the floor and the background surface delineating the ceiling was placed at twice the observer's eye height minus 0.35 m. The test target was placed at one of three distances (4.5, 6 and 7.5 m) either in the lower visual field (height = 0.35 m above the floor) or the upper visual field (height = (2 × observer's eye height) − 0.35 m).

During a trial, the test target was presented in either the upper or lower visual field, in the dark or in the presence of either the ceiling texture or ground texture. In addition, the observer was required to direct attention to either the upper or lower visual field. In all, there were 24 stimulus combinations, comprising 12 for testing the target in the lower field (3 target locations × [dark/attention lower + ceiling texture/attention lower + ground texture/attention upper + ground texture/attention lower]) and 12 for testing the target in the upper field (3 target locations × [dark/attention upper + ground texture/attention upper + ceiling texture/attention upper + ceiling texture/attention lower]).

Eight catch trials were added, comprising targets at 3.25 m with or without the background (4 trials) and targets at 6 and 7 m with the background presented for 500 ms (instead of 150 ms, as in the test trials) (4 trials). Attention was always deployed to the field of the upcoming catch-trial target. The target heights for the catch trials were the same as in the test trials. The purpose of including catch trials in the experiment was to prevent the observer from simply expecting that the targets would only be presented within a narrow range of distance and stimulus duration. Altogether, a total of 128 trials (24 test and 8 catch trials with 4 repeats) were run over four days (sessions). The order of stimulus presentation was randomized, with the first and fourth sessions adopting the same randomization sequence and the second and third sessions having the reversed randomization sequence.

*Stimuli and test environment.* The experiment was performed in a dark room (8 m × 13 m), of which the layout and dimensions were unknown to the observers. Its ceiling (3.1 m high) and walls were painted black and the floor was covered with a black carpet. A 15 m-long rope (0.8 m above the floor) tied to both ends of the room was used to guide the observer while blind walking. One end of the room (1.2 × 2.5 m) served as a waiting area for the observer between trials. The waiting area had a chair for the observer to sit on facing the wall (opposite the test area) and a tungsten lamp for illumination.

A 0.19° red fixation target constructed from a dimly lit LED (0.08 cd m$^{-2}$) housed inside a ping-pong ball was located at the observer's eye level from a viewing distance of 1.5 m. To ensure steady fixation while directing attention to either the upper or lower field, the fixation LED was randomly turned off for 0.1 s and the observer was required to correctly indicate that it was turned off. The random turning off of the fixation target occurred between 1–3 times within a 4–6 s fixation period. If the observer failed to report that the fixation was briefly turned off, the trial would be terminated and repeated at the end of the test block. All observers were able to maintain fixation and successfully reported the fixation turn-off at an average rate of 95.2% ± 1.2%. In addition, if an observer felt their eye moved during a trial, they would indicate this and the trial would be abandoned; this seldom occurred (2.2% ± 1.2%).

Immediately following the fixation period (ISI = 0 s), a textured background comprising a 2 × 4 array of dimly lit red LED elements (0.08 cd m$^{-2}$) spanning an area of 1.4 × 3 m was presented in either the upper or lower field for 0.15 s. Each LED element was housed inside a ping-pong ball with a 2.5 cm (diameter) circular opening. The LED elements to delineate the background were horizontally separated by 1.4 m and placed at 1, 2, 3 and 4 m from the observer. The height of the LED elements was 0.35 m above the floor in the lower field or at twice the observer's eye height minus 0.35 m in the upper field. If no texture background was presented, owing to the trial being conducted in the dark, a darkness period of 0.15 s ensued.

A 5 Hz flickering test target was presented for 1 s immediately following the textured background or darkness presentation (ISI = 0 s). The test target was constructed from a ping-pong ball that was internally illuminated by a green LED (0.16 cd m$^{-2}$) and controlled by a computer. An iris-diaphragm aperture in the front of the ping-pong ball kept its visual angle at 0.19° when measured at eye level.

*Procedure.* At the start of each experimental session, the blindfolded observer was led to the illuminated waiting area (1.2 × 2.5 m) at one end of the test room. The observer then sat on the chair facing the wall (opposite the test area) and removed their blindfold to wait for the start of a trial.

Before each trial, the observer was told to either pay attention to the upper or lower visual field. The experimenter then presented a computer-generated tone to signal the observer to turn off the tungsten lamp in the waiting area, which rendered the test room completely dark. Then, the observer walked to the starting point (the observation position was marked by fluorescent elements on the floor) with the aid of the guidance rope, stood upright and called out, "ready". The trial started after a 5 s delay. The sequence of the trial was: (i) fixation to maintain the observer's eye alignment while directing attention to either the upper or lower field and reporting whenever the fixation flickered off; (ii) textured background or darkness presentation; (iii) target presentation and observer judgement of location. Once the observer was satisfied with their judgment of the target location, they closed their eyes, put on the blindfold and called out "ready to go". Upon which, the experimenter immediately removed the target and shook the guidance rope to indicate the test area was safe for walking. The observer then walked to the remembered target location while sliding one hand along the guidance rope. When the observer reached the remembered location, they stopped and indicated the target height with the tip of a 1-m rod, and reported, "done". The experimenter marked the location of the observer's feet and measured the height of the 1-m rod tip. These measurements were recorded as the judged target location. The observer was then instructed to turn around and walked back to the waiting area, where they removed the blindfold, turned on the lamp, sat on the chair and waited for the next trial. In the meantime, the experimenter prepared the next trial.

The observers were given five practice trials before each experimental session. No feedback regarding their performance was given. During the experiment, music was played to mask possible acoustic cues from revealing the target location.

**Experiment 2.** *Design.* In Experiment 2 the target was positioned at eye level. All aspects of the experiment remained the same as those of Experiment 1, except for the target location and the presentation mode of the textured surfaces. The target was located at the observer's eye level and at one of four distances (3.25, 4.5, 6 and 7.5 m). The target was presented after the ground and ceiling textured surfaces were presented simultaneously (in contrast to Experiment 1, where only one of these surfaces was presented). As in Experiment 1, trials were also conducted under the dark condition. The observers were required to deploy their attention to either the upper or lower visual field. The trials with textured surfaces were repeated three times for each target location, while the trials in the dark were tested once for each target location. The total, 32 trials, were tested in one session.

Similar to Experiment 1, the observers were able to successfully report the turning-off of the fixation target during the fixation period (average success rate = 99.0% ± 0.7%). The average incidence of the observers reporting that they felt their eye moved was also low (3.7% ± 2.0%).

*Data analyses.* Data were analysed using analysis of variance with repeated measures. The Mauchly's test was applied to verify the assumption of sphericity.

**Data availability.** The data sets generated are available from the corresponding authors upon reasonable request.

Received 6 April 2016; accepted 10 June 2016; published 19 December 2016

## References


1. Alhazen, I. *The Optics of Ibn-Haytham* Vol. 1, 1039–1989 (Warburg Institute, 1989).
2. Gibson, J. J. *The Perception of the Visual World* (Houghton Mifflin, 1950).
3. Sedgwick, H. A. in *Handbook of Perception and Human Performance* (eds Boff, K. R. *et al.*) 21.1–21.57 (Wiley, 1986).

4. Sinai, M. J., Ooi, T. L. & He, Z. J. Terrain influences the accurate judgement of distance. *Nature* **395,** 497–500 (1998).
5. Ooi, T. L., Wu, B. & He, Z. J. Distance determined by the angular declination below the horizon. *Nature* **414,** 197–200 (2001).
6. 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).
7. Aznar-Casanova, A., Keil, M. S., Moreno, M. & Supèr, H. Differential intrinsic bias of the 3-D perceptual environment and its role in shape constancy. *Exp. Brain. Res.* **215,** 35–43 (2011).
8. Zhou, L., He, Z. J. & Ooi, T. L. The visual system's intrinsic bias and knowledge of size mediate perceived size and location in the dark. *J. Exp. Psychol. Learn. Mem. Cogn.* **39,** 1930–1942 (2013).
9. Wu, J., He, Z. J. & Ooi, T. L. The visual system's intrinsic bias influences space perception in the impoverished environment. *J. Exp. Psychol. Hum. Percept. Perform.* **40,** 626–638 (2014).
10. 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).
11. Posner, M. I. Orienting of attention. *Q. J. Exp. Psychol.* **32,** 3–25 (1980).
12. He, S., Cavanagh, P. & Intriligator, J. Attentional resolution and the locus of awareness. *Nature* **383,** 334–338 (1996).
13. McCarley, J. S. & He, Z. J. Asymmetry in 3-D perceptual organization: ground-like surface superior to ceiling-like surface. *Percept. Psychophys.* **62,** 540–549 (2000).
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. 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).
16. 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).
17. 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).
18. Loomis, J. M. & Philbeck, J. W. in *Embodiment, Ego-space, and Action* (eds Klatzky, R. L. *et al.*) 1–43 (Psychology Press, 2008).
19. Ooi, T. L. & He, Z. J. Space perception of strabismic observers in the real world environment. *Invest. Ophthalmol. Vis. Sci.* **56,** 1761–1768 (2015).
20. Meng, J. C. & Sedgwick, H. A. Distance perception mediated through nested contact relations among surfaces. *Percept. Psychophys.* **63,** 1–15 (2001).
21. Ni, R., Braunstein, M. L. & Andersen, G. J. Scene layout from ground contact, occlusion, and motion parallax. *Vis. Cogn.* **15,** 46–68 (2007).
22. Chun, M. M. & Wolfe, J. M. in *Blackwell Handbook of Sensation and Perception* (ed. Goldstein, E. B.) 272–310 (Blackwell, 2001).
23. Chun, M. M., Golomb, J. D. & Turk-Browne, N. B. A taxonomy of external and internal attention. *Annu. Rev. Psychol.* **62,** 73–101 (2011).
24. Treisman, A. & Gelade, G. A feature-integration theory of attention. *Cogn. Psychol.* **12,** 97–136 (1980).
25. Wu, B., Ooi, T. L. & He, Z. J. Perceiving distance accurately by a directional process of integrating ground information. *Nature* **428,** 73–77 (2004).
26. Wu, J., He, Z. J. & Ooi, T. L. Perceived relative distance on the ground affected by the selection of depth information. *Percept. Psychophys.* **70,** 707–713 (2008).
27. Gogel, W. C. & Sharkey, T. J. Measuring attention using induced motion. *Perception* **18,** 303–320 (1989).
28. Potter, M. C. Meaning in visual search. *Science* **187,** 965–966 (1975).
29. Greene, M. R. & Oliva, A. Recognition of natural scenes from global properties: seeing the forest without representing the trees. *Cogn. Psychol.* **58,** 137–176 (2009).
30. Greene, M. R. & Oliva, A. The briefest of glances: the time course of natural scene understanding. *Psychol. Sci.* **20,** 464–472 (2009).
31. Fei-Fei, L., Iyer, A., Koch, C. & Perona, P. What do we perceive in a glance of a real-world scene? *J. Vis.* **7,** 1–29 (2007).
32. Greene, M. R. & Wolfe & J. M. Global image properties do not guide visual search. *J. Vis.* **11,** 1–9 (2011).
33. Qiu, F. T., Sugihara, T. & von der Heydt, R. Figure-ground mechanisms provide structure for selective attention. *Nat. Neurosci.* **10,** 1492–1499 (2007).
34. Cavanagh, P., Hunt, A., Afraz, A. & Rolfs, M. Visual stability based on remapping of attention pointers. *Trends. Cogn. Sci.* **14,** 147–153 (2010).

## Acknowledgements
This study was supported by grants from the National Institutes of Health (EY023374 and EY023561) to T.L.O. and Z.J.H. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Author contributions
L.Z. conducted the experiments, performed the data analysis and wrote the paper. C.D. wrote the software for the experiments and participated in testing observers. Z.J.H. and T.L.O. provided the theoretical motivation, performed the data analysis and wrote the paper.

## Additional information
**Supplementary information** is available for this paper.

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

**Correspondence and requests for materials** should be addressed to T.L.O. or Z.J.H.

**How to cite this article:** Zhou, L., Deng, C., Ooi, T. L. & He, Z. J. Attention modulates perception of visual space. *Nat. Hum. Behav.* **1,** 0004 (2016).

## Competing interests
The authors declare no competing interests.