Shape judgments in natural scenes: Convexity biases versus stereopsis

Brittney Hartle Department of Psychology and Centre for Vision Research, York University, Toronto, Ontario, Canada

Aishwarya Sudhama-Joseph Department of Psychology and Centre for Vision Research, York University, Toronto, Ontario, Canada

Elizabeth L. Irving School of Optometry and Vision Science, University of Waterloo, Waterloo, Ontario, Canada

Robert S. Allison Department of Electrical Engineering and Computer Science, York University, Toronto, Ontario, Canada

Mackenzie G. Glaholt Defence Research and Development Canada, Toronto, Ontario, Canada

Laurie M. Wilcox Department of Psychology and Centre for Vision Research, York University, Toronto, Ontario, Canada

Determining the relief of upcoming terrain is critical to locomotion over rough or uneven ground. Given the significant contribution of stereopsis to perceived surface shape, it should play a crucial role in determining the shape of ground surfaces. The aim of this series of experiments was to evaluate the relative contribution of monocular and binocular depth cues to judgments of ground relief. To accomplish this goal, we simulated a depth discrimination task using naturalistic imagery. Stimuli consisted of a stereoscopically rendered grassy terrain with a central mound or a dip with varying height. We measured thresholds for discrimination of the direction of the depth offset. To determine the relationship between relief discrimination and measures of stereopsis, we used two stereoacuity tasks performed under the same viewing conditions. To assess the impact of ambiguous two-dimensional shading cues on depth judgments in our terrain task, we manipulated the intensity of the shading (low and high). Our results show that observers reliably discriminated ground reliefs as small as 20 cm at a viewing distance of 9.1 m. As the shading was intensified, a large proportion of observers (30%) exhibited a strong convexity bias, even when stereopsis indicated a concave depression. This finding suggests that there are significant individual differences in the reliance on assumptions of surface curvature that must be considered in experimental conditions.

In impoverished viewing environments with limiting depth cues, these convexity biases could persist in judgments of ground relief, especially when shading cues are highly salient.

Introduction

Everyday activities such as walking or interacting with objects entail the integration of a complex array of perceptual and motor information. This is especially true for tasks such as locomotion over irregular terrain that requires that observers assess the relief of upcoming paths and adjust their approach accordingly (Barton, Matthis, & Fajen, 2017; Zhao & Allison, 2021). Over rough terrains, observers increase their step planning margins by gazing multiple steps ahead to assess which footholds are traversable (Matthis, Yates, & Hayhoe, 2018). This type of locomotion task necessitates assessment of ground relief well ahead of the observer's steps and quick decision-making regarding safe traversal paths at distances well beyond interaction space (see also Allison, Gillam, & Palmisano, 2009).

Under natural viewing conditions, the perception of ground relief (or other surface shape) is based on the

Citation: Hartle, B., Sudhama-Joseph, A., Irving, E. L., Allison, R. S., Glaholt, M. G., & Wilcox, L. M. (2022). Shape judgments in natural scenes: Convexity biases versus stereopsis. *Journal of Vision*, *22*(8):6, 1–13, https://doi.org/10.1167/jov.22.8.6.

https://doi.org/10.1167/jov.22.8.6 Received March 14, 2022; published July 15, 2022 ISSN 1534-7362 Copyright 2022 The Authors

integration of monocular (e.g., texture and shading) and binocular cues. Although both types of cues provide information regarding the shape of objects, binocular depth cues such as stereopsis provide significant advantages given the visual system's sensitivity to disparities that indicate local changes in curvature (Howard, 2012). Stereoscopic shape perception takes advantage of second-order spatial derivatives that provide information regarding the relative depth and slant of surfaces (Norman et al., 1991; Lappin & Craft, 2000). Most assessments of the role of stereopsis on perceived surface shape are performed at short viewing distances within interaction space where these binocular cues are most precise (Blakemore, 1970; Gogel, 1977; Foley, 1985; Rogers & Bradshaw, 1993). However, stereopsis can support reliable depth and shape estimation at viewing distances well beyond 2 m (Allison, Gillam, & Vecellio, 2009; Palmisano et al., 2010). Given the significant contribution of stereopsis to judgments of perceived surface shape, it should play a crucial role in determining the shape of the ground surface for decision-making before locomotion on foot or in vehicles.

The aim of the current study was to evaluate the relative contribution of monocular and binocular depth cues to judgments of ground relief. To accomplish this goal, we simulated an ecologically relevant "real-world" depth discrimination task modelled after helicopter safe landing decisions using naturalistic imagery. Under typical landing conditions, to achieve a safe landing pilots and flight engineers must ensure that the slope of the ground between the skids of the helicopter is less than approximately 10° (Transport Canada, 2006). We hypothesized that stereopsis was likely to be important for the assessment of ground relief at large distances beyond interaction space, particularly for natural textures such as grass. We assessed the contribution of pictorial depth cues by testing observers both monocularly and binocularly. Further, to determine the relationship between the discrimination of naturalistic relief with traditional laboratory measures of stereopsis, we modified two stereoacuity tasks to assess both local and global stereopsis under the same viewing conditions at a screen distance of 6.1 m. Using the safe landing criteria as a guideline, under these viewing conditions if observers cannot detect a change in ground relief of at least 60 arcsecs (0.38 m), they would not be able to detect ground features that would jeopardize a safe landing.

Methods

Observers

A total of 45 York University students participated in the study. All observers had normal or corrected-to-normal vision. Before testing, informed consent and demographic forms were completed. The demographic form asked observers to indicate their age, sex, handedness, vision correction status, eye dominance, and experience with three-dimensional displays. Randot Stereotest and FLY Stereo Acuity test booklets were used as an initial assessment to determine if observers had stereoscopic vision. In addition, a cover test was conducted to assess binocular alignment of the eyes. When a deviation (strabismus or heterophoria) was detected a note was made on the demographic questionnaire. Two observers exhibited strabismus and were excluded from the study. The remaining 43 observers completed the experiment. The research protocol was approved by York University's Research Ethics Board.

Stimuli

Terrain task

The terrain discrimination task used stereoscopic images depicting a high-resolution grass texture with a mound or depression at the center. The viewing geometry simulated a viewpoint of the observer looking straight down to the ground plane. The position of the texture was jittered between each render so the relative position of the texture elements could not be used as a reference. All stereoscopic images (1920 × 1080) were rendered in Autodesk MAYA 2016. The 28° field of view of the virtual camera matched the visual angle of the display at the viewing distance of 6.1 m. The stereoscopic virtual camera configuration for rendering was set to a nominal interaxial distance of 6 cm. The zero parallax plane matched the distance of the projection screen so that features portrayed at the screen distance were presented without screen disparity (at the same pixel location in both the left and right eyes).

The shape of the feature in the ground terrain followed a radially symmetric Gaussian depth profile that was always rendered at the center of the display. The two-dimensional Gaussian is defined in Equation 1, where A is the amplitude (maximum depth), x_0 and y_0 is the center position, and σ is the standard deviation of 0.4 m. The function was imported into MAYA as a mesh and scaled to create raised or depressed surfaces with a diameter of 2.4 m (which corresponds with ±3 sigma or 99.7% of the height variation) and a range of peak feature heights in the ground terrain (minimum of 0.08 m to a maximum of 1.30 m for both positive and negative ground reliefs). Given the rendering interocular distance of 6 cm and ground plane distance of 9.1 m, the relative disparity between the ground plane and the feature peak ranged from 12 to 227 seconds of arc. The ground plane was presented at a fixed uncrossed

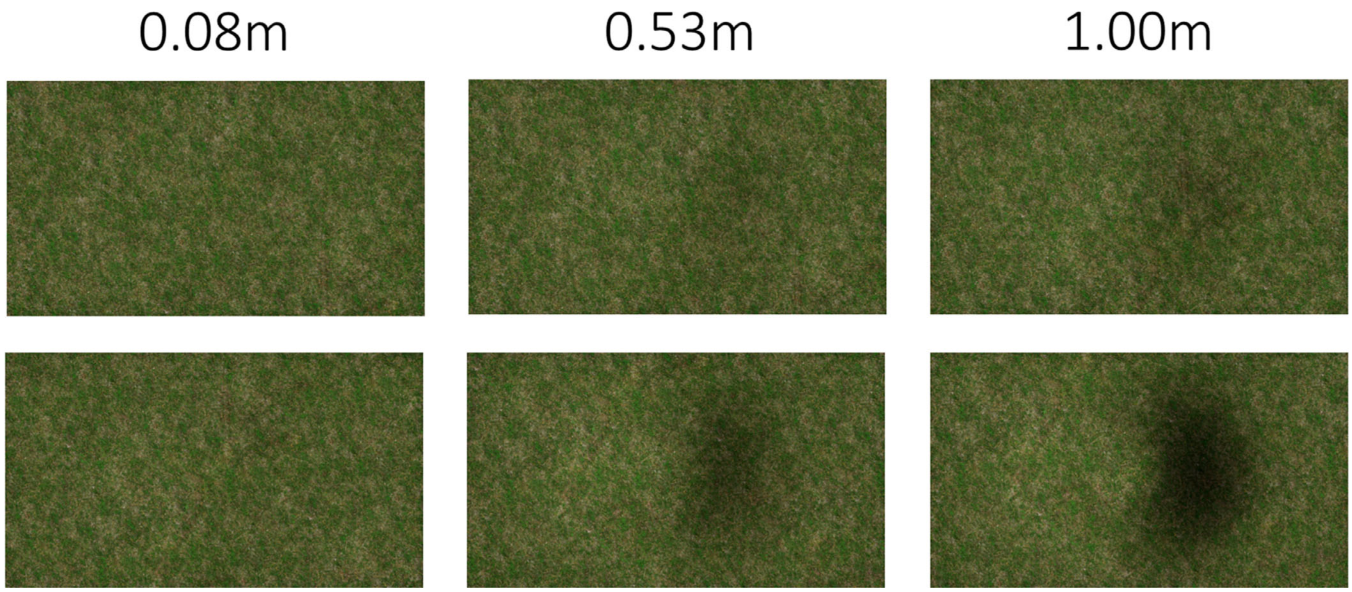


Figure 1. An illustration of the low shading (top) and high shading (bottom) conditions. The sample images above illustrate a 0.08, 0.53, and 1.00 m mound with light from the left. These feature heights represent the low, middle, and highest ground reliefs for an observer with a moderate step size of 0.23 m.

disparity relative to the screen of 0.19° so that it appeared to lie 3.0 m behind the screen plane.

$$f(x, y) = Ae^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\sigma^2}} \quad (1)$$

Each ground relief height was rendered four times to create the low and high shading conditions with each of two directions of lighting. To manipulate the shading intensity, the mean luminance of the frontoparallel ground plane in each condition was held constant, and the ratio between the intensity of directional and ambient light was varied. In the low shading condition, the ratio of directional to ambient light was 1:1, whereas in the high shading condition it was 9:1. The surfaces were not rendered with specular reflections, only diffuse surface reflections. The lighting direction was from the right or left side at an elevation angle of 45° relative to the flat ground plane. The lighting direction was varied in each stimulus condition such that one-half of the trials had lighting from the left and one-half from the right. Sample images for each shading condition are illustrated in Figure 1.

Stereoacuity tests

The two computer-based tasks used to assess each observer's stereoacuity were the 1) ledge and 2) bar tests. The ledge stimulus comprised a bipartite computer-generated random dot pattern with an abrupt disparity transition between the upper and lower areas. On each trial, 5,600 white dots each with a diameter of 0.07° filled the black display area (2.4 × 1.4 m) at a viewing distance of 6.1 m (Figure 2). The bar stimulus consisted of a white frame (0.9° × 1.19°) surrounding a white bar (0.42° × 0.72°) presented at the center of the screen, at a distance of 6.1 m. The width of the outer frame was 0.06° and the frame was separated from the bar by 0.48°. The bar in the center was displaced stereoscopically either in front of or behind the frame (Figure 3). All stereoacuity tests were presented on the same apparatus as the terrain task, but these stimuli only contained disparity cues to depth (see the Apparatus section).

Apparatus

Testing was conducted in an open studio space with the only lighting provided by the projector. Stimuli were back-projected on to a Stewart Film screen (3.0 × 1.7 m) using a Panasonic LCD projector PT-AE7000U (1920 × 1080). The projected image was 2.4 × 1.4 m. Observers wore active three-dimensional eyewear (TY-EW3D2MU) to view the three-dimensional imagery and made their responses using a Logitech F310 gamepad. Observers were seated at a viewing distance of 6.1 m from the screen on an elevated platform with their head at the same height as the center of the projection screen (Figure 4). All testing took place in a darkened room.

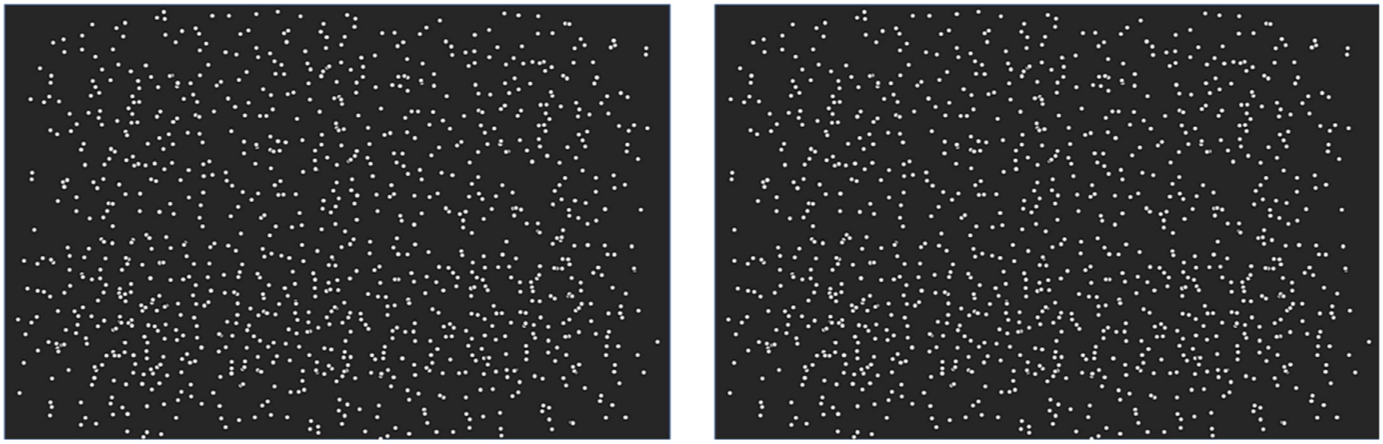


Figure 2. Stereopair illustrating the random dot stimuli used for the ledge task. When cross fused, the upper portion of the array will appear to be further away than the bottom portion.

Figure 3. Illustration of the stimulus used for the bar test. When cross fused, the central white bar will appear to be further away than the outer frame.

Procedure

For the terrain task, a forced-choice method of constant stimuli was used to measure discrimination thresholds. A total of 10 test peak depths of the mound or dip that bracketed zero (flat) were presented 20 times apiece in random order for each shading condition. The step size was either 0.15, 0.23, or 0.30 m and was selected for each observer based on the initial stereoacuity testing. All viewing conditions (shading, light direction, binocular and monocular) were interleaved and the order was randomized across observers. For monocular trials, instead of presenting the left and right eyes images to each eye (like in the binocular condition), the left eye image was presented to both eyes creating a flat zero disparity image at the screen distance. On each trial, observers were asked to indicate whether they saw a mound or dip in the center of the ground plane using the gamepad. Each stimulus was displayed until the observer gave their response, and between trials a white fixation marker was displayed on a black background at the screen plane. Each ground relief height was presented 20 times, for a total of 400 binocular trials (10 feature heights × 2 shading conditions × 2 lighting directions × 10 trials). In addition, 10 monocular trials were included for the largest relief height (0.69, 0.99, or 1.30 m based on the step size) and the smallest relief height (0.08 m) conditions, for a total of 480 trials. Each test session was split into 3 blocks with 160 trials per block to avoid fatigue.

For the stereoacuity assessments, a forced-choice method of constant stimuli procedure was used to assess performance on 11 test disparities presented 20 times apiece. The range of test disparities was adjusted for each observer based on their performance

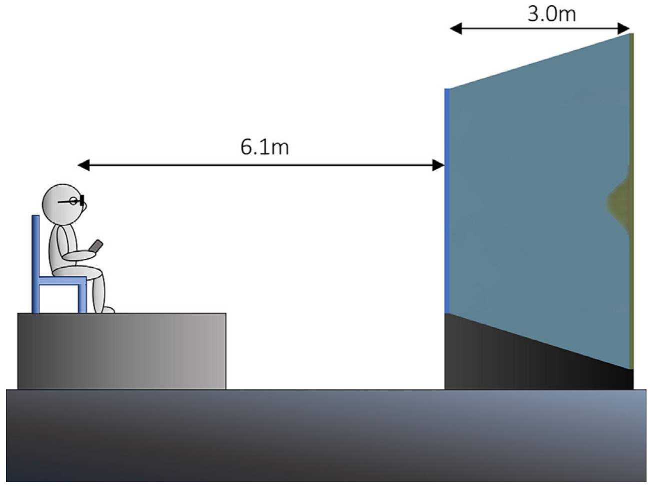


Figure 4. A side view illustration of the experimental layout and stimulus for the terrain task. The vertical blue line represents the projection screen, and the green line represents the ground plane of the three-dimensional imagery. The space between the projection screen (blue line) and the observer is real, while the space from the projection screen to the virtual ground plane (green line) is virtual. Observers were positioned 6.1 m from the projection screen and the ground plane of the stimulus was presented with uncrossed disparity relative to the screen plane producing the impression of the ground plane at a distance of 9.1 m. The size of the mound is not to scale.

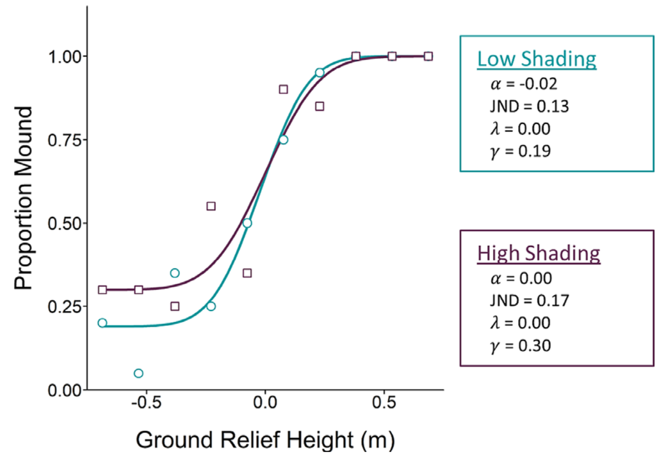


Figure 5. An example of one observer's psychometric function for the binocular condition in the terrain task. The plot shows the proportion of times the observer responded mound as a function of the height of the ground relief for both the low (circles) and high shading conditions (squares). The fitted values for each psychometric function are shown on the right. α represents the inflection point, JND is the difference threshold between 0.50 and 0.75 proportion correct, 1-λ represents the upper asymptote, and γ represents the lower asymptote.

on the book-based stereoacuity tests (i.e., Randot Stereotest and FLY Stereo Acuity test booklets) and a 30-trial practice block. While random dot stereograms rely on the global processing of disparity signals (Julesz, 1960), stereoacuity measured using isolated figures taps into more local disparity mechanisms (Wilcox & Allison, 2009). For each observer, the range of test disparities used was the same for the ledge and bar tests. On each trial, the stimulus remained visible until a response was made. The task in both the ledge and bar tests was to indicate the direction of the depth offset. In the ledge test, observers indicated whether the top or bottom half of the screen appeared closer; in the bar test, observers indicated whether the rectangle appeared in front of or behind the reference frame. Between each trial, a fixation object was presented at the center of the display on the screen plane. For the ledge test, this was a single fixation dot with a diameter of 0.24° and for the bar test it was the outer frame of the stimulus. The results of all tasks were fit with a normal cumulative distribution function that allowed for lapses at both ends of the function. The upper asymptote of the function is represented by $1 - \lambda$, and the lower asymptote was represented by γ. Both λ and γ were fit independently. The functions were fit using MLE and bootstrapped 95% confidence intervals were calculated using Monte Carlo methods (Wichmann & Hill, 2001a, 2001b). From this we computed the just noticeable difference (JND) as the difference threshold between 0.50 and 0.75 proportion correct and inflection point (α) for each psychometric function for all observers (an example function is shown in Figure 5). The goodness of fit measure for each psychometric fit used the method of simulating deviance as described in Wichmann and Hill (2001a). The data were simulated from the fitted model assuming a binomial observer and the function was refit to these simulated data. From these 10,000 simulated fits, deviance was calculated and compared to the deviance of the original dataset. If the deviance of the original dataset exceeded the 95th percentile of the deviances from the simulated fits, then the fit to the original data was considered a failure.

Results

Figure 6 shows the average JND and α values in meters for the two shading conditions. The negative values of α represent dips and positive values represent mounds. A negative α indicates a bias toward perceiving mounds and a positive α indicates a bias toward perceiving dips. The average JND was 0.19 m for the low shading condition and 0.21 m for the high shading condition. The majority of observers (70%) detected changes in ground relief of at least 0.38 m (approximately 60 arcsecs). However, in the high shading condition, 13 of the observers (approximately 30%) always saw convex features; as a result, their data

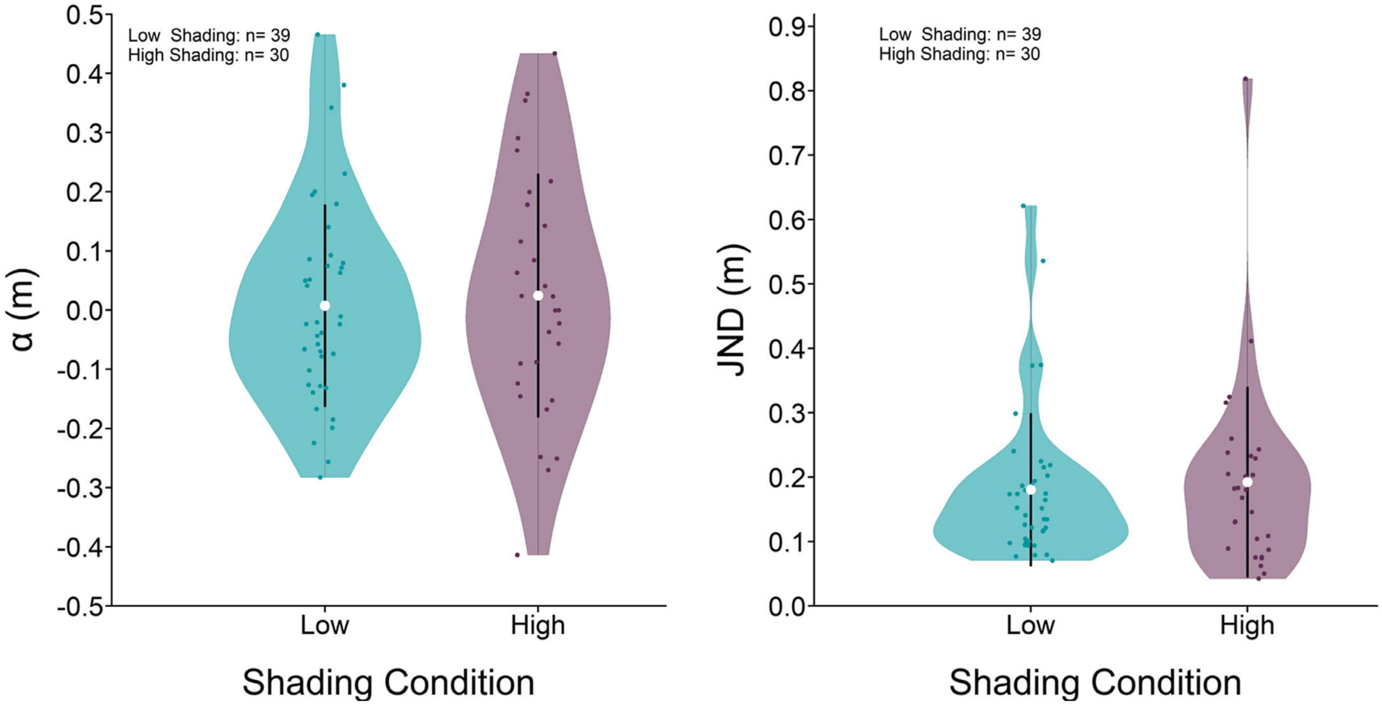


Figure 6. Average α and JND (m) for low shading ($n = 39$), and high shading ($n = 30$) conditions shown as violin plots. The white circle is the mean, the black bar is ± 1 standard deviation, and the faint line is the range of the data. Each colored point represents an individuals' data point, and the shape of the violin represents the distribution of the data. The density estimation was fit using a Gaussian kernel with a smoothing bandwidth using Silverman's rule-of-thumb (or 0.9 times the minimum standard deviation and interquartile range divided by 1.34 times the sample size to the negative one-fifth power). Thirteen participants ($n = 13$) were removed from the high shading condition, and four of these same participants ($n = 4$) were removed from the low shading condition owing to an inability to fit their psychometric functions.

failed our goodness of fit measure and could not be fit with a psychometric function. These results are not represented in Figure 6 and are discussed separately. To determine if there was a significant difference in the JNDs between the low and high shading conditions, a repeated measures analysis of variance was used to assess the data of observers that successfully achieved a psychometric fit in both shading conditions ($n = 30$). These results showed that the mean JND for these observers did not significantly differ between the two shading conditions, $F(1,29) = 0.21$, $p = 0.65$, $\eta^2 = 0.001$. The same analysis was performed on the α values, which also confirmed that there was no significant difference in the mean α value between the two shading conditions, $F(1,29) = 1.77$, $p = 0.19$, $\eta^2 = 0.005$. Further, a t test revealed that the mean α value for both shading conditions did not significantly deviate from zero disparity, $t(29) = -1.57$, $p = 0.13$. However, it is clear from Figure 6 that, despite the average α values being consistent with zero disparity, the range of the biases in both shading conditions is quite large. The average range of the α values in both shading conditions is approximately ± 2.1 JND.

To classify the observers' tendency to perceive convex features even when stereopsis indicated a concave depression, the strength of this effect was categorized according to the value of the lower asymptote (γ) of each psychometric function. We refer to this tendency to perceive convexity even when none is present as the convexity bias (CB). To better understand the nature of the CB we divided the observers into three groups: 1) No CB, 2) Strong CB, and 3) Extreme CB. The Extreme CB group represents the most extreme cases of the CB in which observers responded "mound" regardless of the direction of the ground relief defined by disparity and failed to achieve a psychometric fit. The No CB and Strong CB groups were delineated according to a value of γ. If the γ value was greater or equal to 0.15, the observer was considered to have a Strong CB and was placed in the Strong CB group. This value was chosen to be above a lapse rate of 2 misclicks per 20 trials for each ground relief height. The average proportions at each ground relief height for each observer group is summarized in Figure 7.

As the height of the mound or dip increased (positively or negatively), the shading gradient and thus the salience of the shading cue increased (Figure 1). For observers that exhibited a CB (Strong CB and Extreme CB groups) the presence of salient shading cues at the extremes of the scale caused them to

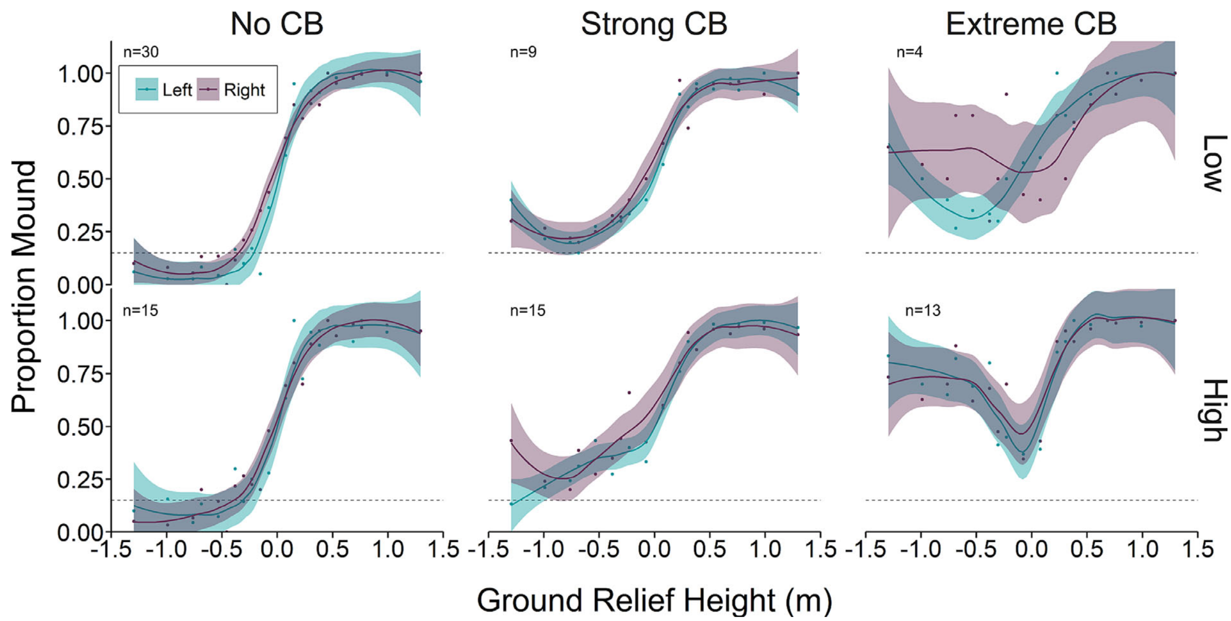


Figure 7. The average proportion mound for each group of observers, (1) No CB, (2) Strong CB, and (3) Extreme CB for the low and high shading conditions with the lighting direction from the left and right. The best fit line represents a loess fit. The shading represents the standard error of the fit. The horizontal dotted line represents the criterion for the γ value. The observers that did not achieve a fit for the psychometric functions exhibited a strong tendency to respond 'mound' especially in the high shading condition as shown by the U-shaped function in the bottom middle. All observers that showed a Strong CB and did not achieve a psychometric fit in the low shading condition ($n = 4$) also showed a Strong CB in the high shading condition.

respond mound more frequently (even when binocular disparity signaled the opposite relief). Further, the lighting direction was not informative about the depth sign and, consistent with this, the magnitude of the CB did not seem to depend on the lighting direction (Figure 7). To verify that the CB for these observers was not due to them having particularly poor stereoacuity, the stereoacuity measures were compared between the three observer groups. We measured stereoacuity with two tasks and the leftmost plot in Figure 8 shows the correlation between these measures. Observers tended to perform worse in the block relative to the ledge task, but the measures were correlated. Therefore, we averaged the two stereoacuity measures for each observer in a subsequent analysis. The stereoacuities for each observer group are shown in the middle bar plot in Figure 8. An analysis of variance confirmed that there was no significant difference between the stereoacuity of the three groups, $F(2,40) = 1.20$, $p = 0.31$, $\eta^2 = 0.06$. Despite the different sample sizes in each observer group, a Levene's test confirmed that there were equal variances between groups, $F(2,40) = 1.33$, $p = 0.28$. Further, to confirm that the magnitude of the CB did not correlate directly with the observers' stereoacuity, the mean stereoacuity of each observer that obtained a psychometric fit (Strong CB and No CB groups) was plotted against their γ value on the terrain task for the low and high shading conditions (Figure 8). The Pearson's correlation coefficient was not significant for either the low shading, $r(37) = 0.11$, $p = 0.51$, or high shading conditions, $r(28) = -0.20$, $p = 0.30$. Thus, the CB exhibited by these individuals was not due to an inability to detect binocular disparity. This finding was also confirmed by the data in the low shading condition as these observers were able to reliably indicate the depth signaled by disparity when shading was less salient.

Further insight into the impact of lighting direction and intensity of shading on the CB was provided by considering performance in the monocular test conditions. To do so, the mean proportion mound responses for each shading condition (left/right) was plotted as a function of the height of the ground relief (Figure 9). Because the step size was determined individually for each participant, test conditions were sampled unequally. Given that lighting direction (left/right) was randomized, we predicted that observers should not be able to use light direction alone as a cue to surface relief and should perform at chance (50%) under monocular viewing. Although most observers detected changes in ground relief of at least 0.38 m (60 arcsecs) under stereoscopic viewing, below this value the intensity of the shading was subtle and difficult to detect (particularly in the low shading condition). Under monocular viewing, without this shading information the terrain discrimination task became quite difficult. Under these conditions when the shading was less salient (ground reliefs of ± 0.08 m),

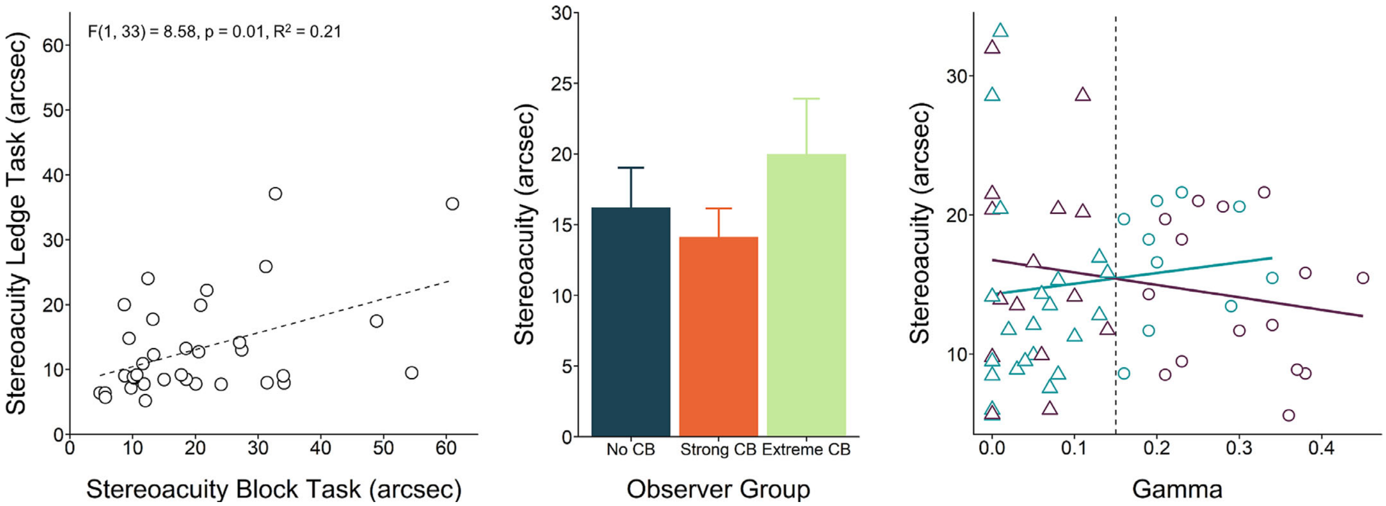
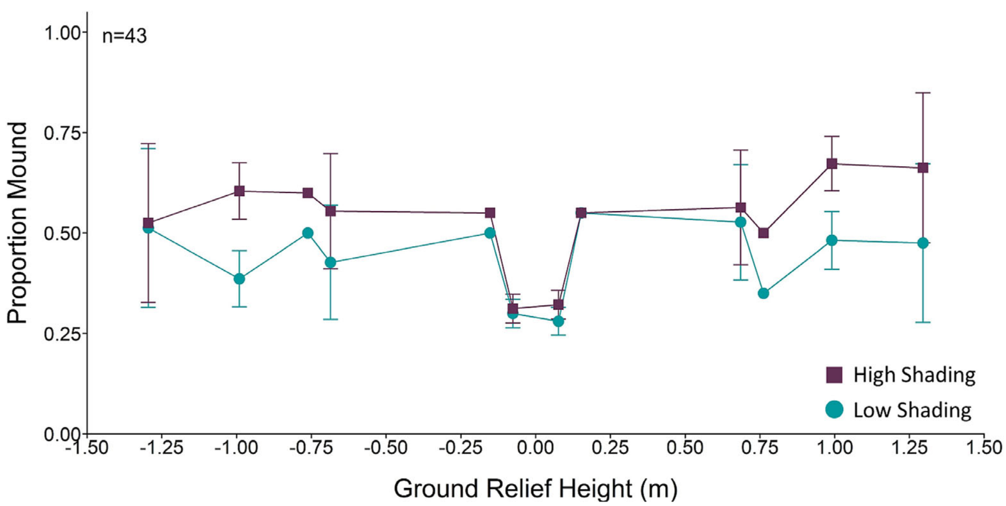


Figure 8. The left scatter plot shows the correlation between the stereoacuity in the ledge task and block task for each observer. The middle bar plot shows the mean stereoacuity of both stereoacuity tests for each observer group. The observer groups consisted of (1) No CB, (2) Strong CB, and (3) Extreme CB ($n = 15$, $n = 15$, and $n = 13$, respectively). The observer groups are split based on the high shading data where the CB was the strongest. The error bars represent one standard error of the mean. The right scatter plot shows the individual stereoacuity for the Strong CB (circles) and No CB (triangles) observers as a function of the γ value of their psychometric functions for the low (blue) and high shading condition (purple). The vertical dashed line represents the 0.15 criterion value for γ. The vertical line easily visualizes the impact of other choices of the criterion as the vertical dashed line would move to the left or right to encompass more or fewer points in each group.



Ground Relief Height (±m)	Number of Observers
0.08	41
0.15	2
0.69	11
0.76	2
1.00	22
1.30	8

Figure 9. Average proportion response mound for the low and high shading conditions for all observers ($n = 43$) for the monocular test conditions. The number of observers for each ground relief is indicated in the insets and error bars indicate ± 1 standard error of the mean. The number of observers varies for the different ground relief heights, because the step size was observer dependent. Error bars have not been plotted for the 0.15 and 0.76 m reliefs because only two observers were tested at these levels.

observers were more likely to report seeing a concavity in both the low and high shading conditions. When ground relief increased and shading cues became more salient, the proportion of mound responses increased for both shading conditions. However, the proportion of mound responses seems to be larger for the high shading relative to the low shading condition, which is consistent with the predisposition to mound responses under binocular viewing and high shading. Given that these monocular trials were interleaved within a much larger subset of stereoscopic trials that exhibited a Strong CB in most conditions, it is possible that observers responded “dip” when the signal was very weak simply because they were more certain it was not a mound.

Given the saliency of shading was reduced at ground reliefs below our safe landing criteria of 0.38 m (60 arcsec), the ground relief conditions were divided into

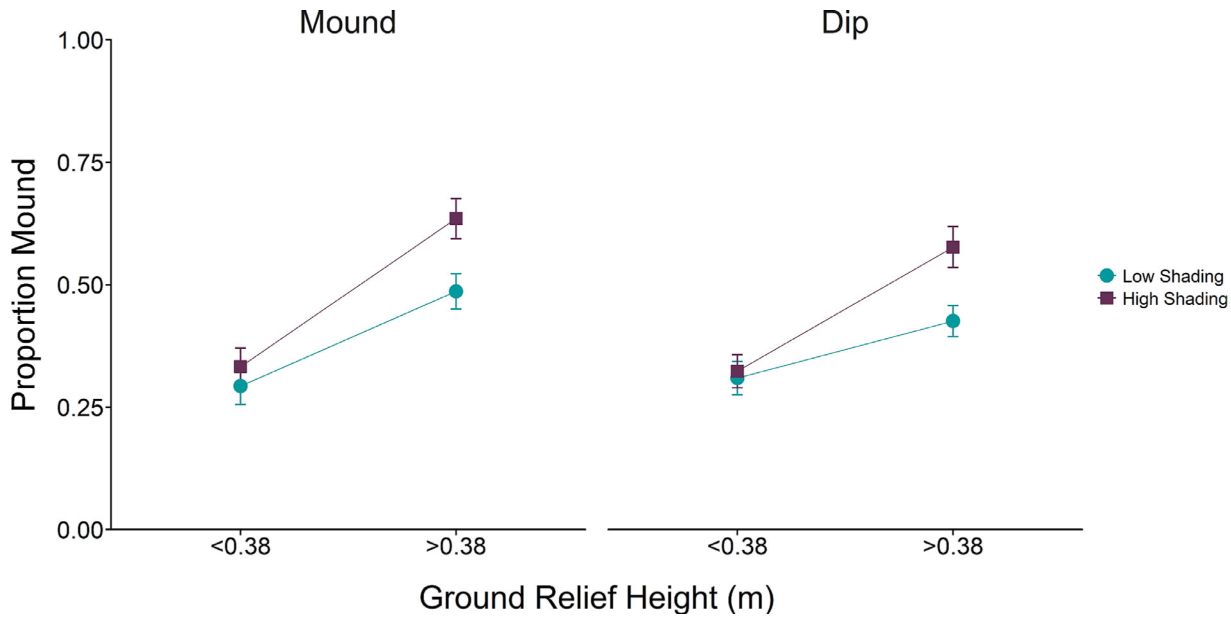


Figure 10. The average proportion response mound for the low and high shading condition for all observers ($n = 43$) for the monocular test conditions. The ground relief heights are averaged for values of less than and greater than 0.38 m. The error bars represent one standard error of the mean.

two groups; 1) less than 0.38 m and 2) greater than 0.38 m (Figure 10). This strategy created equal sized groups for statistical analysis. To determine if the saliency of shading impacted monocular judgments of ground relief, a repeated-measures analysis of variance was conducted on the monocular data in Figure 10. The results revealed no significant three-way interaction between the level of shading (low vs. high), magnitude of ground relief (below vs. above 0.38m), and the direction of the ground relief (mound vs. dip), $F(1,42) = 0.18$, $p = 0.67$, $\eta^2 = 0.0002$. This lack of a three-way interaction suggested that the effect of shading level and magnitude of ground relief were the same, regardless of the direction of surface curvature. This conclusion was supported by the two-way interactions. First, the increase in the proportion of mound responses in the high relative to the low shading condition was the same for mounds and dips, $F(1,42) = 0.15$, $p = 0.70$, $\eta^2 = 0.0001$. Second, the saliency of the shading had a larger impact on the proportion of mound responses at ground reliefs of greater than 0.38 m than of less than 0.38 m, $F(1,42) = 6.38$, $p = 0.02$, $\eta^2 = 0.02$. Thus, at large ground reliefs (>0.38 m) when the shading was more pronounced, observers responded mound more often in the high shading condition relative to the low shading condition. However, at small ground reliefs (<0.38 m) where the shading was less apparent, the responses in the low and high shading conditions were equivalent (Figure 10). Third, the proportion of mound responses was slightly higher for large mounds relative to large dips, but the effect was quite small, $F(1,42) = 4.70$, $p = 0.04$, $\eta^2 = 0.004$. Given the proportion of mound responses for large convex and concave ground reliefs was approximately chance (50%) when only monocular cues were available, this outcome provides another indication that this task was difficult to complete without stereopsis.

Discussion

The aim of the current study was to assess the role of stereopsis in an ecologically valid "real-world" depth discrimination task. We used a naturalistic stimulus where stereopsis was critical to the assessment of ground relief at large distances beyond interaction space. Our results showed that at a portrayed viewing distance of 9.1 m (screen distance of 6.1 m) observers could reliability discriminate ground reliefs producing disparities of 29 to 32 arcsecs (in the low and high shading conditions, respectively), which correspond with ground reliefs of 19 to 21 cm. An analysis of the inflection point (α) of each observer's psychometric functions showed that overall there was no significant bias toward responding mound or dip (Figure 6). However, a closer inspection of interobserver differences showed that, when the shading cue was strong, 30% of observers consistently reported seeing convex surfaces, even when stereopsis indicated a concave depression (Figure 7). Observers that showed this CB in less salient shading conditions also showed a Strong CB in the high shading condition. Importantly, we confirmed that this was not due to the observer's ability to discriminate depth from binocular disparity at these distances (Figure 8). Last, the discrimination results under monocular viewing confirmed that this CB was driven by the saliency of the shading information and increased as the magnitude of ground relief increased (for both convex and concave surfaces), even in the absence of stereopsis (Figure 9).

The tendency toward seeing surfaces as convex is consistent with previous reports of such CBs when viewing face-like stimuli and shaded disks (Gregory, 1970; Perrett & Harries, 1988; Langer & Bulthoff, 2001; Champion & Adams, 2007; Hill & Johnston, 2007; Adams & Elder, 2014). Perceived shape from shading is also constrained by the light-from-above assumption, that is, we assume the light source comes from above (Rittenhouse, 1796; Brewster, 1847; Kleffner & Ramachandran, 1992; Adams, 2007). We deliberately avoided a role for the light-from-above assumption by positioning the light source to the left or right of the feature. Although there is some evidence for a CB in interpreting shading as based on light coming slightly from the left (Sun & Perona, 1998; Mamassian & Goutcher, 2001), this effect seems to disappear when stimuli are presented for an unrestricted duration (McManus, Buckman, & Woolley, 2004; Aubin & Arguin, 2014). Our finding that the CB did not depend on lighting direction supports these observations. Although the bias toward convexity is closely related to the assumption for lighting direction, it has been argued that the CB is stronger than the lighting direction assumption (Liu & Todd, 2004). Even so, it is surprising that such a Strong CB persists in our stimuli given that the shading cue does not disambiguate the direction of the ground relief and normally relies on assumptions of lighting direction to do so (Brewster, 1826; Sun & Perona, 1998; Adams, Graf, & Ernst, 2004).

We used a grass texture to represent a typical covering for the ground surface. At the distance tested in this experiment the texture gradients and other static monocular perspective cues provided by this texture seemed to be of limited usefulness in making the judgments. Thus, although these monocular cues were consistent with depth from stereopsis, on their own they could not be used to make correct depth discrimination judgments. This is evident in Figure 9, where monocular performance was at chance.

When texture and stereopsis cues are present, stereopsis is theoretically more reliable for large slants at shorting viewing distances (Knill & Saunders, 2003; Hillis, Watt, Landy, & Banks, 2004), although observers vary in the relative weight given perspective and disparity (Gillam & Ryan, 1992; Allison & Howard, 2000). When stereopsis is presented in conjunction with inconsistent monocular cues (such as occlusion and texture gradients), the monocular cues could override stereopsis (Braunstein, Andersen, Rouse, & Tittle, 1986; Stevens & Brookes, 1988). However, there are scenarios where stereopsis can override monocular cues, such as the prevention of shape inversion when lighting is from below (Bulthoff & Mallot, 1990). In general, the influence of conflicting cues on shape perception can be complex and vary as a function of scene structure, as has been often reported for pseudoscopic viewing where the left and right images of a stereo pair are swapped (Wheatstone, 1852; Stratton, 1898; Shimojo & Nakajima, 1981; Kalaugher, 1987; Palmisano, Hill, & Allison, 2016). Studies that examined the combination of shading and stereopsis under restricted conditions have concluded that these two cues are processed independently and combined linearly to determine perceived surface shape (Lovell, Bloj, & Harris, 2012; Aubin & Arguin, 2014). However, these experiments were all performed at near viewing distances (<1 m) with reliable binocular disparity information. They also tended to rely on cue conflict scenarios, whereas in the current study monocular and binocular cues were always consistent when present.

It is likely that the use of a relatively large viewing distance of 9.1 m decreased the cue reliability in our study. It has been demonstrated that the reliability of stereopsis decreases as viewing distance increases beyond interaction space (Banks, Hooge, & Backus, 2001; Knill & Saunders, 2003; Hillis, Watt, Landy, & Banks, 2004), and the reliability of texture cues also decreases as stimulus size decreases with distance (Blake, Bulthoff, & Sheinberg, 1993; Knill, 1998). However, it has been argued that stereoscopic thresholds are determined solely by angular disparity, invariant of viewing distance when monocular information is limited (Ogle, 1958; Bradshaw & Glennerster, 2006). Similarly, in studies using naturalistic imagery, stereoscopic discrimination remains relatively precise at large viewing distances (McCann, Hayhoe, & Geisler, 2018). Further, we show that monocular depth cues were unable to support accurate shape judgments when presented in isolation (Figure 9). Under these conditions of high uncertainty, the presence of highly salient shading information seemed to activate assumptions of surface curvature for 65% of our observers. This CB was strong enough to override depth from binocular disparity. However, at shorter viewing distances (1.2 m) there are cases where stereopsis overrides less ambiguous shape from shading cues when the two cues indicate conflicting shape information (Bulthoff & Mallot, 1990). Importantly, it was not that the discrimination task could not be completed with stereopsis at this viewing distance as it was apparent that when the shading cue was less salient (i.e., the low shading condition) most observers had no issues completing the task (Figure 7).

Here, when stereoscopic, texture, and shading cues were all consistent with the true depth of the surface, a bias toward surface convexity was able to assert itself, despite little evidence for it in the scene. The presence of highly salient shading cues resulted in a bias toward perceiving convex features under both monocular (Figure 10) and binocular conditions (Figure 7). The shading cue itself did not provide information regarding the sign or direction of the curvature, which was consistent with the finding that monocular depth cues alone were insufficient to support accurate shape discrimination (Figure 9). Despite this, as the saliency of shading increased, several observers experienced a Strong CB, even when stereopsis indicated that the ground relief was concave (Figure 7). This effect may be related to other phenomena in which depth from disparity augments depth percepts from monocular sources, even when the sign does not match. For instance, it is well-documented that in pseudoscopic viewing vivid percepts of depth, consistent with the monocular depth cues, are experienced (Palmisano, Hill, Allison, 2016). This occurs despite the fact that binocular disparity is reversed throughout the scene. In other instances, binocular disparity provides compelling depth information, but monocular cues dominate the sign of perceived depth. For example, perspective cues can dominate the perception of surface slant even when this information conflicts with binocular disparity signals (Stevens, Lees, & Brookes, 1991; Allison & Howard, 2000). In these examples, binocular disparity seems to provide information regarding the nonplanarity or range of depth in the scene, but the sign of depth depends strongly on monocular information. In our study, although the salient shading information did not specify the sign of the curvature, both shading and binocular disparity indicated that substantial curvature was present. Under these conditions, some observers seemed to apply a bias toward convexity to interpret the shading despite evidence to the contrary from binocular disparity.

Conclusions

Our results show that consistent stereopsis, texture, and salient but ambiguous shading information allowed observers to readily discriminate ground relief at viewing distances beyond interaction space. Stereopsis, in particular, was critical for the reliable discrimination of ground relief at these viewing distances, given that its absence greatly decreased performance. However, under these conditions some observers were apt to rely heavily on assumptions of surface convexity. The extent and frequency of this effect was somewhat surprising and has several implications. First, it highlights the presence of significant individual differences in the reliance on assumptions of surface curvature, a factor that must be considered when assessing performance. The individual nature of the CBs makes them difficult to predict. Second, given that these CBs were quite dramatic, they are likely also present during real life terrain judgments. However, in rich full-cue natural viewing environments, observers usually have no issue locating reliable convex footholds on irregular terrain where visual uncertainty is greater (Hayhoe et al., 2009; Bonnen et al., 2021). One possibility is that, when walking over terrain, observers have more cues to the ground structure given an oblique viewing angle to the ground plane (e.g., mounds would occlude more distant ground terrain). Judgments of ground relief in viewing scenarios where additional depth cues are available (e.g., motion parallax, occlusions, height in the field, or cast shadows) could help to disambiguate ground relief. It is possible that in impoverished viewing environments with limited cues these CBs could persist in judgments of ground relief. For instance, the assessment of ground relief is a requirement during rotary wing landing maneuvers, which require determining ground relief from a top down view at moderately large viewing distances, similar to the viewing geometry in our study. The observed CB, combined with a tendency to assume that light sources are overhead, could produce similar depth reversals under conditions with strong ambiguous shading (such as sunrise or dusk). Accurate assessment of ground relief is an essential and common requirement while traversing our natural environment on foot or in vehicles.

Keywords: binocular disparity, shading, shape, convexity

Acknowledgments

Funded by the Canadian Department of National Defence under the scientific direction of Defence Research and Development Canada.

Commercial relationships: none.
Corresponding author: Brittney Hartle.
Email: brit1317@yorku.ca.
Address: Centre for Vision Research, Lassonde Building, York University, 4700 Keele Street, Toronto, ON, Canada.

References

Adams, W. J. (2007). A common light-prior for visual search shape and reflectance judgements. *Journal of Vision, 7*(11), 1–7.

Adams, W. J., & Elder, J. H. (2014). Effects of specular highlights on perceived surface convexity. *Computational Biology, 10*(5), 1–13.

Adams, W. J., Graf, E. W., & Ernst, M. O. (2004). Experience can change the 'light-from-above' prior. *Nature Neuroscience, 7*(10), 1057–1058.

Allison, R. S., Gillam, B. J., & Palmisano, S. A. (2009) Stereoscopic discrimination of the layout of ground surfaces. *Journal of Vision, 9*(12):8, 1–11.

Allison, R. S., Gillam, B. J., & Vecellio, E. (2009). Binocular depth discrimination and estimation beyond interaction space. *Journal of Vision, 9*(1), 10–10.

Allison, R. S., & Howard, I. P. (2000). Temporal dependencies in resolving monocular and binocular cue conflict in slant perception. *Vision Research, 40*, 1869–1885.

Aubin, M., & Arguin, M. (2014). Stereo and shading contribute independently to shape convexity-concavity discrimination. *Perception, 43*(4), 333–343.

Banks, M. S., Hooge, I. T., & Backus, B. T. (2001). Perceiving slant about a horizontal axis from stereopsis. *Journal of Vision, 1*(2), 1.

Barton, S. L., Matthis, J. S., & Fajen, B. R. (2017). Visual regulation of gait: Zeroing in on a solution to the complex terrain problem. *Journal of Experimental Psychology: Human Perception and Performance, 43*(10), 1773.

Blake, A., Bülthoff, H. H., & Sheinberg, D. (1993). Shape from texture: Ideal observers and human psychophysics. *Vision Research, 33*(12), 1723–1737.

Blakemore, C. (1970). The range and scope of binocular depth discrimination in man. *Journal of Physiology, 211*(3), 599.

Bonnen, K., Matthis, J. S., Gibaldi, A., Banks, M. S., Levi, D. M., & Hayhoe, M. (2021). Binocular vision and the control of foot placement during walking in natural terrain. *Scientific Reports, 11*(1), 1–12.

Bradshaw, M. F., & Glennerster, A. (2006). Stereoscopic acuity and observation distance. *Spatial Vision, 19*(1), 21–36.

Braunstein, M. L., Andersen, G. J., Rouse, M. W., & Tittle, J. S. (1986). Recovering viewer-centered depth from disparity, occlusion, and velocity gradients. *Perception & Psychophysics, 40*(4), 216–224.

Brewster, D. (1826). On the optical illusion of the conversion of cameos into intaglios, and of intaglios into cameos, with an account of other analogous phenomena. *Edinburgh Journal of Science, 4*, 9–108.

Brewster, D. (1847). On the conversion of relief by inverted vision. *Transactions of the Royal Society of Edinburgh, 15*, 657–662.

Bülthoff, H. H., & Mallot, H. A. (1990). Integration of stereo, shading and texture. In *11th European Conference on Visual Perception (ECVP 1988)* (pp. 119–146). New York: Wiley.

Champion, R. A., & Adams, W. J. (2007). Modification of the convexity prior but not the light-from-above prior in visual search with shaded objects. *Journal of Vision, 7*(13), 1–10.

Foley, J. M. (1985). Binocular distance perception: Egocentric distance tasks. *Journal of Experimental Psychology: Human Perception and Performance, 11*(2), 133.

Gillam, B., & Ryan, C. (1992). Perspective, orientation disparity, and anisotropy in stereoscopic slant perception. *Perception, 21*, 427–439.

Gogel, W. C. (1977). An indirect measure of perceived distance from oculomotor cues. *Perception & Psychophysics, 21*(1), 3–11.

Gregory, R.L., 1970. *The intelligent eye*. London: Weidenfeld and Nicolson.

Hayhoe, M., Gillam, B., Chajka, K., & Vecellio, E. (2009). The role of binocular vision in walking. *Visual Neuroscience, 26*(1), 73–80.

Hill, H., & Johnston, A. (2007). The hollow-face illusion: Object-specific knowledge, general assumptions or properties of the stimulus? *Perception, 36*, 199–223.

Hillis, J. M., Watt, S. J., Landy, M. S., & Banks, M. S. (2004). Slant from texture and disparity cues: Optimal cue combination. *Journal of Vision, 4*(12), 1.

Howard, I. P. (2012). *Perceiving in depth, volume 1: basic mechanisms*. Oxford, UK: Oxford University Press.

Julesz, B. (1960). Binocular depth perception of computer-generated patterns. *Bell System Technical Journal, 39*(5), 1125–1162.

Kalaugher, P. G. (1987). Pseudoscopic viewing: Transfer and persistence of reversed depth relations from the viewing of photographs to the real scene. *Perception, 16*, 359–374.

Kleffner, D. A., & Ramachandran, V. S. (1992). On the perception of shape from shading. *Perception & Psychophysics, 52*(1), 18–36.

Knill, D. C. (1998). Surface orientation from texture: ideal observers, generic observers and the information content of texture cues. *Vision Research, 38*(11), 1655–1682.

Knill, D. C., & Saunders, J. A. (2003). Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision Research, 43*(24), 2539–2558.

Langer, M. S., & Bülthoff, H. H. (2001). A prior for global convexity in local shape from shading. *Perception, 30*, 403–410.

Lappin, J. S., & Craft, W. D. (2000). Foundations of spatial vision: From retinal images to perceived shapes. *Psychological Review, 107*, 6–38.

Liu, B., & Todd, J. T. (2004). Perceptual biases in the interpretation of 3D shape from shading. *Vision Research, 44*(18), 2135–2145.

Lovell, P. G., Bloj, M., & Harris, J. M. (2012). Optimal integration of shading and binocular disparity for depth perception. *Journal of Vision, 12*(1), 1.

Mamassian, P., & Goutcher, R. (2001). Prior knowledge on the illumination position. *Cognition, 81*(1), B1–B9.

Matthis, J. S., Yates, J. L., & Hayhoe, M. M. (2018). Gaze and the control of foot placement when walking in natural terrain. *Current Biology, 28*(8), 1224–1233.

McCann, B.C., Hayhoe, M.M., & Geisler, W.S. (2018). Contributions of monocular and binocular cues to distance discrimination in natural scenes. *Journal of Vision, 18*(4), 1–15.

McManus, I. C., Buckman, J., & Woolley, E. (2004). Is light in pictures presumed to come from the left side? *Perception, 33*(12), 1421–1436.

Norman, J. F., Lappin, J. S., & Zucker, S. W. (1991). The discriminability of smooth stereoscopic surfaces. *Perception 20*, 789–807.

Ogle, K.N. (1958). Note on stereoscopic acuity and observation distance. *Journal of the Optical Society of America, 48*(11), 794–798.

Palmisano, S., Gillam, B., Govan, D. G., Allison, R. S., & Harris, J. M. (2010). Stereoscopic perception of real depths at large distances. *Journal of Vision, 10*(6), 19–19.

Palmisano, S., Hill, H., & Allison, R. S. (2016). The nature and timing of tele-pseudoscopic experiences. *I-Perception, 7*(1), 1–24.

Perrett, D. I., & Harries, M. H. (1988). Characteristic views and the visual inspection of simple faceted and smooth objects: 'Tetrahedra and potatoes'. *Perception, 17*(6), 703–720.

Rittenhouse, D. (1786). Explanation of an optical deception. *Transactions of the American Philosophical Society, 2*, 37–42.

Rogers, B. J., & Bradshaw, M. F. (1993). Vertical disparities, differential perspective and binocular stereopsis. *Nature, 361*(6409), 253–255.

Shimojo, S., & Nakajima, Y. (1981). Adaptation to the reversal of binocular depth cues: Effects of wearing left-right reversing spectacles on stereoscopic depth perception. *Perception, 10*, 391–402.

Stevens, K. A., & Brookes, A. (1988). Integrating stereopsis with monocular interpretations of planar surfaces. *Vision Research, 28*(3), 371–386.

Stevens, K.A., Lees, M., & Brookes, A. (1991). Combining binocular and monocular curvature features. *Perception, 20*(4), 425–440.

Stratton, G. (1898). A mirror pseudoscope and the limits of visual depth. *Psychological Review, 5*, 632–638.

Sun, J., & Perona, P. (1998). Where is the sun? *Nature Neuroscience, 1*, 183–184.

Transport Canada. (2006). *Helicopter Training manual.* Available at: https://tc.canada.ca/sites/default/files/migrated/tp9982e.pdf. Accessed XXXX.

Wheatstone, C. (1852). Contributions to the physiology of vision—Part the second. On some remarkable and hitherto unobserved phenomena of binocular vision. *Philosophical Transactions of the Royal Society of London, 142*, 1–17.

Wichmann, F. A., & Hill, N. J. (2001a). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception and Psychophysics. 63*(1), 1293–1313.

Wichmann, F. A., & Hill, N. J. (2001b). The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Perception and Psychophysics. 63*(8), 1314–1329.

Wilcox, L. M., & Allison, R. S. (2009). Coarse-fine dichotomies in human stereopsis. *Vision Research, 49*(22), 2653–2665.

Zhao, J., & Allison, R. S. (2021). The role of binocular vision in avoiding virtual obstacles while walking. *IEEE Transactions on Visualization and Computer Graphics, 27*(7), 3277–3288.