Both parallelism and orthogonality are used to perceive 3D slant of rectangles from 2D images

Jeffrey A. Saunders   Department of Psychology, University of Pennsylvania, Philadelphia, PA, USA

Benjamin T. Backus   Department of Psychology, University of Pennsylvania, Philadelphia, PA, USA

A 2D perspective image of a slanted rectangular object is sufficient for a strong 3D percept. Two computational assumptions that could be used to interpret 3D from images of rectangles are as follows: (1) converging lines in an image are parallel in the world, and (2) skewed angles in an image are orthogonal in the world. For an accurate perspective image of a slanted rectangle, either constraint implies the same 3D interpretation. However, if an image is rescaled, the 3D interpretations based on parallelism and orthogonality generally conflict. We tested the roles of parallelism and orthogonality by measuring perceived depth within scaled perspective images. Stimuli were monocular images of squares, slanted about a horizontal axis, with an elliptical hole. Subjects judged the length-to-width ratio of the holes, which provided a measure of perceived depth along the object. The rotational alignment of squares within their surface plane was varied from 0° (trapezoidal projected contours) to 20° (skewed projected contours). In consistent-cue conditions, images were accurate projections of either a 10°- or 20°-wide square, with slants of 75° and 62°, respectively. In cue-conflict conditions, images were generated either by magnifying a 10° image to have a projected size of 20° or by minifying a 20° image to have a projected size of 10°. For the aligned squares, which do not produce a conflicting skew cue, we found that subjects' judgments depended primarily on projected size and not on the size used to generate the prescaled images. This is consistent with reliance on the convergence cue, corresponding to a parallelism assumption. As squares were rotated away from alignment, producing skewed projected contours, judgments were increasingly determined by the original image size. This is consistent with use of the skew cue, corresponding to an orthogonality assumption. Our results demonstrate that both parallelism and orthogonality constraints are used to perceive depth from linear perspective.

Keywords: linear perspective, depth perception, slant perception, skew symmetry, picture perception, vision

Introduction

When a scene contains rectangular surfaces, linear perspective can convey a strong percept of 3D structure, even from a monocular image. There are two constraints that could provide the basis for linear perspective as a depth cue: parallelism and orthogonality. Consider the image shown in Figure 1a. If one assumed that opposite pairs of edges of the 3D object are parallel (Figure 1b), then their convergence in the projected image provides a 3D cue (Saunders & Backus, 2006a). Alternatively, one could assume that the intersecting edges or axes of the object form orthogonal angles (Figure 1c), in which case, their skewed intersection angles in the projected image provide a 3D cue (see Saunders & Knill, 2001). We will refer to these as *convergence* and *skew* cues, respectively. Either or both of these cues could potentially be used to perceive depth from images of rectangular objects.

Previous work has shown that both perspective convergence and skew can be effective 3D cues. A number of studies have demonstrated that observers can make reliable 3D judgments from minimal stimuli that provide only convergence information (Andersen, Braunstein, & Saidpour, 1998; Clark, Smith, & Rabe, 1955, 1956; Freeman, 1966a, 1966b; Rosinski, Mulholland, Degelman, & Farber, 1980; Saunders & Backus, 2006a; Smith, 1967; Stavrianos, 1945; Todd, Thaler, & Dijkstra, 2005). Other results have shown that convergence contributes to slant perception even in the presence of conflicting stereo information (Attneave & Olson, 1966; Banks & Backus, 1998; Braunstein & Payne, 1969; Gillam, 1968; Saunders & Backus, 2006b; Smith, 1967). The case of skew as a 3D cue has been less studied. Griffiths and Zaidi (2000) observed systematic biases in the perceived 3D orientations of parallelogram-shaped objects, consistent with a misapplied assumption of orthogonal corners. Saunders and Knill (2001) report evidence that skew symmetry is used to perceive 3D orientation from a projected contour, even when stereo information is available.

In an accurate perspective image of a rectangle, convergence and skew specify the same 3D interpretation. However, it is possible to put these cues in conflict. One way is to scale a perspective image, which changes the 3D
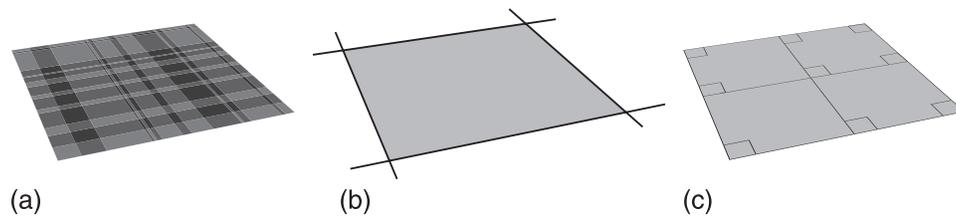
Figure 1. (a) Perspective projection of a slanted rectangular object. (b) Parallel edges converge in the image, providing a *convergence* cue. (c) Orthogonal intersections form skewed angles in an image, providing a *skew* cue.

interpretation from convergence (Farber & Rosinski, 1978; Saunders & Backus, 2006a). Figure 2 illustrates how scaling a perspective image can introduce conflicts. The 3D interpretation of the minified image under an assumption of parallel sides (middle right) is a skewed parallelogram, elongated in depth, that is more slanted than the original square. The axes and corners of this object are nonorthogonal. The planar 3D interpretation that is closest to having orthogonal intersections (lower right) is an object with nonparallel sides, with the same slant as the original square. In the experiment reported here, we used scaled images like in this example to test the roles of parallelism and orthogonality as perceptual constraints, measuring how closely perceived 3D shape and slant agreed with either a parallel-sides interpretation or a maximally orthogonal interpretation of a monocular image.

The conflict between convergence and skew resulting from image scaling depends on how a rectangle is oriented within its surface plane. Following Saunders and Knill (2001), we define the *spin* of a slanted rectangle to be the angle between its vertical axis and the direction along the surface corresponding to increasing depth. This angle, together with *slant* and *tilt* (Stevens, 1983), fully specifies the 3D orientation of a rectangle. In the special case when a rectangle is aligned with the direction of tilt, as illustrated in Figure 3, scaling an image does not introduce conflicting skew information. Regardless of how the image is scaled, its interpretation would be a slanted rectangle. Image size would affect the length-to-width ratio of the 3D rectangle (smaller images correspond to rectangles that are longer relative to their width), but corners would remain orthogonal. This contrasts with the more generic case shown in Figure 2, for which convergence information specifies a 3D object shaped like a skewed parallelogram, whereas maximizing the orthogonality of intersections corresponds to a different 3D object.

For our purposes, a nuisance variable is the slant information provided by contour foreshortening under an assumption of isotropy: Surface slant could be inferred to be that which corresponds to the most circularly symmetric object, rather than the slant specified by convergence or skew. Previous studies of slant-from-texture have found that, in the case of isotropic textures, foreshortening

(or compression) is the dominant texture cue (Buckley, Frisby, & Blake, 1996; Knill, 1998; Rosenholtz & Malik, 1997; Saunders, 2003). However, other results suggest that foreshortening has only a weak influence on perceived slant when in conflict with convergence information (Braunstein & Payne, 1969; Saunders & Backus, 2006a, 2006b). Our focus here is on convergence and skew cues, not foreshortening. We used irregular plaid textures, as in the examples, to degrade foreshortening of texture elements as a cue. Testing squares with varying spins enabled us to control for an influence of contour foreshortening. As can be seen in Figure 4, isotropic interpretations have the same slant regardless of spin, whereas the maximally orthogonal interpretations do not. This invariance to spin distinguishes use of isotropy as a perceptual constraint from use of an orthogonality assumption.

There is evidence from previous work that observers are sensitive to conflicts between convergence and skew cues. In an experiment by Nicholls and Kennedy (1993), subjects evaluated perspective line drawings of obliquely oriented cubes that were scaled by varying amounts. The unscaled perspective drawings, which did not present conflicting convergence and skew information, were rated as the best depictions of cubes. Yang and Kubovy (1999) performed a similar experiment and also observed a preference for consistent perspective images over scaled images with conflicting cues. On the other hand, in both experiments, a range of scaled images was also rated highly, suggesting that observers can accommodate some degree of conflict.

In the experiment reported here, we tested four combinations of object size, slant, and image scaling. In the *small-object* conditions, images were accurate perspective projections of a square that subtended 10° in width, slanted by 75°. In the *large-object* conditions, images were accurate projections of a 20° square, slanted by 62°. In the *magnified* conditions, images of the small object were uniformly scaled by a factor of 2, such that they subtended 20° horizontally, whereas in the *minified* conditions, images of the large object were uniformly scaled by half to subtend 10°. For magnified images, the isotropic interpretation has the same slant used to generate the small-object image, 75°; the interpretation with parallel edges has a lower slant, 62°, and the maximally
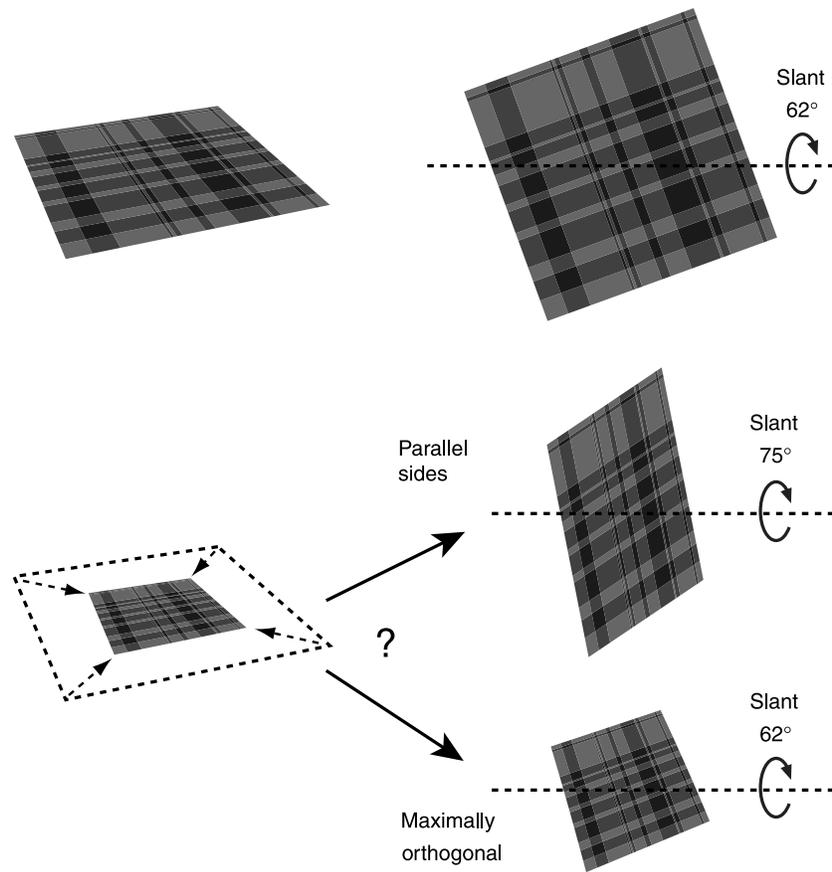
Figure 2. Effect of image scaling on 3D interpretations from convergence and skew information. The base image is a perspective projection of a slanted square (top). When this image is minified (left), the 3D interpretation with parallel sides (middle right) is elongated in depth and nonorthogonal and has higher slant. The interpretation that minimizes the average deviation from orthogonality (bottom right) has the same slant as the original square but does not have parallel sides.
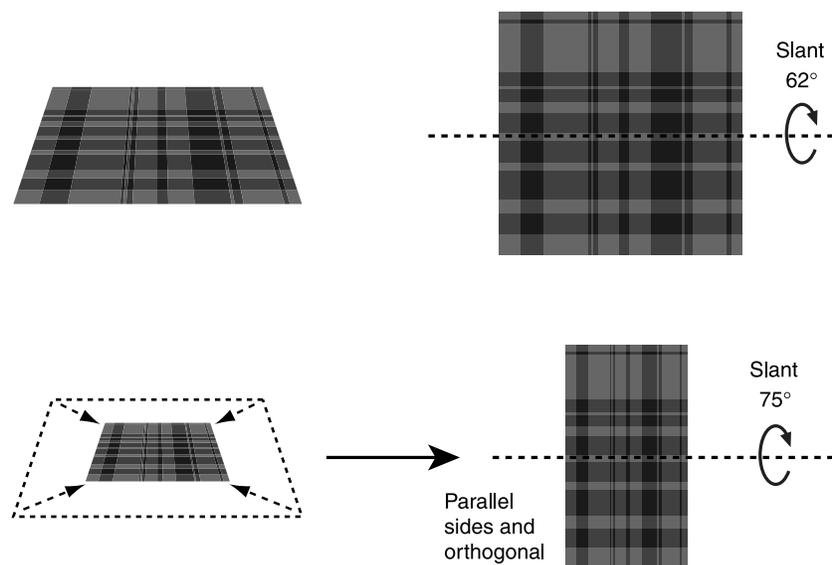

Figure 3. Effect of image scaling when a square is aligned with its direction of tilt. For this special case, a scaled image has an interpretation that is stretched in depth but remains rectangular (parallel sides and orthogonal intersections).
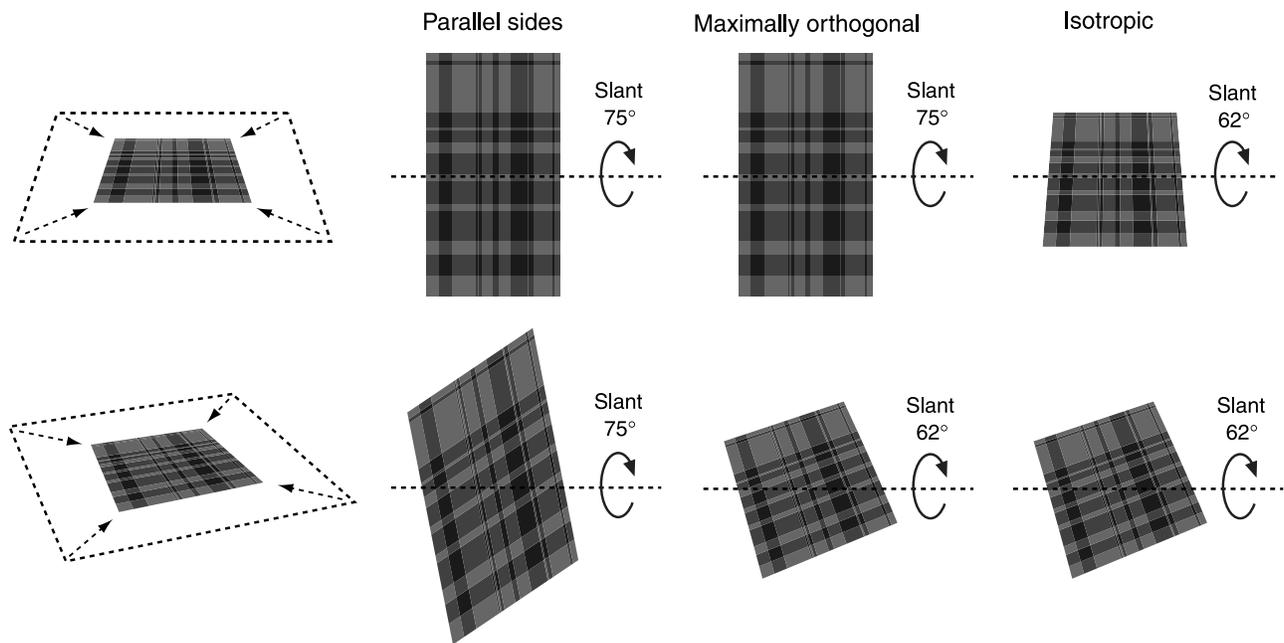
Figure 4. Interpretations of scaled perspective images of squares with 0° spin (top) and 20° spin (bottom), based on different possible constraints: parallel sides (middle left), orthogonal intersections (middle right), or isotropy (right).

orthogonal interpretations have slants between 75° and 62°, depending on spin. For minified images, the relations are the opposite. Note that whereas the unscaled images accurately depict square objects, the magnified and minified images simulate perspective views of objects that are stretched or compressed in depth (as illustrated in Figures 2 and 3). Each of the size and scaling conditions were tested for squares with four magnitudes of spins: 0°, 5°, 10°, and 20°. Figure 5 illustrates the 16 combinations, along with the slants predicted by the different potential cues. Perceived slant was assessed indirectly by having subjects judge the shape of elliptical hole in the surfaces. Figure 6 shows a sample stimulus. The image of the hole by itself provides no information for this task; hence, judgments would necessarily depend on the perceived slant of the overall object.

Methods

Apparatus and display

Stimuli were rear-projected from an InFocus LP350 projector, with a 1,024 × 768 resolution, onto a 166 × 125 cm region of a large screen positioned 2.8 m from the observer. The rectangular projected region subtended 33° horizontally and 25° vertically, and its boundaries were dimly visible. Subjects wore a patch over their left eye throughout the experiment. Subjects were seated on a stool and were instructed to remain stationary during judgments but were not otherwise restricted (no chin rest or bite bar was used). Images were grayscale and antialiased, rendered using OpenGL on a workstation with Nvidia Quadro FX 1000 graphics board.

Stimuli were perspective views of a square textured surface with an elliptical hole, on a black background (Figure 6). Surface textures were random plaid patterns, composed of superimposed horizontal and vertical stripes with irregular size and spacing. The purposes of the texture were to provide convergence and skew information throughout the image and to generally enhance the percept of a 3D object. The horizontal and vertical subdivisions forming the rows and columns were chosen randomly from a uniform distribution. This randomization effectively prevents the texture gradient from providing reliable cues to slant. A different random plaid was generated for each trial.

Squares were slanted around a horizontal axis (i.e., the tilt direction was vertical) and varied in their alignment within their plane (i.e., their spin). Four magnitudes of spin were tested: 0°, 5°, 10°, and 20°. Each of these spins was tested in both positive and negative directions, and trials from both signs were combined for analysis. For small-object conditions, images were accurate perspective projections of a 50-cm (10.2°)-wide square slanted by 75°, whereas for the large-object conditions, the simulated square was 100 cm wide (20.2°) and slanted by 61.8°. The small-object images were uniformly scaled by a factor of 2 to generate the images for magnified conditions, and the large-object images were uniformly scaled by a factor of 1/2 to generate images for the minified conditions. In the case of squares with 0° spin, the projected contours were all trapezoids with converging sides that, by design, were the same orientation, differing from vertical by 18.4°.
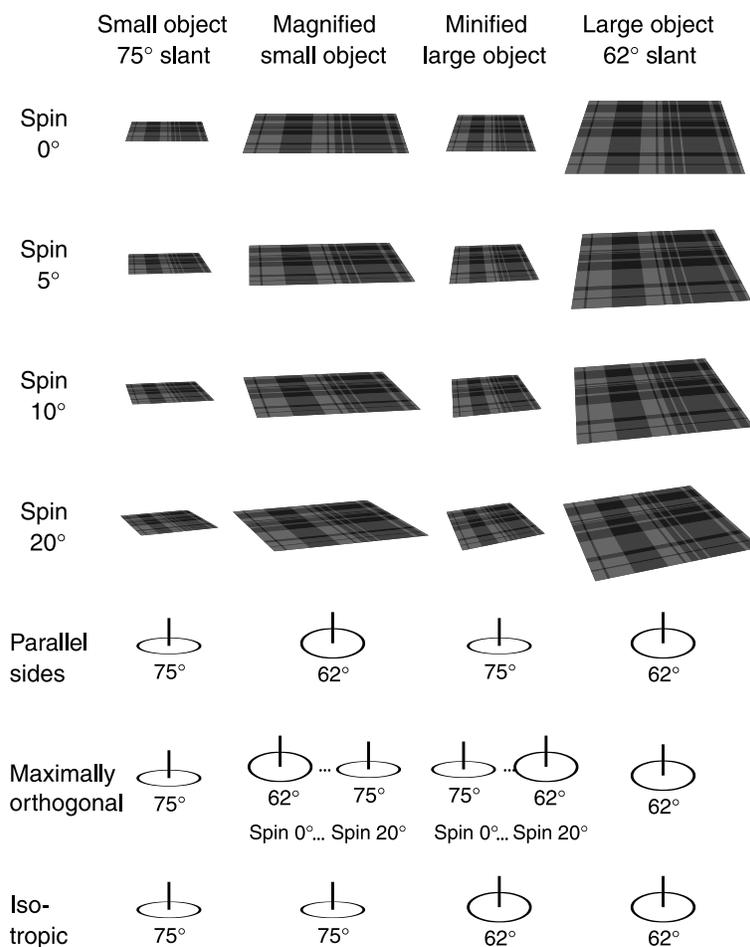
Figure 5. Illustration of the 16 combinations of size, slant and scaling (left to right), and spin (top to bottom) tested in the experiment. Each image was also presented left to right reversed, corresponding to negative spins (not shown). The bottom three rows show the predicted perceived slant depending on what computational assumptions are used.

The square objects were extruded to have a slight thickness, equal to 2% of its width. The thickness served two purposes. One was to generally enhance the perception that the ellipse was a hole in a 3D object, rather than being detached. Second, the visible lips on the hole and on the square provided a cue that objects were slanted around a horizontal axis (tilt direction was vertical). As we will discuss later, subjects may still have perceived the objects as having nonzero tilt. The lip was thin to prevent its left and right edges in the projected image from providing a useful convergence cue (i.e., specifying a third, downward vanishing point). The top and bottom edges of the square's visible lip provided a convergence cue that was redundant with that provided by the top face.

The elliptical holes had varying length-to-width ratios (see the Procedure section) and were always aligned vertically. Lengths and widths both varied, such that the area of the hole along the simulated surface was constant. In the case when the hole was circular on the surface, its width was one fourth of the outer square's width.

Procedure

Subjects made forced-choice judgments of whether the hole in the 3D object appeared longer versus wider than a circle. They were instructed to base their judgments on the shape of the hole along the surface, not on its screen projection. Trials were self-paced and subjects received no feedback.

The length-to-width ratios of holes were varied across trials using a minimized expected entropy adaptive staircase method (Saunders & Backus, 2006a). Judgments from a condition were fit to a cumulative Gaussian psychometric function, using maximum likelihood criteria. The mean of the best fitting function was taken as the point of subjective equality (PSE), which, for this task, indicates the length-to-width ratio for which a hole appears circular. The difference between the 75% point and the PSE was taken as the just-noticeable-difference (JND) threshold.

Prescaled size and slant, projected size, and rectangle alignment were randomized within blocks. The experiment
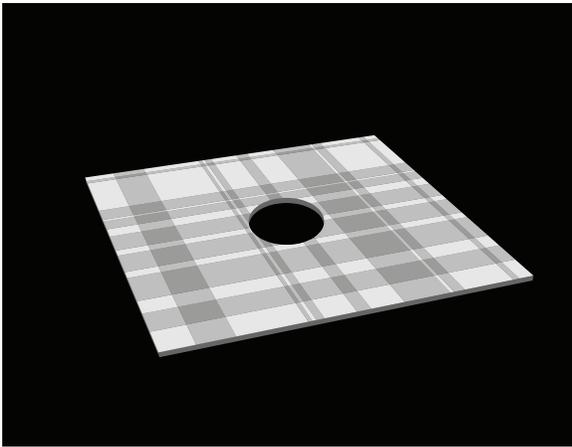
Figure 6. Sample stimulus. Subjects judged whether the hole on the surface was longer versus wider than a circle.

consisted of two blocks of 320 trials each, completed in a 1-hr session. This yielded 40 trials for each of the 16 conditions.

Subjects

Seventeen subjects participated in Experiment 1. Two were the authors. The others were naive to the purposes of the experiment and were paid for participating. All subjects had normal or corrected-to-normal vision. Subjects gave informed consent in accordance with a protocol approved by the IRB panel at the University of Pennsylvania.

Results

For each subject and condition, we estimated the height-to-width ratio of the ellipse that would be perceived to be a circular hole on the surface (see the Methods section). Figure 7 plots the mean height-to-width ratios, averaged

across subjects. The four panels correspond to the four spins tested. The graphs also plot predicted results in the extreme cases where judgments were based entirely on convergence (thin solid), skew (dashed), or foreshortening (dotted).

The data provide evidence that convergence, skew, and foreshortening all contributed. An ANOVA on the mean height-to-width ratios revealed significant main effects of the projected size of an image, $F(1, 240) = 104$, $p < .001$ (with smaller images seen as if more slanted), the slant of the object used to generate the prescaled image, $F(1, 240) = 161$, $p < .001$ (with more slanted objects seen as if more slanted), and the spin of the object, $F(3, 240) = 16$, $p < .001$ (with rotated objects appearing as if more slanted). There was also an interaction between the generating slant and spin, $F(3, 240) = 10.8$, $p < .001$. No other interactions were significant in the ANOVA, size/slant: $F(3, 240) = 1.7$, $p = .19$, *ns*; size/spin: $F(3, 240) = 2.2$, $p = .09$, *ns* (but see further analyses below). The main effect of projected size could be attributable to either convergence or skew cues because both predict greater slant for small images when spin is zero. However, the effect of size was also present in the 20° spin conditions when analyzed separately, $F(1, 48) = 15.3$, $p < .001$, indicating at least some contribution from convergence. The main effect of generating slant could be due to either skew or foreshortening. This effect was significant in the 0° spin conditions when analyzed separately, $F(1, 48) = 7.8$, $p = .007$, implying that at least some of this effect was due to foreshortening. However, both the main effect of spin and the interaction between slant and spin can only be attributed to the influence of skew because use of the other cues (convergence and foreshortening) predict that judgments would be invariant to spin.

Figure 8 shows four conditions that illustrate the interaction between image scaling and spin. For 0° spin conditions, objects appear more slanted in the minified 62° image (left) than in the magnified 75° slant image (right), as predicted by convergence cues, despite the fact that foreshortening is greater in the magnified condition. For
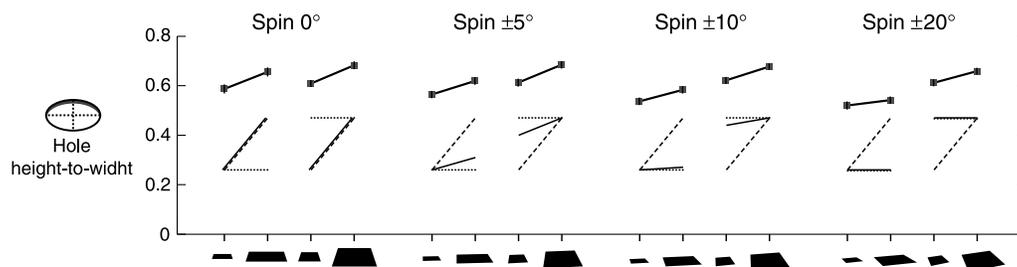

Figure 7. PSE results. Graphs plot the mean projected height-to-width ratios of holes that appeared circular, for different spins. The four points on each graph correspond to the slant and scaling conditions depicted on the x-axis (see Figure 5). Lines connect conditions that differ only by image scaling. The thin lines show the predicted results based on different computational assumptions: parallel sides (dashed), maximally orthogonal (solid), or isotropic (dotted). See the Appendix for derivations. Note that the predictions based on orthogonality change as a function of spin, and the observed PSEs change in a consistent direction.

Spin 0°

Minified
62° slant

Magnified
75° slant
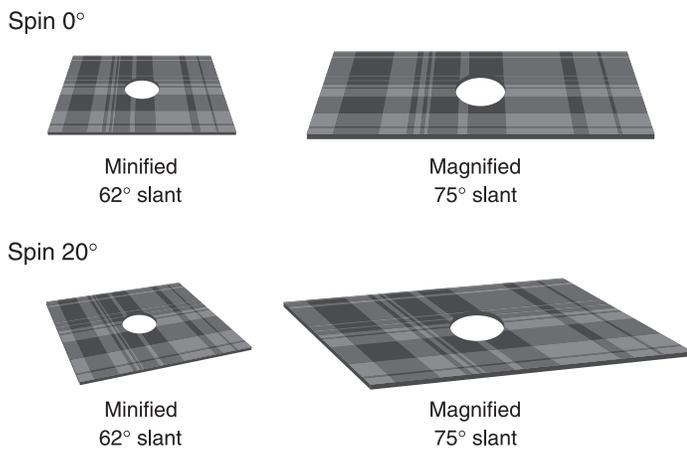
Spin 20°

Minified
62° slant

Magnified
75° slant

Figure 8. Four conditions that illustrate the interaction between image scaling and spin. Holes have aspect ratios equal to the mean observed PSEs for the conditions. When spin is 0° (top), the minified image of an object with low slant (top left) appears more slanted than the magnified image of an object with higher slant (top right), as predicted by the convergence cue. When spin is 20° (bottom), perceived slant is more dependent on the slant used to generate the image prior to scaling, such that the difference reverses: The minified image of the low-slant object (bottom left) appears less slanted than the magnified image of the higher slant object (bottom right).

20° spin conditions, the skew cue overcomes the conflicting convergence information, such that perceived slant is larger for the magnified 75° slant image (right) despite the comparative lack of convergence relative to the minified 62° slant image (left).

We performed a regression analysis to estimate the relative contributions of each cue, correlating the observed PSEs against the predictions shown in the right graphs of Figure 7. The resulting regression weights were as follows: $r = .18$ for convergence ($p = .04$), $r = .26$ for skew ($p = .02$), and $r = .16$ for foreshortening ($p = .09$, *ns*). These weights are generally consistent with the findings of the ANOVA, in that both convergence and skew showed significant nonzero weight, whereas foreshortening has smaller, marginal influence.

If skew information contributes to perceived slant, then conditions with skewed contours should be less affected by image scaling. The ANOVA on aspect ratios is not a sensitive test of this prediction because of individual differences in the sizes of the main effects. For a more sensitive test, we computed the differences in aspect ratios for pairs of conditions that differ only in image size, which provides a direct measure of the effect of image magnification or minification. The mean differences are shown in Figure 9. For both slant conditions, image scaling had less effect as spin increased, 75° slant: $F(3, 48) = 7.4$, $p < .001$; 62° slant: $F(3, 48) = 3.8$, $p = .02$. The modulation was more pronounced for the 75° slant conditions (Figure 9,

left), which could be due to the larger amount of skew in these images.

The holes that were perceived as circular had projected shapes that were taller than would be predicted by any cue. The direction of overall bias is consistent with perceptual underestimation of slant because circles on a surface with a low slant projected to taller ellipses in an image than circles at higher slants. A similar overall bias was observed in an experiment by Saunders and Backus (2006a) that measured perceived length-in-depth based on convergence, for monocular images of rectangles with 0° spin. Such biases could be due to conflicting information that would suggest frontal orientation, such as the absence of an accommodative gradient or the visible frame of the projection screen.

Judgments also showed less effect of image scaling than expected, even in the 0° spin conditions where convergence and skew cues do not conflict. The predicted PSEs change from .26 to .47, whereas the observed PSEs increased by only .07 on average, corresponding to a gain of .33. Other studies have similarly observed a smaller-than-predicted effect of size on perceived slant or depth from convergence (Saunders & Backus, 2006a; Smith, 1967; Tibau et al., 2001). One factor might have been the foreshortening of the elliptical hole itself, which could be used as a cue to surface slant and would indicate differing slants from trial to trial. This cue is not informative with respect to our task, and its effect would have been uniform across conditions, but it might have reduced the effects of the convergence and skew cues if weight were given to it. The smaller-than-predicted modulation by image size could also be a consequence of perceptual compression in depth
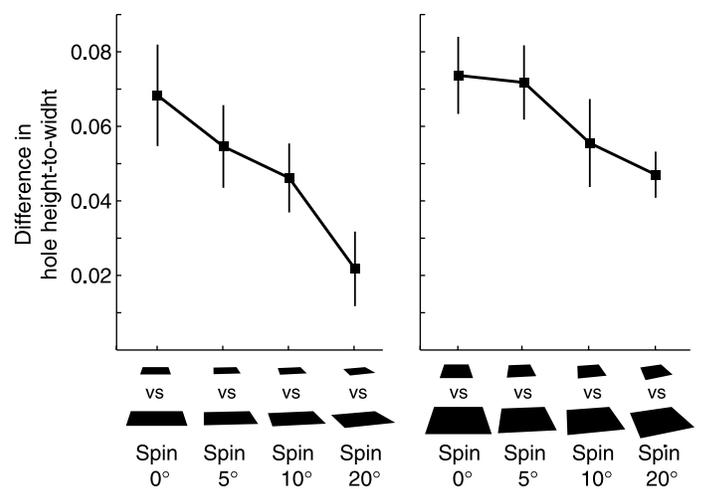

Figure 9. Effect of image magnification or minification on the shape of the hole perceived to be circular. The left graph plots differences between PSE aspect ratios for magnified and unscaled images of small objects with 75° slant, with various spins. The right graph plots the differences between PSE aspect ratios for unscaled and minimized images of large objects at 62° slant, with various spins.
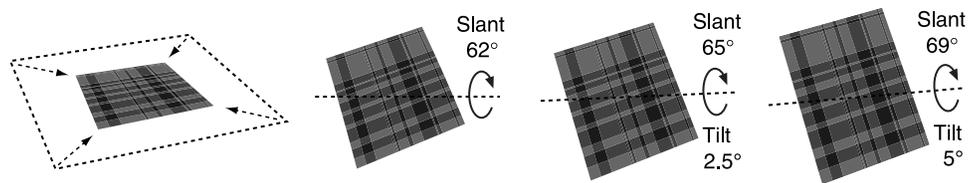
Figure 10. If tilt is allowed to vary, there exist interpretations of scaled images that have less conflict between parallel-sides and orthogonal interpretations.

(see Saunders & Backus, 2006a) or simply reflect errors in interpreting convergence and skew information.

Our analysis assumed that the tilt of the surfaces was accurately perceived to be vertical. The shape of the hole and its extrusion edges would have encouraged a percept of vertical tilt. However, there is also reason to expect some bias. Saunders and Knill (2001) found that the perceived tilt of slanted symmetric figures was biased depending on their spin. For spins of 15° or 30°, the tilt bias was about 3°. As illustrated in Figure 10, the maximally orthogonal interpretations of our cue-conflict stimuli change depending on the assumed direction of tilt. If tilt biases were in the direction observed by Saunders and Knill (in the direction of spin, as in Figure 10), orthogonal interpretations would have slants closer to that specified by convergence (compared to that with unbiased tilt). Therefore, to the extent that tilt is not perceived to be vertical, our analysis may have underestimated the contribution of skew relative to convergence.

Figure 11 plots the mean JND thresholds, averaged across subjects, expressed as Weber fractions. The most pronounced effect was that judgments were more consistent (lower JNDs) for large objects than for objects with identically shaped but smaller projected contours. This effect was revealed in an ANOVA as a main effect of projected size, $F(1, 240) = 24$, $p < .001$. No other effects or interactions were significant. Saunders and Backus (2006a) observed an improvement with projected size on the ability to discriminate length-in-depth of slanted

rectangles based on contour information and were able to model the results with a Bayesian ideal observer for slant-from-convergence that incorporates noise in image measures of orientation.

Discussion

Our conditions dissociated convergence, skew, and foreshortening cues to slant, and the results provide evidence that all three contribute to slant perception. Magnifying or minifying a projected contour affected judgments, in the direction expected based on convergence information. Image scaling had less effect for skewed projected contours, as expected if skew information also contributes, and the addition of skew tended to increase perceived slant overall. The overall foreshortening of the projected contours also contributed, as evidenced by the effect of generating slant when convergence and skew cues were matched (0° spin conditions in Figure 7).

The foreshortening of the projected contours had a relatively weak influence compared to that of convergence and skew. This is consistent with our attempt to minimize this cue in the texture and with earlier results. Braunstein and Payne (1969) observed that varying the ratio of vertical to horizontal spacing of a rectangular grid texture (with 0° spin), which changes foreshortening of texture elements, had comparatively less effect on slant judgments
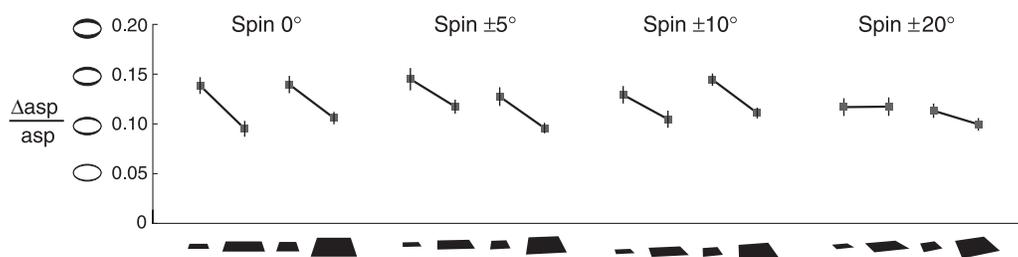

Figure 11. JND results. For each subject and condition, we computed the difference between the aspect ratios corresponding to the PSE and the 75% point of the psychometric function (∆asp) and then divided it by the PSE aspect ratio to obtain a Weber fraction (∆asp/asp). The graphs plot the mean Weber fractions averaged across subjects. Conditions are arranged in the same way as in Figure 7. The icons to the left of the y-axis graphically depict the range of aspect ratios corresponding to a given Weber fraction.

than varying the amount of convergence. Saunders and Backus (2006a) similarly found that foreshortening of trapezoidal projected contours had little effect on perceived slant. In the present experiment, foreshortening might appear to have had larger effect when the square was rotated away from the tilt direction: Results for 20° spin conditions are closer to the isotropic predictions. However, this interaction can be attributed to the influence of skew information. When skew is taken into account, the residual effect attributable to foreshortening was not significant.

Whereas no previous studies have attempted to distinguish the contributions of convergence and skew as we do here, an experiment by Tibau et al. (2001) did test stimuli within which these cues conflicted. They measured slant judgments for scaled images of surfaces textured by rectangular grids with varying spins. In most conditions, the projected textures provided conflicting convergence and skew cues, as in our experiment. Tibau et al. described their stimuli in terms of the surface slant and focal distance used to generate the perspective projection and viewing distance of the observer. The ratio of focal distance to viewing distance is equivalent to the amount of image magnification relative to an accurate perspective projection. For conditions that differed only in viewing distance (50 vs. 100 cm), slant judgments differed by approximately 2–5° on average. We computed the difference expected based on convergence information for their near and far conditions to be 15–17°. Thus, the effect of image scaling on judgments in the Tibau et al. study was about 20% of the magnitude predicted by convergence cues, which is comparable to what we observed. Tibau et al. suggest that this modest effect of viewing distance (or image scaling) is due to the influence of texture foreshortening. However, they averaged across spin when analyzing the effect of image scaling; hence, one cannot determine the relative contributions of foreshortening and skew. Given our results, we suspect that skew was the larger factor. The smaller-than-predicted effect of image scaling could also be due to the influence of other absent or conflicting depth cues (e.g., accommodative gradient) or some a priori bias (Saunders & Backus, 2006a).

Conflicting convergence and skew cues, as tested here, are often present when viewing photographs or perspective pictures. A perspective image accurately reproduces a view of a scene only when it is viewed from a particular location—the picture's *center of projection*. If viewed from a station point farther or nearer than the center of projection, the optic array presented by a picture is generally no longer consistent with the depicted scene (Farber & Rosinski, 1978; Nicholls & Kennedy, 1993; Sedgwick, 1980, 1991). The question of why 3D structure in pictures does not appear more noticeably distorted has been a topic of considerable debate (e.g., see Koenderink, van Doorn, Kappers, & Todd, 2004; Kubovy, 1986). In the case of pictures viewed obliquely, there is evidence that information about the slant of the picture surface itself

allows the visual system to compensate for viewing angle (Goldstein, 1987, 1988; Halloran, 1993; Perkins, 1974; Rosinski et al., 1980; Vishwanath, Girshick, & Banks, 2005; Wallach & Marshall, 1986). However, knowledge about the picture surface is not informative about whether viewing *distance* is consistent with a picture's center of projection. For example, same-sized photographs taken with telephoto or wide-angle lens require different viewing distances to be optically correct; thus, even if the size and distance of a photograph are known, this knowledge is not sufficient to accurately interpret the projected image.

In the case of scenes containing rectangular surfaces, we can be specific about the conflict introduced by image rescaling: The 3D structure specified by convergence under an assumption of parallelism changes with size or viewing distance, whereas the 3D interpretation based on an assumption of either orthogonality or isotropy is relatively unaffected (see Figure 4). Thus, if perspective cues in a scaled picture are interpreted naturally, that is, as if the picture were a window onto the scene, then the amount of perceptual distortion would depend on the relative contributions of scale-dependent and scale-independent cues during perception. Our results indicate that both scale-dependent information from convergence and scale-invariant information from skew and foreshortening are used. Thus, one would expect neither complete invariance to changes in viewing distance nor distortions as large as predicted by use of convergence alone.

Previous discussions of robustness of perception in pictures have often failed to distinguish between pictorial cues that are and are not affected by changes in viewpoint. Here, we have identified two 3D cues that are differentially affected by changes in image size or viewing distance: convergence and skew. We have further shown that both these cues make measurable contributions to perceived slant of rectangular surfaces. Thus, when considering the problem of perceiving pictures from various viewpoints, it is important to take into account the information provided by available pictorial cues.

Conclusion

Two computational assumptions that could be used to perceive slant-in-depth from linear perspective are that converging lines in a projected image are parallel on a surface and that skewed angles formed by intersecting axes and edges in an image are orthogonal on a surface. Our results indicate that both of these constraints are utilized. Scaling a perspective image changes its perceived slant in the direction of the parallel-sides interpretation, even when conflicting skew information is present. However, image scaling has a smaller effect when skew cues conflict with convergence, indicating a tendency to perceive objects as

orthogonal. We conclude that perception of 3D structure from linear perspective makes use of both parallelism and orthogonality as computational constraints.

Appendix A

This appendix describes how we predicted observers' "apparently circular" settings (Figure 7) based on different possible constraints: maximizing parallelism, orthogonality, or isotropy of the 3D object, given the images. The generation of stimuli is described in the Methods section. The orientations of internal edges within the surface texture were determined by the orientations of the outer edges, and the spacing between internal edges was randomized. Therefore, the stimuli can be characterized (up to the random spacing) by their projected contours. These contours were quadrilaterals in the image plane. When back-projected onto a slanted planar surface, the result is another quadrilateral. The overall size of a back-projected quadrilateral depends on the distance of the slanted surface, and its shape depends solely on surface orientation (slant and tilt) relative to the viewer.

To derive predictions based on parallelism, we first computed the surface slant for which the back-projected contour had parallel sides. We define slant as the angular difference between the surface normal and the line of sight to the geometric center of the projected figure (where the elliptical hole was positioned). Slant was assumed to vary only around a horizontal axis (vertical tilt direction), in agreement with stimulus generation. For our stimuli, there was always a slant for which both pairs of sides of the quadrilateral were parallel, S_{parallel}. We then computed the aspect ratio of a projected circular hole on a surface having slant S_{parallel}. The results are shown by the dashed lines in Figure 7. These aspect ratios were also used for the regression analysis (see the Results section).

To derive predictions based on an orthogonality constraint, we used a similar procedure as for parallelism, except that we sought the back-projected quadrilateral that was closest to being rectangular. In many conditions, there was no slant for which the back-projected quadrilateral was perfectly rectangular. As a measure of overall deviation from orthogonality, we used the RMS deviation from orthogonality at each of the corners: $\sqrt{(\sum (A_k - 90°)^2/4)}$, where A_k are the four internal angles of the back-projected quadrilateral. For the projected contours in each condition, there was some slant (S_{ortho}) that minimized this orthogonality measure. The minimal RMS deviation from orthogonality varied between 0° and 1.7°. The predicted aspect ratio for a given condition (thin solid lines in Figure 7) was the aspect ratio of a projected circular hole on a surface having slant S_{ortho}.

Similarly, for the isotropy constraint, we computed the slant for which the back-projected contour was most isotropic (circularly symmetric) and then computed the

projected aspect ratio of a circular hole with that slant (dotted lines in Figure 7). As a measure of deviation from isotropy, we used the ratio of the moments of inertia of a back-projected quadrilateral, with a ratio of 1 indicating isotropy. For our stimuli, there was always some slant for which the back-projected quadrilateral was isotropic by this measure (i.e., had equal moments of inertia).

References

Andersen, G. J., Braunstein, M. L., & Saidpour, A. (1998). The perception of depth and slant from texture in three-dimensional scenes. *Perception, 27,* 1087–1106. [PubMed]

Attneave, F., & Olson, R. K. (1966). Inferences about visual mechanisms from monocular depth effects. *Psychonomic Science, 4,* 133–134.

Banks, M. S., & Backus, B. T. (1998). Extra-retinal and perspective cues cause the small range of the induced effect. *Vision Research, 38,* 187–194. [PubMed]

Braunstein, M. L., & Payne, J. W. (1969). Perspective and form ratio as determinants of relative slant judgments. *Journal of Experimental Psychology, 81,* 584–590.

Buckley, D., Frisby, J. P., & Blake, A. (1996). Does the human visual system implement an ideal observer theory of slant from texture? *Vision Research, 36,* 1163–1176. [PubMed]

Clark, W. C., Smith, A. H., & Rabe, A. (1955). Retinal gradient of outline as a stimulus for slant. *Canadian Journal of Psychology, 9,* 247–253. [PubMed]

Clark, W. C., Smith, A. H., & Rabe, A. (1956). The interaction of surface texture, outline gradient, and ground in the perception of slant. *Canadian Journal of Psychology, 10,* 1–8. [PubMed]

Farber, J., & Rosinski, R. R. (1978). Geometric transformations of pictured space. *Perception, 7,* 269–282. [PubMed]

Freeman, R. B., Jr. (1966a). Absolute threshold for visual slant: The effect of stimulus size and retinal perspective. *Journal of Experimental Psychology, 71,* 170–176. [PubMed]

Freeman, R. B., Jr. (1966b). Function of cues in the perceptual learning of visual slant: An experimental and theoretical analysis. *Psychological Monographs, 80,* 1–29. [PubMed]

Gillam, B. J. (1968). Perception of slant when perspective and stereopsis conflict: Experiments with aniseikonic lenses. *Journal of Experimental Psychology, 78,* 299–305. [PubMed]

Goldstein, E. B. (1987). Spatial layout, orientation relative to the observer, and perceived projection in pictures viewed at an angle. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 256–266. [PubMed]

Goldstein, E. B. (1988). Geometry or not geometry? Perceived orientation and spatial layout in pictures viewed at an angle. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 312–314. [PubMed]

Griffiths, A. F., & Zaidi, Q. (2000). Perceptual assumptions and projective distortions in a three-dimensional shape illusion. *Perception, 29,* 171–200. [PubMed]

Halloran, T. O. (1993). The frame turns also: Factors in differential rotation in pictures. *Perception & Psychophysics, 54,* 496–508. [PubMed]

Knill, D. C. (1998). Ideal observer perturbation analysis reveals human strategies for inferring surface orientation from texture. *Vision Research, 38,* 2635–2656. [PubMed]

Koenderink, J. J., van Doorn, A. J., Kappers, A. M., & Todd, J. T. (2004). Pointing out of the picture. *Perception, 33,* 513–530. [PubMed]

Kubovy, M. (1986). *The psychology of perspective and renaissance art.* Cambridge: Cambridge University Press.

Nicholls, A. L., & Kennedy, J. M. (1993). Angular subtense effects on perception of polar and parallel projections of cubes. *Perception & Psychophysics, 54,* 763–772. [PubMed]

Perkins, D. N. (1974). Compensation for distortion in viewing pictures obliquely. *Perception & Psychophysics, 14,* 13–18.

Rosenholtz, R., & Malik, J. (1997). Surface orientation from texture: Isotropy or homogeneity (or both)? *Vision Research, 37,* 2283–2293. [PubMed]

Rosinski, R. R., Mulholland, T., Degelman, D., & Farber, J. (1980). Picture perception: An analysis of visual compensation. *Perception & Psychophysics, 28,* 521–526. [PubMed]

Saunders, J. A. (2003). The effect of texture relief on perception of slant from texture. *Perception, 32,* 211–233. [PubMed]

Saunders, J. A., & Backus, B. T. (2006a). The accuracy and reliability of perceived depth from linear perspective as a function of image size. *Journal of Vision, 6*(9):7, 993–954, http://journalofvision.org/6/9/7/, doi:10.1167/6.9.7. [PubMed] [Article]

Saunders, J. A., & Backus, B. T. (2006b). Perception of surface slant from oriented textures. *Journal of Vision, 6*(9):3, 882–897, http://journalofvision.org/6/9/3/, doi:10.1167/6.9.3. [PubMed] [Article]

Saunders, J. A., & Knill, D. C. (2001). Perception of 3D surface orientation from skew symmetry. *Vision Research, 41,* 3163–3183. [PubMed]

Sedgwick, H. A. (1980). The geometry of spatial layout in pictorial representation. In M. A. Hagen (Ed.), *The perception of pictures* (vol. 1, pp. 33–90). New York: Academic Press.

Sedgwick, H. A. (1991). The effect of viewpoint on the virtual space of pictures. In S. R. Ellis, M. K. Kaiser, & A. C. Grunwald (Eds.), *Pictorial communication in virtual and real environments* (pp. 460–469). London: Taylor & Francis.

Smith, A. H. (1967). Perceived slant as a function of stimulus contour and vertical dimension. *Perceptual and Motor Skills, 24,* 167–173.

Stavrianos, B. K. (1945). The relation of shape perception to explicit judgments of inclination. *Archives of Psychology, No. 296.*

Stevens, K. A. (1983). Slant–tilt: The visual encoding of surface orientation. *Biological Cybernetics, 46,* 183–195. [PubMed]

Tibau, S., Willems, B., Van den Bergh, E., & Wagemans, J. (2001). The role of the centre of projection in the estimation of slant from texture of planar surfaces. *Perception, 30,* 185–193. [PubMed]

Todd, J. T., Thaler, L., & Dijkstra, T. M. (2005). The effects of field of view on the perception of 3D slant from texture. *Vision Research, 45,* 1501–1517. [PubMed]

Vishwanath, D., Girshick, A. R., & Banks, M. S. (2005). Why pictures look right when viewed from the wrong place. *Nature Neuroscience, 8,* 1401–1410. [PubMed]

Wallach, H., & Marshall, F. J. (1986). Shape constancy in pictorial representation. *Perception & Psychophysics, 39,* 233–235. [PubMed]

Yang, T., & Kubovy, M. (1999). Weakening the robustness of perspective: Evidence for a modified theory of compensation in picture perception. *Perception & Psychophysics, 61,* 456–467. [PubMed]