# Looking at Perspective Pictures From Too Far, Too Close, and Just Right

Igor Juricevic and John M. Kennedy University of Toronto, Scarborough

A central problem for psychology is vision's reaction to perspective. In the present studies, observers looked at perspective pictures projected by square tiles on a ground plane. They judged the tile dimensions while positioned at the correct distance, farther or nearer. In some pictures, many tiles appeared too short to be squares, many too long, and many just right. The judgments were strongly affected by viewing from the wrong distance, eye height, and object orientation. The authors propose a 2-factor angles and ratios together (ART) theory, with the following factors: the ratio of the visual angles of the tile's sides and the angle between (a) the direction to the tile from the observer and (b) the perpendicular, from the picture plane to the observer, that passes through the central vanishing point.

Keywords: spatial perception, perspective, constancy, picture perception

When we walk in front of a masterpiece such as Raphael's "School of Athens," showing scholars discussing in a great hall, we are entertaining a scene drawn in perspective, a format invented as a crowning glory of the intellectual advances of the 15th century. But even in the time of its invention, those adept in linear perspective, such as Leonardo da Vinci, admitted it created a mysterious mixture of acceptable and distorted effects. That is, when looking at some pictures drawn with perfect adherence to perspective, observers were struck by areas in which the picture looked realistic (perceptual constancy) and areas in which the picture looked distorted. Here, we respond to the mystery with a new theory about the visual angles of the sides of an object and, revealingly, the angle between two directions: (a) the direction to the object from the observer.

In our experiments, we examine a problem that originated during the Renaissance—the problem of viewing in perspective, particularly of viewing pictures from different distances. This problem has been the subject of heated debate in experimental psychology, developmental psychology, cross-cultural psychology, philosophy, semiotics, engineering, physics, and art history. There are few topics in psychology on which so much has been written within psychology and outside it, for centuries, by many of the best minds in scholarship. Is perspective a cultural convention? Is it readily used by perception? This problem is at the core of theories of constancy, ambiguity of our sensory input, and Gibsonian realism—in other words, the long history of research on perception. Furthermore, perspective displays are very often used as surrogates for real-world stimuli in many kinds of experiments, video displays, and flying and driving simulators.

Can perceptual constancy be reconciled with its opposite concept, distortion (Koenderink & van Doorn, 2003; Kubovy, 1986; Sedgwick, 2003)? Our aim was to study pictures and perspective, but ultimately we ask about a general account of perspective in vision. The implications are many—not just for psychology but for photography, movies, and art history for example.

Figure 1 is a perspective picture of tiles on a ground plane (Gibson, 1966). The tiles project many different shapes. Do they all suggest square tiles? No. Some look far from square. But why? To answer, let us consider the essence of linear perspective and then vision's reaction to it.

Linear perspective dictates how a scene should be depicted from a particular vantage point, with the picture set at a particular location. When viewing a picture, vision's task is "inverse projection" (Niall, 1992; Niall & Macnamara, 1989, 1990; Norman, Todd, Perotti, & Tittle, 1996; Wagner, 1985). Every perspective picture has a correct viewing distance from which the perspective projection was determined. Call this "the artist's (or the camera's) distance." Strictly speaking, if a picture is viewed from farther than the artist's distance, and if vision followed perspective exactly, then the pictured scene should expand in depth. From double the artist's distance, what was originally depicting a set of square tiles should be seen as depicting elongated tiles, twice as long as broad (Kennedy & Juricevic, 2002; La Gournerie, 1859; Pirenne, 1970). Similarly, halve the viewing distance and the tiles should appear stubby, cut in half. There is a simple reason for the multiplication. Consider a point on the picture projected to a viewer's vantage point; it will be a projection of a point on the ground plane. Slide the viewer back from the picture plane to double the viewing distance, and, by similar triangles, the point projected on the ground plane must also slide back, away from the picture plane, and its distance must also double (see Figure 2).

It is well-known that a perspective picture, such as a photograph, can be viewed from varied distances without all parts of the

Igor Juricevic and John M. Kennedy, Department of Life Sciences-Psychology, University of Toronto, Scarborough, Toronto, Ontario, Canada.

We acknowledge helpful comments from Dejan Todorovic, Michael Kubovy, Paul Milgram, Mark Schmuckler, and Matthias Niemeier and the influence of Rudolf Arnheim, now in his 102nd year, the major theorist of the psychology of art of the 20th century; and John Willats, the foremost theorist of perspective drawing in children, who passed away on April 12, 2006.

Correspondence concerning this article should be addressed to John M. Kennedy, Department of Life Sciences-Psychology, University of Toronto, Scarborough, 1265 Military Trail, Toronto, Ontario, M1C 1A4 Canada. E-mail: kennedy@utsc.utoronto.ca



*Figure 1.* A perspective picture of a series of square tiles on a ground plane. The picture is rendered in one-point perspective, meaning that the edges of the tiles are either orthogonal to the picture plane (e.g., the right and left edges) or parallel to the picture plane (i.e., the closer and farther edges). The central vanishing point for all tiles is also indicated.

picture shrinking and expanding in the fashion we have just described. So, vision does not use exact perspective. Indeed, some theories have gone so far as to suggest that perceptual constancy holds across perspective changes, and vision can ignore perspective's multiplication effects by means of many subterfuges, top-down or bottom-up, conscious or unconscious (Gibson, 1979, 1947/1982; Koenderink, van Doorn, Kappers, & Todd, 2001; Kubovy, 1986; Pirenne, 1970; for discussion, see Rogers, 1995, 2003).

It is less widely appreciated that when perspective effects become extreme, vision does become wildly distorted (Kennedy & Juricevic, 2002; Kubovy, 1986). The margins of wide-angle pictures induce vivid perceptual effects if the pictures are viewed from afar, that is, much farther than the artist's distance. Just so, tiles in the very bottom margins of Figure 1 often appear much too long to be square. It is because these vivid perceptual effects are often most pronounced in the periphery of a perspective picture that they are called *marginal distortions*. However, as will become evident, central distortions may arise from extensive foreshortening.

Marginal distortions caused artists to use rules of thumb such as "paint only narrow-angle views" (say 12° on either side of the vanishing point) when depicting a scene and caused camera makers to adopt lenses that only take in narrow visual angles. Central distortions lead artists to hide distant squares in tiled piazza pictures behind foreground objects such as people.

Our goal was to reconcile distortion and constancy. To begin, we contend that many extant theories can explain one effect, not both.

To relate the different major theories, we describe a single "pseudoperspective" function (one related to perspective geometry), which deals with average tile length in a picture. Then, after Experiment 1, we examine the angles and ratios together (ART) theory, which reaches beyond average tile lengths, and reconciles distortions and constancy. The ART theory treats individual tiles and relates the ratio of the visual angles projected by sides of each tile to its direction from its central vanishing point.

For the first major theory, consider "projective" theories. In this approach, an observer perceives the width and length (i.e., the z dimension, or depth) of each tile in Figure 1 according to the laws

of projective (perspective) geometry. They require perceived elongation of depth when an observer is farther than the artist's distance, compression when an observer is too close (Kennedy & Juricevic, 2002). Call the ratio of the depth to the width of each tile its "relative depth." Their function is perceived relative depth = k(correct relative depth) × (observer's distance)<sup>*d*</sup>/(artist's distance)<sup>*j*</sup>, where k = 1, d = 1, and j = 1.

The ratio of observer's and artist's distance is directly linearly related to perceived relative depth, as in projective geometry.

Many approaches can be expressed with similar pseudoperspective functions. "Perceived relative depth" is a tile's perceived depth divided by perceived width. "Correct relative depth" is the actual relative depth, and for squares is 1. This term is multiplied by a constant "k," which is 1 if the tiles are all perceived as squares at the artist's distance. If k < 1, then the tile appears compressed, and if k > 1, then the tile is elongated. Perceived depth in pictures is often flattened (by 15%, e.g., Koenderink & van Doorn, 2003), and it is possible that k is the only term needed to account for this.

An exponent, "d," modifies "observer's distance," the physical distance of the observer from the picture surface. Doubling the distance doubles perceived relative depth if the exponent d = 1. In compensation theories, "observer's distance" does not affect depicted extents and has an exponent of d = 0 (so this term in the equation is simply equal to 1). Larger exponents increase the effect of the observer's distance.

"Artist's distance" is the distance used to create the perspective picture and is the correct distance from which to view it. In correct perspective, doubling the observer's distance should double the apparent depth of the tiles, so an artist's distance half the observer's distance could make the tiles seem especially long. To reflect this, artist's distance is in the denominator of the equation (i.e., dividing by one half increases apparent size). Effects of artist's distance may not be exactly one-to-one, so it is given an exponent "j." In compensation theories, j = 0 and does not affect "perceived



*Figure 2.* Observer 1 (O<sub>1</sub>) looking at Point C at a distance of D<sub>1</sub> from the picture plane (P). Point C is a projection of Point G<sub>1</sub> on the ground. The triangle defined by the observer and the projected Point G<sub>1</sub> ( $\Delta O_1 D_1 G_1$ ) and the triangle defined by Point C on the picture and Point G<sub>1</sub> ( $\Delta CPG_1$ ) are similar triangles. As such, the distance from the picture plane to the observer (D<sub>1</sub>) is geometrically similar to the distance from the picture plane to the point on the ground plane (G<sub>1</sub>). Doubling the observer's distance (to D<sub>2</sub>) will therefore double the distance of the point projected on the ground (to G<sub>2</sub>).

relative depth." The larger the *j*, the greater the effect of movements away from the "artist's distance."

The size of j depends on the units used for the pseudoperspective function. This is simply a mathematical consequence of exponents. So, for convenience, j will always be calculated here with respect to an artist's distance less than 1 unit (i.e., less than 1 m), and roughly arm's length or within.

With respect to the projective theories, this approach could fail on two accounts. First, distortions are predicted throughout the picture rather than selectively for some tiles. Second, an incorrect amount of distortion is predicted in many situations.

Next, consider the "compensation" argument that the visual system determines the artist's distance from information present within the picture and adjusts for this when undertaking inverse projection. Compensation predicts that regardless of the position of the observer, this ratio is perceived as constant, which can be summarized as perceived relative depth = k(correct relative depth) × (observer's distance)<sup>*d*</sup>/(artist's distance)<sup>*j*</sup>, where k = 1, d = 0, and j = 0.

Marginal distortions, according to compensation theories, occur when the process of compensation breaks down. But there is, as yet, no accepted explanation of why this breakdown in apparent depth constancy occurs in the periphery of pictures of ground planes (though see Kubovy, 1986; and Yang & Kubovy, 1999, for excellent discussions of apparent angular distortions of cubes). Furthermore, compensation theories make no allowance for distortions that may occur in central regions where there is extreme foreshortening.

In the "invariant" approach, Gibson (1979) argued that perception is governed by contents of the optic array, especially one projected by the ground plane. We follow him on this point but argue that invariants are only one kind of function carrying the optic array's information. For Gibson, a spatial property (e.g., a certain size or certain shape) can produce an optic invariant that is specific to that property. For example, if a pole on the ground plane has a top just below the horizon line, and another pole's top is above the horizon, then the one above is taller.

Many invariants remain no matter what in which direction the observer moves in front of the picture; for example, a pole's top is always depicted above or below the horizon. Invariant relations of this type (call them "horizon-ratio" type) are present regardless of the observer's distance from the picture. Hence, their function is identical to compensation's: perceived relative depth = k (correct relative depth) × (observer's distance)<sup>*d*</sup>/(artist's distance)<sup>*j*</sup>, where k = 1, d = 0, and j = 0.

As with compensation, invariants of the horizon-ratio type are unable to account for constancy and distortions within one picture. The invariants are present in both the apparently distorted area of the picture and its perceptually constant neighbor.

The "compromise" approach proposes effects from the flatness of the picture surface. Perceived flatness diminishes perceived tile proportions (Koenderink & van Doorn, 2003) and may make the ground appear sloped, that is, closer to the slant of the picture surface (Miller, 2004; Rosinski & Farber, 1980; Rosinski, Mulholland, Degelman & Farber, 1980; Sedgwick & Nicholls, 1993). In its pseudoperspective function, *k* is less than 1, shrinking as the picture surface is made more salient, for example, by adding texture (Sedgwick, 2001) or by instructing the observer to pay attention to the surface (Miller, 2004): perceived relative depth =

k(correct relative depth) × (observer's distance)<sup>d</sup>/(artist's distance)<sup>j</sup>, where 0 < k < 1, d = 1, and j = 1.

Any compromise should occur across the entire picture because information for depth and flatness is present across the entire picture. However, this does not occur when peripheral areas show distinctive distortions (Niederée & Heyer, 2003), for example, if they look full of especially elongated tiles.

Finally, an "approximation" approach supports the argument that vision's inverse projection is just "ballpark-perspective"; it may work well at moderate distances but veers from proper perspective in less-restricted tests, for example, a wide range of artist's distances.

Cross-scaling theory (Smallman, Manes, & Cowen, 2003; Smallman, St. John, & Cowen, 2002) is a useful example of a theory in which an approximation approach is used. In Figure 1, the tiles have two sets of parallel edges, one running left to right, the other in depth. The lines in the picture are parallel from left to right and converge from bottom to top. The length of a line projected onto the picture surface by a left-to-right tile edge decreases linearly as the depth to the tile increases. In contrast, the converging lines decrease in length as a square function of each tile's depth. This true mathematical perspective is not used by vision, the cross-scaling model proposes. Rather, vision "ballparks" that the lines projected by both the left-to-right tile edges and the tile edges in depth decrease linearly with depth. Differences between the ballpark function and true perspective's quadratic function become sizable in the far distance.

Unfortunately, cross-scaling cannot account for both constancy and distortion. All the tiles in a row, such as the third row from the bottom in Figure 1, should appear the same. If the center tile appears square (perceptual constancy), while the leftmost tile clearly does not (marginal distortion), then this contradicts cross-scaling.

However, we believe the approximation approach holds the most promise for a theory of vision's use of perspective. Cross-scaling is simply the wrong theory. Here, vision's approximation is shown to depart sizably from perspective proper by setting the observer, like Goldilocks, too close to the picture (artist's distance large), too far from the picture (artist's distance small), and just right, which, in the present study, is a picture with an artist's distance of 0.36 m.

We varied artist's distance in Experiment 1 and sought a pseudoperspective function, which looks for constancy and distortion in one and the same picture. We follow this approach by introducing ART theory factors governing regions of constancy and distortion.

#### Experiment 1

## Method

Subjects. Twelve first-year students (7 women, with mean age = 19.9, SD = 1.9) participated. All subjects were psychology students from the University of Toronto, had normal or corrected-to-normal vision (self-reported), and were naïve about the purpose of the study.

*Stimuli.* Perspective pictures were projected as panoramic images onto a large translucent back-projection screen using an EPSON PowerLite 51c LCD projector (Model EMP-51). The resolution of the projector was  $800 \times 600$ . Projected, each picture measured 0.64 m (high)  $\times$  1.28 m (wide) and subtended 79.3°  $\times$  121.3° of visual angle at a distance of 0.36 m. The stimuli were presented to the limits of fidelity. That is, the farthest row of tiles shown to subjects (in this case, row 9) was chosen because it was the last row for which tile proportions could be resolved distinctly from tile proportions in the next possible row.

The perspective pictures each depicted 153 square tiles (17 columns  $\times$  9 rows) on a ground plane (see Figures 1 and 3). The rows were numbered from 1 (near) to 9 (far), beginning with the row depicted closest to the observer (i.e., the row that projects to the lowest part of the picture plane). The columns were also numbered from 1 (center) to 9 (left), beginning with the center column (1 center) and increasing for each column to the left (9 left). Columns to the right of the center column were not used in the experiment because they are symmetrical with those to the right. Inspection and informal testing found no differences in the visual response between right and left stimuli. (Figures 4, 6, 9, and 11 in the *Results* sections below are shown symmetrical for clarity of presentation.) Any tile's position can, of course, be described by giving the tile's row and column number.

The tiles were depicted in one-point perspective; that is, the two receding edges of each tile were perpendicular to the picture plane, and the other two were parallel. Oblique lines depicting the receding edges converged in the picture to a single, central vanishing point. The width of the tiles was such that the closest edge of the tile in row 1 near, column 1 center subtended 6.1° of visual angle when viewed at a distance of 0.36 m.

The tiles were depicted using seven different artists' distances. The distances were all on the normal from the horizon. The observer's vantage point was in front of the central column of tiles (column 1), and varied in its distance from the picture plane. The seven distances varied by 0.09 m and were at 0.09, 0.18, 0.27, 0.36, 0.45, 0.54, and 0.63 m.

The tiles tested were those located in the factorial combinations of rows 1, 3, 5, 7, and 9 and columns 1, 3, 5, 7, and 9. They were indicated to the subjects by using bold lines (three times the thickness of the other lines in the picture) to depict the closest and rightmost edge of the tiles. In each picture, only one tile was depicted with bold edges.



*Figure 3.* Seven different perspective pictures of the same set of square tiles. The pictures are all rendered using different artist's distances. The artist's distance for each picture (in meters) refers to when the picture is presented at a scale of 0.64 m (high)  $\times$  1.28 m (wide). The artist's distances are (a) 0.09, (b) 0.18, (c) 0.27, (d) 0.36, (e) 0.45, (f) 0.54, and (g) 0.63 m.

The 25 different tiles tested were factorially combined with the seven artist's distances to produce 175 pictures that were used in the experiment.

*Procedure.* Each subject was tested individually and instructed to judge the length of the right edge of an indicated tile (one of the converging lines) relative to the closest edge of the tile (a horizontal line). They were told that the judgment was relative to the closest edge, set at 100 units. Thus, if the right edge appeared to be as long as the closest edge, then the subject would judge it to be 100 units. If it appeared longer or shorter, then the subject would judge its length proportionately.

The subject viewed each picture monocularly. To control the position from which the subject viewed the picture, a bar parallel to the floor was positioned 0.36 m from the picture plane. For a subject using the right eye, the bar was positioned in front of the picture plane, on the right side of the picture. The end of the bar was at the height of the horizon in the picture, approximately 3 cm to the right of the central vanishing point. The end of the bar touched the subject's temple at eye height, just to one side of the corner of the right eye. Subjects were instructed to maintain the temple's contact with the bar. For subjects using their left eye, the position of the bar was reversed. In this way, the subject was positioned so that her or his eye was in front of the central vanishing point, in line with the foot of the normal, and he or she was free to turn their eyes and head. Each picture was presented with no time limit. Once the subject made her or his judgment, the screen went black for 2 s, and the next picture was displayed.

Subjects were asked to judge the length of the tile, not the lines in the picture. They were reminded that, in a picture, a mountain off in the distance may be drawn with smaller lines than a person who is nearby.

### Results

*Dependent measure.* The dependent measure was perceived relative depth, obtained by dividing the responses by 100. Tiles longer than their width have ratios greater than 1, shorter less than 1, and tiles perfectly square 1.

To fit the function perceived relative depth =  $k(\text{correct} \text{ relative depth}) \times (\text{observer's distance})^1/(\text{artist's distance})^j$ , a choice has to be made as to the exponent for observer's distance. Fortunately, for theories in which the artist's distance affects perceived relative depth, the observer's distance has an exponent of 1 (i.e., projective and compromise approaches). We may set aside for the moment theories in which the exponent on observer's distance should be set to 0 (as in the invariant and compromise approaches).

Repeated measures analysis of variance (ANOVA). For this and all subsequent analyses, an alpha level of .05 was used.

Three independent variables were tested: artist's distance, column, and row in a 7 (artist's distance)  $\times$  5 (column)  $\times$  5 (row) repeated measures ANOVA. In brief, centers of pictures often had perceived square tiles, but tiles in leftmost columns stretched, tiles in top rows compressed, and bottom rows lengthened in depth (see Figure 4).

The ANOVA revealed a main effect of artist's distance, F(6, 66) = 63.82,  $\eta_p^2 = .85$ . Perceived relative depth increased as the artist's distance decreased. Bonferroni a posteriori comparisons revealed significant differences between all artist's distances (all p < .03). Figure 4 illustrates this effect, as the number of tiles that appear square changes dramatically from Figure 4a, in which all tiles are elongated, to Figure 4g, in which all are compressed, covering both extremes.

The main effect of column, F(4, 44) = 27.10,  $\eta_p^2 = .71$ , was the result of tiles to the side being judged longer than central ones. Bonferroni a posteriori comparisons revealed significant differences between column 9 and all other columns (all p < .09),

*Figure 4.* Experiment 1 Vantage Point × Column × Row interaction. For the sake of simplicity, mean perceived relative depths have been divided into three groups: (a) compressed (mean perceived relative depth < 0.9), (b) square (mean perceived relative depth = 0.9-1.1), and (c) elongated (mean perceived relative depth > 1.1). The artist's distances are (a) 0.09, (b) 0.18, (c) 0.27, (d) 0.36, (e) 0.45, (f) 0.54, and (g) 0.63 m.

column 7 and columns 3–1 center (all p < .04), and between column 5 and column 1 center (p = .01).

The main effect of row, F(4, 44) = 78.92,  $\eta_p^2 = .88$ , indicates near tiles in the scene appeared longer than far tiles. Bonferroni a posteriori comparisons revealed significant differences between all rows (all p < .05).

The ANOVA revealed significant Artist's Distance × Column, F(24, 264) = 3.25,  $\eta_p^2 = .23$ , and Artist's Distance × Row,  $F(24, 264) = 37.98 \eta_p^2 = .78$ , interactions, meaning the tiles to the far side are markedly different than ones in the central column and nearer rows at the smaller artist's distances. The Row × Column interaction did not reach significance, F(16, 1768) = 1.38, p = .16,  $\eta_p^2 = .11$ . However, the three-way Artist's Distance × Row × Column interaction did, F(96,1056) = 1.73,  $\eta_p^2 = .14$  (see Figure 4). This indicates that tiles in the extreme side columns and bottom rows are especially enlarged at small artist's distances.

*Perceived relative depth function.* We can begin to understand the complex effects of row, column, and artist's distance by first devising a pseudoperspective function for the average tile in a picture for each artist's distance. The result is perceived relative depth = k (correct relative depth) × (observer's distance)<sup>d</sup>/(artist's distance)<sup>j</sup>, where k = 1.30, d = 1 (fixed a priori), and j = 0.67. The 95% confidence intervals for k and j were  $1.24 \le k \le 1.35$  and  $0.64 \le j \le 0.71$ . The pseudoperspective function is highly significant, F(1, 5) = 1645.37, MSE = .001, and fits the data almost perfectly, with  $R^2 = .98$ .

## Discussion

Artist's distance affects perceived relative depth less than predicted by perspective geometry. For an observer at 0.36 m viewing pictures that have artist's distances of 0.63–0.09 m, perspective predicts a sevenfold increase in perceived relative depth, from 0.57 to 4.0, respectively. The actual values changed less than fourfold, from 0.61 to 2.3.

In the pseudoperspective functions for the compromise and projective theories, j = 1 (the exponent on artist's distance), and in compensation and invariant theories, j = 0. Significantly different from both in the function derived here, j = 0.67 (95% confidence interval  $0.64 \le j \le 0.71$ ). Furthermore, in the pseudoperspective functions for the compromise, invariant, and projective theories, k = 1, and in compensation theories, 0 < k < 1. Once again, the function derived here is significantly different from both, with k = 1.30 (confidence interval  $1.24 \le k \le 1.35$ ).

The value of 0.67 for the mediator j needs to be interpreted in light of the constant k, which was 1.30. One factor alone cannot predict the depth distortions. Consider that many researchers argue that a perceived "flattening" of depth regularly occurs when viewing pictures (Koenderink & van Doorn, 2003; Miller, 2004; Sedgwick, 2003; Woodworth & Schlosberg, 1954). For example, Koenderink and van Doorn (2003) found flattening to 85% of real depth (a compression of 15%). If there were no mediator j, then this flattening of 85% would predict a constant k of 0.85 not the 1.30 that was found. In fact, a constant k of 1.30 alone implies that a perceived "elongation" of depth occurs when viewing pictures, a sort of "hyper-depth" perception. The factor that is preventing the apparent depth being pushed to 1.30 is the mediator j; its value of 0.67 balances the effect of the constant k. Koenderink's 0.85 is a product of two functions.

It has further been pointed out that observers do not notice change in apparent depth as they move pictures to and fro. In the pseudoperspective function, this is also achieved by both the constant k and the mediator j. Perceived relative depth varies less for smaller values of the mediator j. As j shrinks toward 0, the artist's distance factor approaches 1. This is a key factor in constancy, producing much less elongation of depth than perspective predicts. However, too small an exponent j leads to square tiles being perceived as compressed—too stubby—when the observer is closer to the picture than the artist.

Recall that the pseudoperspective function merely deals with the average perceived relative depth per picture. We need to envisage extra factors involved with individual tiles, as Figure 4 clearly indicates constancy can be shown by one tile and distortion by its neighbor.

To simplify, we define three categories as follows: let compressed tiles have a perceived relative depth less than 0.9, square tiles a perceived relative depth between 0.9 and 1.1 (inclusive), and elongated tiles a perceived depth greater than 1.1. Their locations are far from random. Compressed tiles are in centermost regions. Elongations are in the periphery, and happily, of course, square tiles always occupy the region between the two. Categories appear to spread out from the central vanishing point in reasonably concentric bands or crescents, shown in Figure 4d, beginning with compressed tiles, followed by square, and then elongated tiles.

Two very influential implications follow: First, the values for k and j in the pseudofunction can be easily modified. It is important



that we point this out emphatically. The crucial fact is that one could simply add more tiles to pictures in the apparently compressed bands (near the central vanishing point) to decrease the value of the constant k. If k deals with average lengths, then adding more apparently short tiles will reduce k. To increase k, one could simply add tiles to the periphery, in the apparently elongated band. If j operates on rates of change, then shortening or lengthening all the tiles equally would not affect j, but modifying the apparent rate of compression and elongation across pictures would. It is absolutely clear that, whereas the basic form of the function will not change, the specific values of k or j are not set in stone, as our later experiments show. For any set of pictures, the values are easily shifted for good reasons that we need to explore.

The second implication has to do with how perceptual constancy has failed altogether for some pictures in the study (e.g., Figure 4a), illustrating the power of the pseudoperspective function. Some pictures are considerably beyond the limits of constancy. The challenge now is to understand the factors producing these limits. To this end, we propose an angles and ratios together (ART) theory.

Some combination of optical features signals the relative width and depth of a depicted square tile (Gibson, 1979). The ART theory proposes that the perception is determined by a combination of "visual angle ratio" and "angle from normal" (see Figure 5). The *visual angle ratio* is the ratio of the visual angle of the depth of an object divided by the visual angle of the width of an object. The *angle from normal* is defined as the angle between the line joining the observer to the central vanishing point and the line to a point on the object (see Figure 5). For convenience, the object's point (*N*) is chosen to be on the base of the object closest laterally from the observer. The line joining the observer to the central vanishing point is traditionally referred to as the "normal" to the plane. The normal and the vanishing point are conventionally defined with



*Figure 5.* Consider an Observer (O) standing in front of a ground plane covered with tiles. The *visual angle ratio* of a tile is defined as  $\perp$  DON/ $\perp$  HON (the angle defined by two lines in the figure, one joining Point D to Point O, and the other joining Point O to Point N). The *angle from normal of a tile* is defined as the  $\perp$  VON. Both these concepts are integral to the angles and ratios together (ART) theory of spatial perception.

respect to a flat picture plane, but they can be considered to be a function of parallel lines and visual angles. The direction of the normal to the vanishing point is also the direction of a line from the observer parallel to the receding sides of a set of tiles. This concept will be important when considering the ART theory's relation to direct perception. For now, consider that many theories have dealt with the visual angles of sides of squares, but here, we have added an angle-from-normal factor, in a novel way.

A priori, one can see that visual angle ratio and angle from normal together determine the perceived relative depth. A given visual angle ratio has to produce a compressed tile for a large angle from the normal and a square tile as the angle from normal decreases. Let us see why. A square on the ground directly below the observer is at 90° from the normal and has a visual angle ratio of 1. A square that is directly in front of the observer and very far away is at a very small angle from the normal and has a very small visual angle ratio because, as it recedes, the visual angle of the square's depth approaches 0° faster than the visual angle of its width. But the small visual angle ratio is visually indeterminate because rectangles approaching the horizon also have a visual angle limit of 0°. In practice, vision rejects the indeterminate and sees slim (horizontally elongated) rectangles in keeping with the foreshortened forms.

A square that is to one side of an observer and very far away will have a very large visual angle ratio. This is because the visual angle of its width approaches 0° faster than the visual angle of its depth. The square's visual angle ratio, approaching infinity as its distance from the observer increases, is visually indeterminate because, once again, all rectangles approach infinity in this fashion. Vision once again sees rectangles, but elongated in depth, the z dimension. Overall, then, the visual angle ratio for an object in front of the observer can range from 0 to infinity, with 1 being specific to a square for objects on the ground below the observer.

Given that the visual angle ratio range (zero to infinity) is far larger than the angle-from-the-normal range  $(0-90^\circ)$ , one might expect the visual angle ratio to make a larger contribution to perceived relative depth than angle from normal. Also, in principle, visual angle ratio has to be a major influence because angle from normal is not information about object shape.

If moving the observer to and fro in front of the picture does not change the observer's/artist's distance ratio much, then the visual angle ratios and angles from the normal also do not change much, which will lead to perceptual constancy for a particular tile. Notice that Figures 4d, 4e, and 4f reveal large regions in which tiles remain square, especially 4e and 4f (artist's distances of 0.45–0.54 m). In this fashion, most movies viewed in theaters are viewed from too close. The artist's distance is at the projector; only here would the observer be at the correct position. Audiences in a movie theater fall in this area of moderate constancy. Little wonder our experience with movies is often acceptable, despite being forward of the projector.

A single picture can have tiles both within the boundaries for square tiles (perceptual constancy) and outside (distortions). Furthermore, distortions occur in the center as well as in the periphery of pictures, for some tiles near the center seem compressed (too small a visual angle ratio). The ART theory, unlike others, can accommodate distortions throughout the picture.

Although the size of the contributions of the factors of the ART theory to perceived relative depth are purely empirical, the choice of the factors is not. These factors fit the argument that all objects that are perceived as equal in relative depth (i.e., square) project visual signals that the object's sides are equal (Gibson, 1966). The most basic element of the information available to the visual system is the visual angle. Angle from normal, importantly, changes as an object moves on the ground plane. It is direction information. Direction and information about a horizontal plane specify the 3-D location of the object. Once the direction and location on a plane, such as the ground plane, is known then, theoretically, the visual angle ratio indicates the perceived relative depth.

We can conclude from first principles that visual angle ratio and angle from the normal belong in the ART theory. To evaluate their empirical contributions in practice, we ran a linear regression analysis, relating visual angle ratio and angle from normal to perceived relative depth of each tile in Experiment 1. That is, whereas the pseudoperspective function was based on mean sizes per picture, the regression analysis was based on every tile. The predictors were entered into the linear regression analysis using stepwise criteria, with both predictors passing criteria.

Because of its larger range, and greater expected contribution to perceived relative depth, visual angle ratio was the first variable entered into the model and, as expected, explained a significant amount of the variance, F(1, 173) = 1032.6, MSE = .069, with  $R^2 = .86$ . Angle from normal was the second variable entered into the model. It produced a significant increase in the amount of variance explained, F(1, 172) = 110.4, MSE = .043, and increased the  $R^2$  of the model to .91. The overall model, then, was highly significant, F(2, 174) = 866.8, MSE = .043, with an  $R^2 = .91$ . The regression function is perceived relative depth = a + b<sub>1</sub>(visual angle ratio) + b<sub>2</sub>(angle from normal), where a = 0.64, b<sub>1</sub> = 1.22, and b<sub>2</sub> = -0.012.

If the ART theory reflects vision's approximation to perspective, then it can predict mean depth perception of a new sample of pictures. Its predictions should fit the function: actual perceived relative depth = s(ART theory prediction), where s = 1. Notice that "s" is the slope of the function. If s = 1, then the ART theory can be said to successfully predict perceived relative depth. However, if s > 1, then the ART theory is underestimating perceived relative depth, whereas an s < 1 would indicate that the ART theory is overestimating perceived relative depth. This will be called the "slope" test.

Second, it is possible to compare the accuracy of the ART theory's predictions with those of the compensation, projective, invariant, and compromise approaches. Their pseudoperspective functions can be used to make precise predictions for each and every tile tested. The prediction can be compared with the mean and standard deviation of the judgments of that tile by the subjects in a given experiment. The ART theory's success rate (the percentage of successful predictions) can be compared with those of the other approaches. This second test is called the "individual tiles" test.

Consider Experiment 1. The relation between the ART theory predicted values and the actual perceived relative depth is actual perceived relative depth = s(ART theory prediction), where s = 0.95 (*SD* = .32). A two-sided *t* test revealed that the ART theory's predictions were successful, as s did not differ significantly from a slope of 1, t(173) = 1.97, p = .057, *MSE* = .024.

Second, was the ART theory more successful at predicting the perceived relative depths of the tiles, obtained from the 12 subjects in Experiment 1, than the other approaches? As with the slope test, predictions for the ART theory were calculated using its ballparkperspective function. Predictions for the other four approachescompensation, projective, invariant, and compromise-were calculated using their pseudoperspective functions. Because the pseudoperspective functions of the compensation and invariant approaches are identical, their predictions are considered together. These predictions were then tested to see whether they differed significantly from the actual perceived relative depths. Bonferroni adjusted t tests were performed to test the difference between the predictions and the actual perceived relative depths for each individual tile. A significant difference was counted as a failure, and the percentage of successful predictions were calculated for the ART theory and the compensation, projective, invariant, and compromise approaches. Note that for the compromise approach, a value of k was chosen so that the average predicted perceived relative depth equaled the average obtained perceived relative depth. This post hoc manipulation of the value of k maximized the fit of the pseudoperspective function for the compromise approach and, as such, greatly favored the success rate of the compromise approach.

A one-way repeated measures ANOVA, with theory as the independent variable (ART theory, compensation/invariant, projective, and compromise), was performed with "successful prediction" as the dependent variable. The variable successful prediction takes on a value of 1 when there is no significant difference between the prediction and the obtained perceived relative depth for an individual tile (as revealed by the *t* test comparing mean and standard deviation of the judgments of the 12 subjects to the predicted value) and a value of 0 when there is a difference. The average successful predictions for each theory is equal to its percentage of successful predictions.

The ANOVA revealed that the theories differed in their rates of successful predictions, F(3, 519) = 12.01,  $\eta_p^2 = .065$ . More important, Bonferroni a posteriori comparisons revealed that the ART theory had more successful predictions (96.6%) than any of the other approaches: compensation/invariant (73.6%), projective (79.9%), or compromise (79.9%) (all p < .001).

The successes of the ART theory here are not a fair measure because the ballpark-perspective function was derived from and tested on the same results. What is needed is a test in new conditions, for example, increasing the observer's distance from 0.36 to 0.54 m.

#### Experiment 2

An increase in observer's distance to 0.54 m puts the observer far from the shortest artist's distance (0.09 m). Will perceptual effects fit with ART theory?

#### Method

Subjects. Twelve first-year students (7 women, mean age = 19.6, SD = 1.9) participated.

Stimuli. The apparatus was the same as in Experiment 1.

*Procedure.* Observers viewed the pictures from a larger distance than before (0.54 m). Perspective predicts that the tiles with an artist's distance of 0.09 m should now appear fully 6.0 times longer than wide, rather than

4.0 times, as in Experiment 1. Hence, Experiment 2 may be a more sensitive test.

## Results

*Dependent measure*. The dependent measure was as before, that is, perceived relative depth.

*Repeated measures ANOVA.* Three independent variables were tested: artist's distance, column, and row in a 7 (artist's distance)  $\times$  5 (column)  $\times$  5 (row) repeated measures ANOVA. Once again, central tiles were generally compressed and peripheral ones elongated (see Figure 6).

As artist's distance grew, tile judgments shrank, F(6, 66) = 42.48,  $\eta_p^2 = .79$ . Bonferroni a posteriori comparisons revealed significant differences between all artist's distances (all p < .007).

Tiles in peripheral columns were judged especially large, F(4, 44) = 54.50,  $\eta_p^2 = .83$ . Bonferroni a posteriori comparisons revealed significant differences between all pairs of columns (all p < .016), except for columns 3 and 5 (p = .58).

Tiles in lower rows were judged particularly large, F(4, 44) = 49.26,  $\eta_p^2 = .82$ . Bonferroni a posteriori comparisons revealed significant differences between all rows (all p < .004).

The ANOVA revealed significant Artist's Distance × Column, F(24, 264) = 7.24,  $\eta_p^2 = .40$ , and Artist's Distance × Row, F(24, 264) = 38.75,  $\eta_p^2 = .78$ , interactions. There was also evidence of a Row × Column interaction, F(16, 1768) = 4.42,  $\eta_p^2 = .29$ . This interaction was nonsignificant in Experiment 1. Evidently, the more extreme conditions in Experiment 2 allowed this interaction to become significant. This might be expected from the significant



*Figure 6.* Experiment 2 Vantage Point × Column × Row interaction. For the sake of simplicity, mean perceived relative depths have been divided into three groups: (a) compressed (mean perceived relative depth < 0.9), (b) square (mean perceived relative depth = 0.9-1.1), and (c) elongated (mean perceived relative depth > 1.1).

three-way Artist's Distance × Row × Column interaction in both Experiments: here, F(96, 1056) = 2.34,  $\eta_p^2 = .18$ .

*Slope test.* The relation between the ART theory predicted values and the actual perceived relative depth is actual perceived relative depth = s(ART theory prediction), where s = 0.98 (*SD* = .25). A two-sided *t* test revealed that the ART theory's predictions were successful, as s did not differ significantly from 1, t(173) = 1.11, p = .27, *MSE* = .019.

Individual tiles test. A one-way repeated measures ANOVA, with theory (ART theory, compensation/invariant, projective, and compromise) as the independent variable and successful prediction as the dependent variable, revealed that the theories differed in their rates of successful predictions, F(3, 522) = 53.15,  $\eta_p^2 = .23$ . More important, Bonferroni a posteriori comparisons revealed that the ART theory had higher predictive success (97.1%) than any of the other approaches: compensation/invariant (84.6%), projective (45.1%), or compromise (78.3%) (all p < .001).

#### Discussion

The ART theory applies at the new observer distance. The effects of the change were much less than perspective predicts. For example, when the artist's distance changed from 0.54 to 0.63 m, 40% of tiles (10 out of 25) changed less than 10%. That is, some perceptual constancy occurred, in keeping with common experience that many pictures look the same when viewed from different distances. However, in revealing cases, there was far less constancy. For instance, when the artist's distance changed from 0.09 to 0.18 m, only a mere 4% of tiles (1 out of 25) changed less than 10%.

More important, the ART theory was able to predict both the constancy and the distortions. Constancy occurred mostly when the relative change in artist's distance was small (e.g., increasing from 0.54 to 0.63 m) and may be the result of minor changes in visual angle ratios and angles from the normal. Distortions occurred predominately when the relative change in artist's distance was large (e.g., from 0.09 to 0.18 m), implying that many distortions occur because of large changes in the visual angle ratios and angles from the normal.

The observer's distance from the picture plane is one of the three variables that fully determine a perspective picture. The remaining two are (a) the observer's position above the ground plane and (b) the orientation in the plane of the objects within the scene. If the ART theory is general, then it applies to these variables. Experiment 3 was designed to test the observer's position above the ground plane.

## Experiment 3

Three perspective pictures of tiles on a ground plane are illustrated in Figure 7. Each has a different artist's vantage point or "eye" height. They can be called "standard view," "child's view," and "worm's-eye view." What does perspective geometry propose should happen as eye height diminishes? No change should occur for tile length, though the vantage point of the observer should appear to lower.

Of great importance to the ART theory is that the visual angle ratios and angles from the normal of all the tiles change with eye height. Consider the entire range of eye heights, from 0 (i.e., at the



*Figure 7.* Three perspective pictures of the same tiles from three different eye heights (going from highest to lowest): standard view (A), child's view (B), and worm's-eye view (C).

ground) to infinitely high. From infinitely high, every square projects an equal visual angle for depth and width and has a visual angle ratio of 1, the ratio specific to a square on the ground. From eye heights approaching ground level, the visual angle for depth decreases to 0, and the visual angle ratio approaches 0. The same ratio is projected by any rectangle, and hence shape is visually indeterminate.

What about angle from the normal? The set of angles from the normal is compressed in Figure 7's pictures as eye height lowers.

In summary, the ART theory is tested at three different eye heights in Experiment 3.

## Method

Subjects. Twelve first-year students (8 women, mean age = 18.5, SD = 1.6) participated.

Stimuli. The apparatus was the same as in Experiments 1 and 2.

The perspective pictures for Experiment 3 are based on the perspective pictures in Experiment 1. Only three of the seven artist's distances were used, namely, 0.18, 0.36, and 0.54 m. These three artist's distances were factorially combined with three different eye heights. The eye heights for each picture can be expressed as a percentage of the eye height used in Experiment 1. The percentages for the standard view, the child's view, and the worm's-eye view are 100%, 71%, and 42%, respectively. The observer's distance was 0.36 m (as in Experiment 1).

Note that the standard view is, in essence, a "reduced" replication of Experiment 1. The tiles that were tested are the same as those in Experiments 1 and 2, namely, those tiles located in the factorial combinations of

rows 1, 3, 5, 7, and 9 and columns 1, 3, 5, 7, and 9. All other aspects of the stimuli were exactly as those in Experiments 1 and 2.

The 25 different tiles tested factorially combined with the three artist's distances and three eye heights produced the 225 pictures used in Experiment 3.

*Procedure.* The procedure was the same as in Experiment 1, with the subjects positioned at a distance of 0.36 m from the picture surface.

#### Results

*Dependent measure.* The dependent measure was the same as in Experiments 1 and 2, that is, perceived relative depth.

*Repeated measures ANOVA.* Four independent variables were tested—eye height, artist's distance, column, and row—in a 3 (eye height)  $\times$  3 (artist's distance)  $\times$  5 (column)  $\times$  5 (row) repeated measures ANOVA (see Figures 8 and 9).

The ANOVA revealed that tile sizes decreased as eye height decreased, F(2, 18) = 168.20,  $\eta_p^2 = .95$ . Bonferroni a posteriori comparisons revealed significant differences between all eye heights (see Figure 8).

The ANOVA revealed that tile size increased as artist's distance decreased, F(2, 18) = 152.77,  $\eta_p^2 = .94$ . Bonferroni a posteriori comparisons revealed significant differences between all artist's distances (see Figure 9).

Tile size increased toward peripheral columns, F(4, 36) = 165.05,  $\eta_p^2 = .95$ . Bonferroni a posteriori comparisons revealed significant differences between all columns. Furthermore, tile size increased toward bottom rows, F(4, 36) = 121.36,  $\eta_p^2 = .93$ . Bonferroni a posteriori comparisons revealed significant differences between all rows.

All two-way interactions were significant (all F > 4.53,  $\eta_p^2$ .26). The three-way Eye Height × Artist's Distance × Column interaction approached significance, F(16, 144) = 1.62, p =.07,  $\eta_p^2 = .15$ . All other three-way interactions were significant (all F > 2.23,  $\eta_p^2 > .20$ ), as well as the four-way Eye Height × Artist's Distance × Column × Row interaction, F(64, 576) =1.50,  $\eta_p^2 = .14$  (see Figure 10). Tile size increased toward bottom peripheral tiles as artist's distance decreased, especially for lower eye heights.

*Slope test.* The relation between the ART theory's predicted values and the actual perceived relative depth determined for each eye height is (a) standard view: actual perceived relative depth = s(ART theory prediction), where s = 0.94 (SD = .32); (b) child's view: actual perceived relative depth = s(ART theory prediction), where s = 0.95 (SD = .30); (c) worm's-eye view: actual perceived relative depth = s(ART theory prediction), where s = 0.92 (SD = .39).

A two-sided *t* test with a Bonferroni adjustment revealed that the ART theory's predictions were successful, as s did not differ significantly from 1 for any of the eye heights, all t(73) < 1.89, p > .063, MSE < .45.

*Individual tiles test.* Because the ART theory passed the slope test for each eye height, the individual tiles in each eye height were pooled for the individual tiles test. A one-way repeated measures ANOVA, with theory (ART theory, compensation/invariant, projective, and compromise) as the independent variable, found differences in the rates of successful predictions, F(3, 672) = 11.24,  $\eta_p^2 = .05$ . More important, Bonferroni a posteriori comparisons revealed that the ART theory had more successful predictions (86.2%) than any of the other approaches: compensation/invariant



*Figure 8.* Experiment 3 main effect of artist's distance (with standard error bars). For all eye heights, as artist's distance increases, mean perceived relative depth per picture decreases.

(68.0%), projective (65.3%), or compromise (69.3%) (all p < .001).

# Discussion

It is evident that ART theory applies across eye heights. Of most interest, in Experiment 3, the ART theory succeeded though there

was very little perceptual constancy across eye heights. Specifically, the perceived relative depth of many tiles decreased noticeably as eye height decreased—fully 81% of tiles (61 out of 75) decreased by 10% or more as eye height decreased from the standard to the worm's-eye views. It appears that the ART theory can handle situations in which there is a lot of apparent constancy



*Figure 9.* Experiment 3 Eye Height ×Vantage Point × Column × Row interaction. For the sake of simplicity, mean perceived relative depths have been divided into three groups: (a) compressed (mean perceived relative depth < 0.9), (b) square (mean perceived relative depth = 0.9-1.1), and (c) elongated (mean perceived relative depth > 1.1).

View



*Figure 10.* A perspective picture of a series of square tiles rotated at  $45^{\circ}$  on a ground plane.

(see Experiment 2) as well as situations in which constancy fails (see Experiment 3).

The remaining degree of freedom for objects on a ground plane is rotation, which is tested in Experiment 4.

## Experiment 4

Changing the orientation of a group of tiles from squares to diamonds results in their diagonals receding directly from the observer (see Figure 10). The vanishing point for the diagonals is implicit because they are not represented by actual lines. Use of diagonals increases the depth of each of the tiles and the total depth of the set of tiles. The relative depth, that is, the depth to width ratio, remains unchanged. The effect is that the mean visual angle ratios of the pictures are increased, from 0.79 (see Experiment 1) to 0.84 (see Experiment 4). Also, from picture to picture, the rate of change in visual angle ratio for Experiment 4 (decrease of 14%) is smaller than in Experiment 1 (decrease of 17%).

Furthermore, changing the orientation of tiles also changes the angles from the normal. In the same way that depth was increased, width is also increased. Coupled with the changes in depth, this produces an entirely new set of angles from the normal. In summary, changing the orientation of the tiles is yet another way to manipulate the visual angle ratios and the angles from the normal.

#### Method

Subjects. Twelve third-year students (7 women, mean age = 22.8, SD = 3.2) participated.

Stimuli. The apparatus used in Experiment 4 is as in Experiments 1–3, but with tiles rotated at 45° (see Figure 10). The depth of a diamond tile in Experiment 4 (a diagonal) is greater than the depth of a square tile (an edge) in Experiment 1 by a factor of  $\sqrt{2}$ . The same applies to width. Because of this increase in width, only 13 columns were depicted (one center column and six on either side). The tiles tested in Experiment 4 consisted of those tiles located in the factorial combinations of rows 1, 3, 5, and 7 and columns 1, 2, 3, 4, 5, and 6. These tiles were indicated to the subjects by using bold lines (three times the thickness of the other lines in the picture) to depict the depth and width of the tiles. The width was depicted at the corner of the tile closest to the observer, whereas the depth was depicted at the left corner of the tile.

The 24 different tiles tested were factorially combined with the seven artist's distances to produce 168 pictures that were used in the experiment.

## Results

Repeated measures ANOVA. Three independent variables were tested—artist's distance, column, and row—in a 7 (artist's distance)  $\times$  6 (column)  $\times$  4 (row) repeated measures ANOVA (see Figure 11).

Tile size increased with decreases in artist's Distance, F(6, 66) = 47.05,  $\eta_p^2 = .81$ . Bonferroni a posteriori comparisons revealed significant differences between all artist's distances (all p < .012).

Tile size increased toward peripheral columns, F(5, 55) = 62.10,  $\eta_p^2 = .85$ . Bonferroni a posteriori comparisons revealed significant differences between all pairs of columns (all p < .023) except for columns 1 and 2 (p = .99), columns 1 and 3 (p = .68), and columns 5 and 6 (p = .073).

Tile size increased toward bottom rows, F(3, 33) = 57.92,  $\eta_p^2 = .84$ . Bonferroni a posteriori comparisons revealed significant differences between all rows (all p < .01).

The ANOVA revealed significant Artist's Distance × Column, F(30, 330) = 2.12,  $\eta_p^2 = .16$ , and Artist's Distance × Row, F(18, 198) = 40.67,  $\eta_p^2 = .79$ , interactions. The Row × Column interaction did not reach significance, F(15, 165) = 1.43, p = .14,  $\eta_p^2 = .12$ . Finally, the three-way Artist's Distance × Row × Column interaction was significant, F(90, 990) = 1.69,  $\eta_p^2 = .13$ . Tiles in the periphery and bottom rows increased in perceived relative depth the most as artist's distance decreased.

*Slope test.* The relation between the ART theory predicted values and the actual perceived relative depth is actual perceived relative depth = s(ART theory prediction), where s = 0.94 (*SD* = .22). A two-sided *t* test revealed that the ART theory's predictions



*Figure 11.* Experiment 4 Vantage Point × Column × Row interaction. For the sake of simplicity, tiles have been divided into four groups: (a) compressed (mean perceived relative depth < 0.9), (b) square (mean perceived relative depth = 0.9-1.1), (c) elongated (mean perceived relative depth > 1.1), and (d) untested tiles.

deviated slightly but significantly, and the slope was not equal to 1, t(167) = 3.35, p = .001, MSE = .017.

*Individual tiles test.* A one-way repeated measures ANOVA, with theory (ART theory, compensation/invariant, projective, and compromise) as the independent variable and successful predictions as the dependent variable, was performed.

The ANOVA revealed that the theories differed in their rates of successful predictions, F(3, 501) = 9.40,  $\eta_p^2 = .053$ . More important, Bonferroni a posteriori comparisons revealed that the ART theory had more successful predictions (86.3%) than any of the other approaches: compensation/invariant (69.6%), projective (60.7%), or compromise (67.3%) (all p < .002).

## Discussion

Again, the ART theory makes the best predictions. Of most interest is that it was imperfect on the slope test and overestimated the actual perceived relative depth by 6%. Although this is an extremely small overestimation, it does pose some interesting possibilities. The overestimation may have been because of the diagonal tiles being perceived as resting on a tilted ground plane and foreshortened less than they would be if horizontal. Alternatively, the diamonds' vanishing point from which the angles from the normal are measured is not explicitly represented. If this leads to underestimations of the angles from the normal, then it would produce the overestimations. The last possibility to be considered is that it is simply the result of a response bias. Observers may have been reluctant to report large perceived relative depths. The preponderance of apparently compressed tiles may have caused observers to bias their judgments toward lower perceived relative depths. This possibility is bolstered by the fact that, even though Experiments 1-3 all passed the slope test, the slopes were all in the direction of overestimated predictions. If so, then the 6% overestimation here is an interesting procedural artifact rather than a genuine perceptual result.

Comparing common tiles in Experiments 1 and 4 reveals very little constancy; only 21% of tiles (18 of 84) changed less than 10%. So, the ART largely accounted for perceived relative depth again, even though constancy failed.

## General Discussion

The ART theory predicted tile perception across distance, eye height, and tile rotation better than alternatives tested with highly favorable assumptions. Though devised using squares, ART theory may apply widely. In Figure 12, the relative depth of Object 1 is simply its length divided by its width. It has both a visual angle ratio and an angle from the normal. Therefore, ART theory can be applied to solid objects. It also applies to perception of spaces. In Figure 12, the space between Objects 1 and 2 has both a visual angle ratio and an angle from the normal (from the central vanishing point to the intersection of Arrows C and D).

Thus far, ART theory has been applied to the relative depths of objects. However, some of the tiles in the periphery of pictures may seem not only elongated but also not to have 90° corners, that is, not to be rectangular. The perception of the angles at corners is another important aspect of shape perception. Indeed, some theo-

*Figure 12.* Object 1 and Object 2 are standing on the ground plane, the central vanishing point being clearly illustrated. Object 1 has a width indicated by Arrow A and a depth indicated by Arrow B. The visual angle ratios and angles from the normal of both Arrows A and B can be determined. From this information, the angles and ratios together (ART) theory can predict a perceived relative depth for Object 1. The same logic applies to the relative distance between Objects 1 and 2, where lateral distance is indicated by Arrow C, whereas distance in depth is indicated by Arrow D.

ries, for example, Perkins' laws, indicate when corners of cubes appear correct versus distorted, that is "90°" versus "not 90°" (Cutting, 1987; Kubovy, 1986; Perkins & Cooper, 1980). Usefully, however, the ART theory can also be applied to the perception of angles.

Assume that the horizontal parallel lines on the screens in Experiments 1–3 (the lines running left and right) are perceived as showing parallel edges on the ground, an assumption justified by geometric constraints on "assuming good form" (Perkins & Cooper, 1980). Call this the assumption of "two parallel edges on the ground." Given this "two parallel edges" assumption, the ART theory predicts changes in perceived angle. For example, for tiles at or very near the center of the picture (e.g., tiles in the central column and the adjoining ones), the edges shown by converging lines in the picture (that is, the perceived left and right sides of the tile) are equal or nearly equal. Together with the "two parallel edges" assumption, this requires perceived angles of 90° or very close.

For tiles near the periphery, the perceived lengths of the left and right sides of the tile are not equal. However, given the "two parallel edges on the ground" assumption, the ART theory predicts the perceived angles of the four corners of the tile. For example, consider a case in which the tile in the central column is perceived as square. Now consider a tile near the periphery. If the length of the right side is 1.1 units and the left side is 1.2 units, and the base is 1.0 (the closer of the two parallel edges on the ground), then trigonometry predicts the perceived corner angles to be  $112^{\circ}$  (bottom right),  $52^{\circ}$  (bottom left),  $68^{\circ}$  (top right), and  $128^{\circ}$  (top left). Of course, it is important to check the



ART theory's predictions empirically. Vision may adopt somewhat independent approximations for length and angle in perspective pictures. Our point here is simply that the ART theory is consistent with changes in angle perception as well as length. Indeed, the ART theory might be integrated smoothly with Perkins' laws of angles at cubic corners because it indicates when tile edges are at or far from 90°. Perkins' laws are "all-or-nothing" however, whereas the ART theory predicts gradual changes in perceived angle.

Both one- (see Experiments 1, 2, and 3) and two-point perspective pictures (see Experiment 4) were tested here. Three-point perspective results if the tiles are on a cube tilted with respect to the picture plane (see Figure 13). The top of the cube is the equivalent of a square tile on a horizontal plane, and the sides are the equivalent of square tiles on vertical planes. The orientation of the planes is not a factor in the ART theory; it can apply at all orientations and to each face of a cube independently. For sure, in Figure 13, cubes look distorted. So constancy and distortion need to be reconciled for three-point pictures, and ART theory's factors may be key. For example, differential elongation of sides can produce angular distortions at corners.

The ART factors are present in the 3-D world. Visual angle ratio is simply the visual angle of an object's depth divided by the visual angle of the object's width. The central vanishing point is a direction to which parallel edges recede. Hence, *angle from the normal* can be defined as the angle between the line beginning at the observer and parallel to the ground and an object's parallel receding edges, and the line to a point on the object (see Figure 5). Indeed, some of the effects that the ART theory can account for in picture perception occur in the 3-D world. Perceived compression at great distances is an often-reported phenomenon (Baird & Biersdorf, 1967; Foley, 1972; Gilinsky, 1951; Harway, 1963; Wagner, 1985). Perceived elongation, another effect in ART theory, although not as widely reported, has also been found (Baird & Biersdorf, 1967; Harway, 1963; Heine, 1900, as cited in Norman et al., 1996; Wagner, 1985).

In summary, ART theory is an approximation theory, proposing that optical features (visual angle ratio and angle from normal) determine the perception of relative depth; it predicts when constancy fails and by how much and explains the factors responsible for the perspective effects that puzzled Renaissance artists.



*Figure 13.* A three-point perspective picture results if the tiles were placed on the top of gray cubes, tilted with respect to the picture plane.

## References

- Baird, J. C., & Biersdorf, W. R. (1967). Quantitative functions for size and distance judgments. *Perception & Psychophysics*, 2, 161–166.
- Cutting, J. E. (1987). Rigidity in cinema seen from the front row, side aisle. Journal of Experimental Psychology: Human Perception and Performance, 13, 323–334.
- Foley, J. M. (1972). The size-distance relation and intrinsic geometry of visual space: Implications for processing. *Vision Research*, 12, 323–332.
- Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston: Houghton Mifflin.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Gibson, J. J. (1982). Pictures as substitutes for visual realities. In E. Reed & R. Jones (Eds.), *Selected essays of James J. Gibson* (pp. 231–240). Hillsdale, NJ: Erlbaum. (Original work published 1947)
- Gilinsky, A. S. (1951). Perceived size and distance in visual space. *Psy-chological Review*, 58, 460–482.
- Harway, N. I. (1963). Judgment of distance in children and adults. *Journal of Experimental Psychology*, 64, 385–390.
- Kennedy, J. M., & Juricevic, I. (2002). Foreshortening gives way to forelengthening. *Perception*, 31, 893–894.
- Koenderink, J. J., & van Doorn, A. J. (2003). Pictorial space. In H. Hecht, R. Schwartz, & M. Atherton (Eds.), *Looking into pictures: An interdisciplinary approach to pictorial space* (pp. 239–299). Cambridge, MA: MIT Press.
- Koenderink, J. J., van Doorn, A. J., Kappers, A. M. L., & Todd, J. T. (2001). Physical and mental viewpoints in pictorial relief [Electronic version]. *Journal of Vision*, 1, 39a.
- Kubovy, M. (1986). The psychology of perspective and Renaissance art. Cambridge, MA: Cambridge University Press.
- La Gournerie, J. D. (1859). Trait de perspective lin'eaire contentant les trac'es pour les tableaux plans et courbes, les bas-reliefs et les d'ecorations th'eatrales, avec une th'eorie des effets de perspective [Treatise on linear perspective containing drawings for paintings, architectural plans and graphs, bas-reliefs and theatrical set design; with a theory of the effects of perspective]. Paris: Dalmont et Dunod.
- Miller, R. J. (2004). An empirical demonstration of the interactive influence of distance and flatness information on size perception in pictures. *Empirical Studies of the Arts*, 22, 1–21.
- Niall, K. K. (1992). Projective invariance and the kinetic depth effect. Acta Psychologica, 81, 127–168.
- Niall, K. K., & Macnamara, J. (1989). Projective invariance and visual shape constancy. *Acta Psychologica*, 72, 65–79.
- Niall, K. K., & Macnamara, J. (1990). Projective invariance and picture perception. *Perception*, 19, 637–660.
- Niederée, R., & Heyer, D. (2003). The dual nature of picture perception: A challenge to current general accounts of visual perception. In H. Hecht, R. Schwartz, & M. Atherton (Eds.), *Looking into pictures: An interdisciplinary approach to pictorial space* (pp. 77–98). Cambridge, MA: MIT Press.
- Norman, J. F., Todd, J. T., Perotti, V. J., & Tittle, J. S. (1996). The visual perception of three-dimensional length. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 173–186.
- Perkins, D. N., & Cooper, R. G., Jr. (1980). How the eye makes up what the light leaves out. In M. A. Hagen (Ed.), *The perception of pictures: Vol. 2, Durer's devices: Beyond the projective model of pictures* (pp. 95–130). New York: Academic Press.
- Pirenne, M. H. (1970). Optics, painting, and photography. Cambridge, MA: Cambridge University Press.
- Rogers, S. (1995). Perceiving pictorial space. In W. Epstein & S. Rogers (Eds.), Handbook of perception and cognition: Perception of space and motion (2nd ed., Vol. 5, pp. 119–163). San Diego, CA: Academic Press.
- Rogers, S. (2003). Truth and meaning in pictorial space. In H. Hecht, R. Schwartz, & M. Atherton (Eds.), *Looking into pictures: An interdisci*

plinary approach to pictorial space (pp. 301–320). Cambridge, MA: MIT Press.

- Rosinski, R. R., & Farber, J. (1980). Compensation for viewing point in the perception of pictured space. In M. A. Hagen (Ed.), *The perception of pictures: Vol. 1, Alberti's window: The projective model of pictorial information* (pp. 137–176). New York: Academic Press.
- Rosinski, R. R., Mulholland, T., Degelman, D., & Farber, J. (1980). Picture perception: An analysis of visual compensation. *Perception & Psychophysics*, 28, 521–526.
- Sedgwick, H. A. (2001). Visual space perception. In E. B. Goldstein (Ed.), *Blackwell handbook of perception* (pp. 128–167). Oxford, England: Blackwell Publishers.
- Sedgwick, H. A. (2003). Relating direct and indirect perception of spatial layout. In H. Hecht, R. Schwartz, & M. Atherton (Eds.), *Looking into pictures: An interdisciplinary approach to pictorial space* (pp. 239– 299). Cambridge, MA: MIT Press.
- Sedgwick, H. A., & Nicholls, A. L. (1993). Cross talk between the picture surface and the pictured scene: Effects on perceived shape. *Perception*, 22(Suppl.), 109.
- Smallman, H. S., Manes, D. I., & Cowen, M. B. (2003, October). Measuring and modeling the misinterpretation of 3-D perspective views.

*Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting* (pp. 1615–1619). Human Factors and Ergonomics Society, Santa Monica, CA.

- Smallman, H. S., St. John, M., & Cowen, M. B. (2002, October). Use and misuse of linear perspective in the perceptual reconstruction of 3-D perspective view displays. *Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting* (pp. 1560–1564). Santa Monica, CA: Human Factors and Ergonomics Society.
- Wagner, M. (1985). The metric of visual space. Perception & Psychophysics, 38, 483–495.
- Woodworth, R. S., & Schlosberg, H. (1954). Experimental psychology (2nd ed.). New York: Holt, Rinehart & Winston.
- 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.

Received May 19, 2005 Revision received March 14, 2006 Accepted March 17, 2006

# **Correction to Perruchet et al. (2006)**

In the article, "Do We Need Algebraic-Like Computations? A Reply to Bonatti, Pena, Nespor, and Mehler (2004)," by Pierre Perruchet, Ronald Peereman, and Michael D. Tyler (*Journal of Experimental Psychology: General*, 2006, Vol. 135, No. 2, pp. 322–326), the page numbers that Dr. Perruchet cited from Dr. Bonatti et al.'s article were printed incorrectly. These page numbers should appear as follows:

p. 322, left column, Paragraph 1, Line 5: Replace "(p. 21)" with "(p. 320)"

p. 322, left column, Paragraph 2, Line 3: Replace "(e.g., pp. 7, 8, 11)" with "(e.g., pp. 317, 318, 320)"

p. 322, left column, Paragraph 2, Line 4: Replace "(e.g., p. 12)" with "(e.g., p. 320)"

p. 324, right column, Footnote 2, Line 11: Replace "(p. 12)" with "(p. 317)"

p. 325, left column, Paragraph 1, Line 8: Replace "(p. )" with "(p. 318)"

p. 326, left column, Paragraph 2, Line 10: Replace "(p. 8)" with "(p. 316)"