

# Depth Perception in Motion Parallax and Stereokinesis

Corrado Caudek and Dennis R. Proffitt

Perceived depth in the stereokinetic effect (SKE) illusion and in the monocular derivation of depth from motion parallax were compared. Motion parallax gradients of velocity can be decomposed into 2 components: object- and observer-relative transformations. SKE displays present only the object-relative component. Observers were asked to estimate the magnitude and near-far order of depth in motion parallax and SKE displays. Monocular derivation of depth magnitude from motion parallax is fully accounted for by the perceptual response to the SKE, and observer-relative transformations absent in the SKE are of perceptual utility only as determinants of the near-far signing of perceived sequential depth. The amount of depth and rigidity perceived in motion parallax and SKE displays covaries with the projective size of the stimuli. The monocular derivation of depth from motion is mediated by a perceptual heuristic of which the SKE is symptomatic.

The perception of depth from monocular motion information occurs in three situations, two yielding veridical depth percepts and the third giving rise to an illusion. First, depth is perceived when one views an object that is rotating around some axis other than the line of sight. Second, depth is seen when a viewer moves past stationary objects, or conversely when objects translate by a stationary observer on some path other than the line of sight. The depth-evoking optical transformations that occur in these situations are called *motion parallax*. Finally, depth is observed when one views certain two-dimensional patterns that are rotating in the picture plane. Called the *stereokinetic effect* (SKE) by Musatti (1924), these phenomena are illusory because here the distal objects are, in fact, two-dimensional.

In this article, we propose that the SKE illusion is symptomatic of the processes by which the visual system derives depth in everyday situations. Proffitt, Rock, H. Hecht, and Schubert (1992) decomposed rigid object rotations into two motion transformations, one of which defines the stimulus basis for the SKE. They showed that the perceptual response to the SKE is indicative of how people perceive depth when they observe small object rotations of 15° or less. Through a similar analysis, we show here that motion parallax can be decomposed into two transformations, one of which again defines the stimulus basis for the SKE. In four experiments, we show that when they perceive depth magnitude in motion parallax displays, people use only that motion component that exists in SKE displays. The remaining motion transformation has perceptual utility only for the specification of depth order or the near-far signing of perceived sequential depth. Finally, we show that because depth mag-

nitude is not geometrically specified by the SKE transformation, people must derive depth magnitude in a heuristical manner, about which we make some speculations.

## Stimulus Bases for the Stereokinetic Effect

The SKE is a visual illusion in which two-dimensional stimuli rotating in the frontoparallel plane appear to be rigid objects moving in three-dimensional space (Musatti, 1924; Zanforlin, 1988). The most frequently studied SKE display, depicted in Figure 1, is a pattern consisting of an eccentric arrangement of nested circles that are attached to a slowly rotating turntable. Three aspects of the perception of this display need mentioning. First, as the display revolves, those motions occurring in the same direction as the contours cannot be detected (Hildreth, 1984; Wallach, Weisz, & Adams, 1956). This results in what Musatti (1924) called *orientational stability*. That is, the contours are seen to move in relation to each other while maintaining a constant orientation in the picture plane. Second, the pattern's appearance is that of a moving three-dimensional cone or funnel, with the latter being a depth-order reversal of the former. Third and most important for our present purposes, the cone or funnel is seen to have a definite tip-to-base depth that varies very little between or within individuals (Musatti, 1928; Proffitt et al., 1992; Zanforlin, 1988).

Proffitt et al. (1992) defined the stimulus basis for the SKE as the presence of only a subset of those transformations exhibited by the projection of a rotating three-dimensional object onto a two-dimensional plane. Proffitt et al. related the motions inherent in SKE displays to those present in parallel projections of rotating objects. To ease exposition, parallel projection is again assumed in this discussion of object rotation. The two transformations of interest are presented in Figures 2 and 3. There we represent two instances in the rotation of a cone's cross-section within a Cartesian coordinate system depicted from above. In these figures, the z-axis coincides with the depicted observer's line of sight, the x- and y-axes define this observer's picture plane, and the y-axis is vertical in relation to this plane. A rigid object rotating around the y-axis projects motions onto

---

Corrado Caudek and Dennis R. Proffitt, Department of Psychology, University of Virginia.

This research was supported by U.S. Air Force Office of Scientific Research Grant AFOSR-91-0057 and National Aeronautics and Space Administration Grant NCA2-468. All computer graphics were programmed by Steve Jacquot.

Correspondence concerning this article should be addressed to Dennis R. Proffitt, Department of Psychology, Gilmer Hall, University of Virginia, Charlottesville, Virginia 22903-2477.

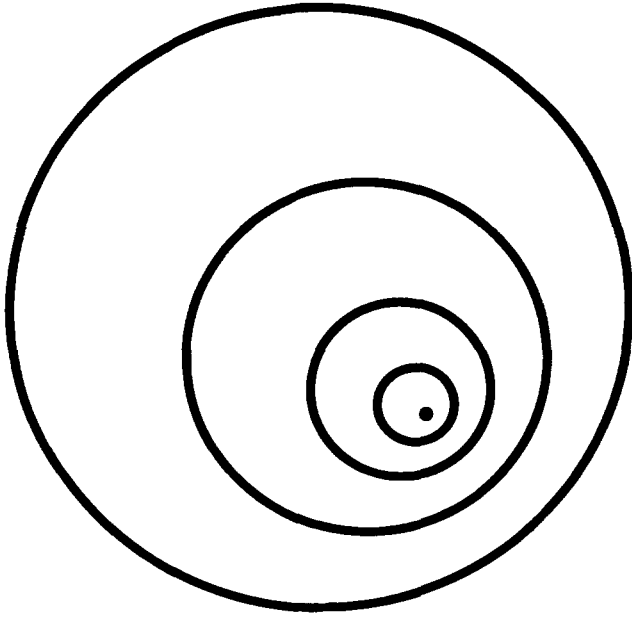


Figure 1. A pattern of eccentric rings that produces the stereokinetic effect when placed on a slowly rotating turntable.

the picture plane,  $P$ , that can be decomposed into two distinct transformations. Of course, other and more sophisticated analyses abound in the literature (Koenderink, 1986; Longuet-Higgins & Prazdny, 1980; Todd, 1982; Ullman, 1979); however, this current decomposition is motivated by the desire to isolate those transformations related to the SKE.

As depicted in Figure 2, the first transformation relates to the motions of points on the object that are initially at different object-relative depths, that is, points differing in their  $z$ -axis values. For the depicted configuration, two points,  $b$  and  $c$ , initially lie in the picture plane, whereas a third point,  $a$ , is maximally displaced within the object toward the observation point,  $O$ . The depth of the triangle  $abc$  is  $d$ . A rotation of the configuration around the  $y$ -axis located at the midpoint of the base produces a projected displacement of Point  $a$  (the position of Point  $a$ ,  $p[a]$ , on the projection plane) that is equal to  $d$  times the sine function of the angle of rotation,  $\theta$ . This projected displacement relates to the depth of the object.

The second transformation manifested by small object rotations is discussed in reference to Figure 3. The distance between Points  $b$  and  $c$ , which were initially in the picture plane, foreshortens as a cosine function of the angle of rotation. This transformation relates to changes in the object's orientation in relation to the observer.

Proffitt et al. (1992) defined the stimulus basis for the SKE as consisting of only the first motion component inherent in the projection of a three-dimensional object undergoing a small rotation. Hereafter, it is referred to as the *SKE transformation*. The foreshortening motions of each contour that would be consistent with the object's orientation change are completely absent. For example, consider the cone depicted in Figure 1. Rotating this pattern in the

picture plane causes the contours to achieve orientational stability, and thus they appear to move in relation to each other in a manner consistent with the projection of a cone or funnel undergoing small rotations. On the other hand, the contours themselves do not foreshorten as is appropriate for their apparent continuous change in observer-relative orientation. Thus, the projected transformations of rigid objects undergoing small rotations can be decomposed into two transformations: One forms the basis for the SKE, and the other is completely absent in this illusory event.

Two important implications of this definition need emphasis. First, SKE transformations do not themselves specify magnitudes of object-relative depth. As is illustrated in Figure 2, the projected displacement of a cone's tip, resulting from a rotation around the midpoint of its base, is a function of the angle of rotation and its object-relative depth. That is,  $e = d \sin(\theta)$ . To derive a depth magnitude from this projected object-relative displacement, the magnitude of the rotation must be known. For rigid object rotations, the magnitude of rotation can be derived from knowing the magnitude of contour foreshortening. As Figure 3 illustrates, foreshortening occurs as a cosine function

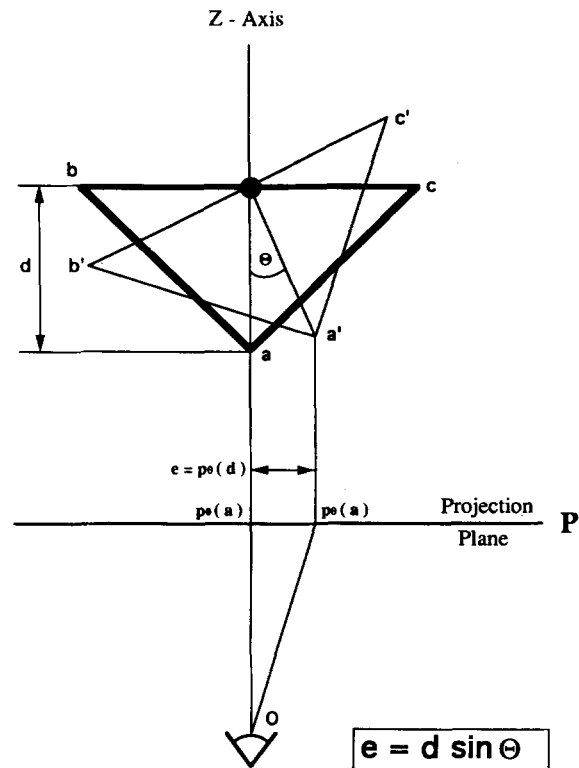


Figure 2. The counterclockwise rotation of a cone in relation to an observer is drawn in a coordinate system as if viewed from above. (In this representation, the  $z$ -axis coincides with the observer's line of sight and the  $y$ -axis defines the projection panel's vertical axis. Configurations  $abc$  and  $a'b'c'$  represent the cross-sections of the cone before and after the rotation through the angle  $\theta$ . The projection of Point  $a$  onto the projection plane  $P$ ,  $p[a]$ , is displaced to the right as a sine function of the angle of rotation  $\theta$  and the depth,  $d$ , of Configuration  $abc$ .)

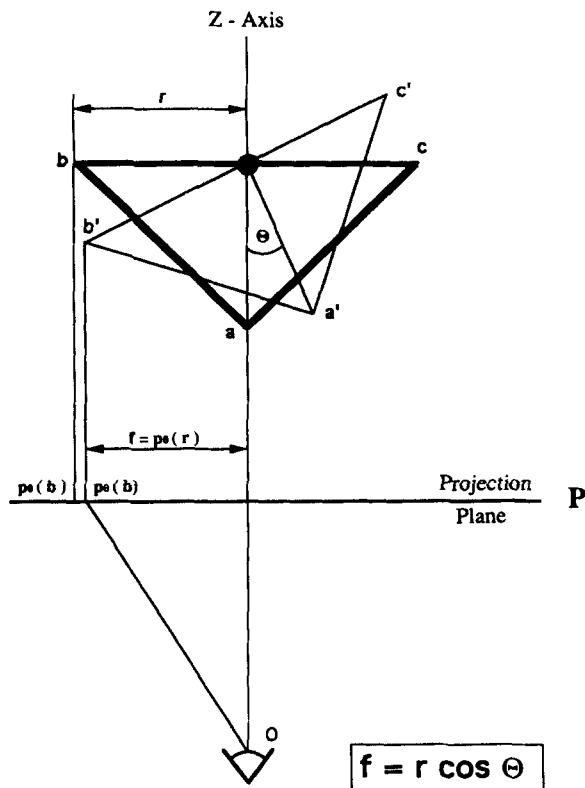


Figure 3. Projective foreshortening of the base of a cone represented by Configuration *abc* consequent to the counterclockwise rotation through the angle  $\theta$ . (The projection of Point *b* onto the projection plane *P*,  $p_0(b)$ , moves to the right as a cosine function of the angle of rotation  $\theta$  and of the distance *r* [radius of the cone]. Correspondingly, the projection of Point *c* will be displaced to the left by an equivalent amount.)

of the angle of rotation and the size of the cone's base,  $2r$ . In the case depicted, the size of the cone's base is initially given by the distance between Points *b* and *c* because they fall in the picture plane. More generally, for small cycling rotations such as those apparent in SKE phenomena, the maximum distance between Points *b* and *c* occurs repeatedly whenever they pass through the picture plane; at these instants, their extent is given in the projection. Thus, both transformations are required to derive depth magnitude from rigid object rotations. In the SKE, however, there are no contour-foreshortening motions, and thus depth magnitude is not geometrically specified (unless it is assumed that the rotation is infinitesimal, thereby requiring an infinite depth percept:  $d = e/\sin[\theta]$ , where  $\theta = 0$ ). As was previously mentioned, people perceive a definite depth magnitude in SKE displays with little between- and within-subject variability. Moreover, the depth seen is of a magnitude that differs little from the diameter of the apparent cone's base. In terms of a depth-to-width ratio, Proffitt et al. (1992) found that the apparent depth in stereokinesis was contained in a range approximately between 0.7 and 1.2. Thus, the perceptual process by which depth magnitude is derived in the SKE cannot be one that is consistent with a canonical

geometrical analysis of the projective transformations inherent in the display.

The second issue to note here is that for small rotations of rigid objects, only the SKE transformations are sufficiently salient to be of perceptual utility. Contour foreshortening is minuscule. Again referring to Figures 2 and 3, it can be seen that displacements of the cone's tip occur as a sine function of rotation times object-relative depth, whereas foreshortening of the base is a cosine function of rotation times the cone's diameter. The derivative of the sine function for small angles is enormous compared with that for the cosine function; in fact, the derivative of the cosine function around  $0^\circ$  is 0. Moreover, a  $15^\circ$  rotation of the cone results in a projected tip displacement that is 26% of its depth, whereas the base is foreshortened at this angle by only 3%. Proffitt et al. (1992) argued that for small rotations of around  $15^\circ$  or less, contour foreshortening is too small to be perceptually useful, and thus the perceptual system must derive depth magnitude solely from the salient motions that define the SKE transformation. For this reason, the SKE is symptomatic of the process by which the perceptual system derives depth when small rotations are observed: In both cases, depth magnitude is derived solely from the SKE motion component.

### Perceptual Heuristics

SKE transformations are not informative about depth magnitudes; however, they do signal the presence of depth, thereby motivating the derivation of depth magnitude from this incomplete information. The consistent empirical finding that SKE displays evoke perceptions of objects that have definite depth magnitudes implies that the perceptual system can derive depth magnitude from their motions even though such an analysis must be inconsistent with the laws of projective geometry. Following Braunstein (1976), we propose that the perception of depth from monocular motion information is mediated by *perceptual heuristics*.

We hypothesize that the derivation of depth from SKE transformations is performed by the visual system in conformity with a *regularity bias*. In particular, we speculate that it uses a *compactness assumption* such that whenever SKE transformations are the only source of salient depth-relative information, the best guess of the system is that the z-axis extent of the visual object equals the smallest of the dimensions exhibited by the projection of the object in the x-y plane. Consistently with depth assessments in the SKE literature (Musatti, 1924; Proffitt et al., 1992; Zanforlin, 1988), we propose that depth magnitude is assigned by the perceptual system as a function of two factors: the magnitude of the SKE transformations and a default regularity bias of the system that is based on the compactness assumption. When SKE transformations are small, apparent depth is some fraction of the width of the SKE displays. When the magnitude of these transformations increases, perceived depth increases, although the depth-to-width ratio is never much greater than unity. Loomis and Eby (1988) found con-

verging evidence for our account by creating a compelling demonstration of a failure of shape constancy in the perception of rotating objects. They suggested that the apparent shape of rotating objects is influenced by the dimensions of their projected outlines.

By this account, stereokinesis is symptomatic of a perceptual heuristic that is evoked whenever SKE transformations are the only salient source of depth-relative information. According to this hypothesis, the strategy used by the perceptual system to derive depth in stereokinesis is not specific to this visual illusion. Rather, it reflects a perceptual heuristic that is used in many everyday situations.

### SKE and Motion Parallax

The purpose of the present investigation is to relate stereokinesis to the perceptual response to motion parallax. The term *motion parallax* is commonly used to refer to the gradient of the angular velocities in the optical flow field produced by the relative motion between the observer and the environment. As is depicted in Figure 4, for a horizontal dihedral angle pointing toward the picture plane, a motion parallax velocity flow field can be decomposed into two components. This decomposition can be exemplified by considering the paradigm introduced by Rogers and Graham (1979), in which motion parallax was examined in two conditions. In the self-produced parallax condition, a motion parallax flow field was created by the movement of the observer's head that systematically distorted a random-dot pattern on a stationary oscilloscope. In the externally produced parallax condition, a motion parallax flow field was created by the movement of an oscilloscope in relation to a stationary observer and by the corresponding motion of the random dots displayed on the oscilloscope. Two distinct components of the velocity flow field can be distinguished in this experimental setting. The first of these two components is the common motion component of the whole ve-

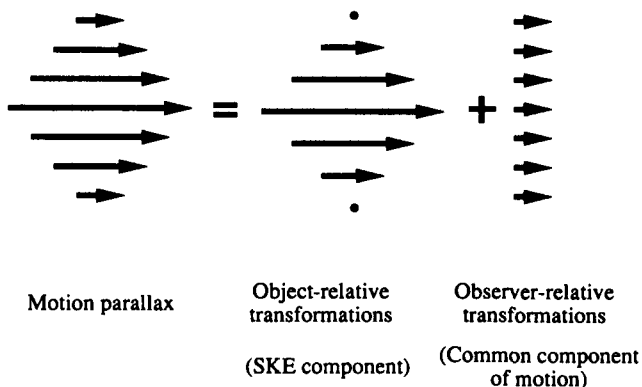


Figure 4. Decomposition of a motion parallax velocity gradient. (Line lengths represent relative velocities in the optic flow field. For illustrative purposes, the magnitude of the object-relative vectors has been greatly exaggerated. In actual displays, the common component of motion component is many times greater than the largest object-relative vector.)

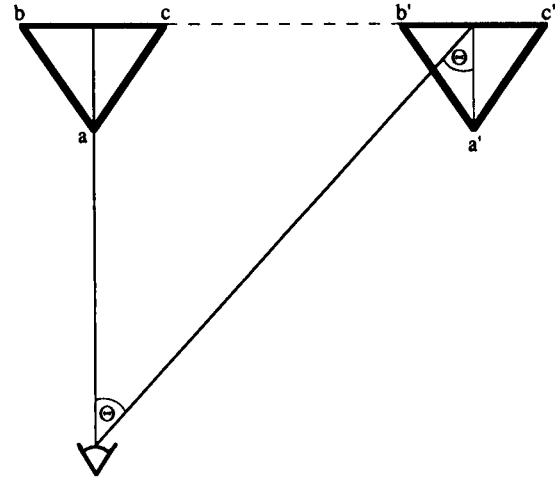


Figure 5. Representation of the cross-section ( $abc$ ) of a cone in two moments of a translation along a path orthogonal to the initial line of sight. (Angle  $\theta$  corresponds to the observer-relative displacement of the object as well as its effective observer-relative rotation.)

locity field (provided either by the motion of the oscilloscope or by the motion of the observer's head). The second of these components is obtained by subtracting this common motion component from the whole velocity field (provided by a gradient of velocity presented on the oscilloscope).

The common component of motion parallax corresponds to changes in the angular displacement of the object in relation to the observer. In motion parallax, the common motion component is an *observer-relative transformation*. The second component in motion parallax consists of the differential angular velocities of the optic flow field that are determined by the depth within the distal object: These motions are called the *object-relative transformations*. In accordance with the definition of the stimulus bases of the SKE proposed by Proffitt et al. (1992), the object-relative transformations of a motion parallax velocity flow field are equivalent to the transformations produced by an SKE display. That is, the magnitude of object-relative motions is a function of each point's  $z$  depth and the sine of the angular displacement of the object in relation to the observer.

There is, however, a very important difference between motion parallax and object rotations regarding the salience of the information that specifies the angle of effective object rotation. In object rotations, the cosine of this angle relates to the projected foreshortening produced by the rotation of the distal object in relation to the observer. In motion parallax, on the other hand, this angle is manifested in a far more salient information source. Figure 5 depicts a cross-section of a cone moving in relation to a stationary observer along a path orthogonal to the initial line of sight. Note that this object undergoes an effective observer-relative rotation that is equivalent to its angular displacement. As with rigid object rotations, the visual angle between Points  $b$  and  $c$  will be foreshortened as a cosine function of this angle; however, in motion parallax, the angle of effective rotation need not

be computed from this minuscule foreshortening because it is given robustly by the extent of the common motion component,  $\theta$ .

A canonical geometrical analysis for recovering depth from motion parallax can be achieved through a variety of perspective representations (Graham, Baker, M. Hecht, & Lloyd, 1948; Koenderink, 1986; H. Ono & Comerford, 1977; M. Ono, Rivest, & Ono, 1986). The following experiments demonstrate, however, that perceived depth through motion is derived more heuristically. We propose that the effective stimulus for perceiving depth in motion parallax consists of a subset of the available information, that being the SKE component of the optical flow field. These object-relative motions are related to the default regularity bias of the system in deriving perceived depth magnitudes. The observer-relative transformations are not used in deriving depth magnitudes, even though in motion parallax this information is highly salient. Geometrically, both the object-relative and the observer-relative components of motion parallax are necessary to derive veridical depth magnitudes (modulo a scaling factor for distance) from velocity flow fields. As was previously discussed when considering rigid object rotations, for motion parallax it is also true that the presence of the SKE motion component alone is geometrical consistent only with an infinite object-relative depth (Braunstein & Andersen, 1981).

### Overview of Experiments

The following experiments on perceiving depth in motion parallax demonstrate that the perception of depth magnitude but not depth order is fully accounted for in terms of the perceptual response to the SKE. In four experiments, SKE displays were matched with ones exhibiting motion parallax. That is, pairs of displays were shown in which object-relative transformations were identical; however, one member of the pair, the motion parallax display, presented an appropriate common motion component, whereas the other, the SKE display, presented none. We found that the presence of the common motion component had no effect on perceived depth magnitudes. Its influence was on only the specification of depth order or the near-far signing of object features.

By relating the current studies to those of Proffitt et al. (1992), we note that their definition for the stimulus basis for the SKE eliminates the limitations of the turntable methodology in producing SKE displays. Traditional SKE displays produced by the turntable methodology exhibit equal  $x$ -axis and  $y$ -axis sinusoidal velocities. Proffitt et al. created SKE displays in which only the horizontal motion component was presented. An example of this kind of event is shown in Figure 6. We found that such displays are perceived to have the same depth magnitude as those exhibiting the full range of  $x$ - and  $y$ -axis motions. In the current Experiment 3, we show that when a common motion component was added to this stimulus, thereby producing a motion parallax display, it does not influence the perception of depth magnitude.

### Experiment 1

In this experiment, we presented motion parallax and SKE displays that had been equated in all aspects except for the presence or absence of the common motion component. Observers provided indications of perceived depth magnitude and depth order.

#### Method

**Subjects.** Sixteen University of Virginia undergraduates (8 women and 8 men) participated in this experiment. All of them were naive to the purposes of this experiment, and none had previously participated in depth-from-motion experiments. They were paid \$5 for their participation. The data from 1 additional subject were not included in the analysis because of failure to follow instructions.

**Stimuli.** A stimulus display consisted of a field of random dots (hereafter referred to as the *screen window*) that simulated a motion parallax gradient of velocity. Each stimulus was located in a  $5.08 \times 5.08$ -cm square area that had a resolution of  $200 \times 200$  pixels. The simulated object was a dihedral angle, or wedge, made up of two slanted planes meeting at a horizontal ridge in relation to the observer. The width of the base of the wedge matched the width of the screen window. Four different distances between the ridge and the base of the wedge were simulated (0.25, 0.50, 0.75, and 1.50 in terms of the depth-to-width ratio), for both convex and concave angles. The random-dot displays

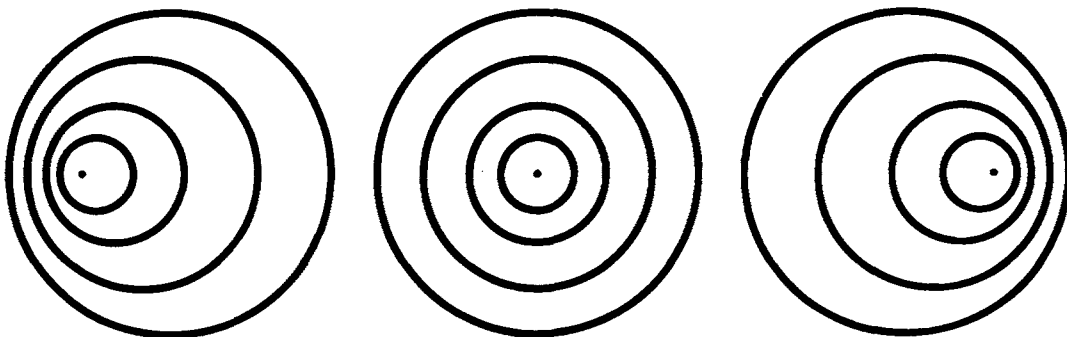


Figure 6. Three frames of a stereokinetic effect (SKE) display having only horizontal contour motions.

were generated by choosing a random distribution of positions on the projection plane (not random positions on the simulated wedge) and by displacing these dots to the surface of the simulated wedge. No gradient of texture was provided by any of the frames simulating the velocity flow field produced by the relative motion between a point of observation and the wedge. Approximately 200 dots were used in each display. Every time a dot was occluded by an edge of the window, a new dot was introduced on the opposite side of the window. The stimuli were realized by using a polar projection with a simulated distance equal to the actual viewing distance.

As in Figure 7, four kinds of velocity flow fields were simulated. The first motion parallax condition simulated the movement of an observer in relation to a stationary object, the self-produced parallax condition in the terminology of Rogers and Graham (1979). In this case, only the object-relative (SKE) component of a motion parallax velocity gradient was provided on the terminal screen. The movement of the observer's head supplied the common motion component (moving-observer condition). The observer's head was positioned on a chinrest that yoked the movement of the dots on the screen to the movements of the observer. Observers were instructed to move their heads at a comfortable speed. The range of the allowed excursion of the chinrest was equal to 8 cm. The second motion parallax condition simulated the relative motion between a moving object and a stationary observer, the externally produced parallax condition. In this case, both the object-relative component (SKE) and the observer-relative component (common component) of motion parallax were provided on the screen (stationary-observer condition). The range of the translation of the screen window was equivalent to the allowed excursion of the chinrest in the moving-observer condition. The magnitude of the observer-relative transformations was fixed across trials. Both the first and the second motion parallax conditions adequately reproduced the velocity flow field of the relative motion between an observer and an object having the shape of a dihedral angle. One SKE condition was identical to the first motion parallax condition, except that the observer's head was kept stationary (stationary-observer condition). The other SKE condition was identical to the second motion par-

allax condition except that the common component of the motion parallax gradient of velocity was nullified by a corresponding movement of the observer's head (moving-observer condition). The movement of the screen window was yoked to the movement of the chinrest so that no relative motion occurred between the two of them. In both of the SKE conditions, only the object-relative motion component was presented to the observer.

An icon representing the side view of the dot displays was shown in the upper right part of the terminal screen. The base of this icon was matched in size with the base of the dihedral angles simulated by the dot displays. Both the direction of pointing in relation to the observer and the depth of the simulated wedge could be adjusted from the keyboard.

**Apparatus.** The stimuli were presented on a high-resolution color monitor ( $1,152 \times 900$  addressable locations) under the control of a Sun 3/60 Workstation. The screen had a refresh rate of 60 Hz and was approximately photometrically linearized. An antialiasing procedure was used; for point lights falling at locations other than a pixel's center, the screen luminance was proportionally adjusted at the two relevant addressable locations. The graphics buffer was 8 bits deep (256 gray levels).

A chinrest was used that set the eye-to-screen distance at 33 cm. Cardboard models of both a concave and a convex wedge were used to demonstrate how the displays were to be interpreted. White dots were randomly scattered on the surface of both wedges.

**Design.** Four independent variables were examined in a mixed factorial design: stimulus type (motion parallax vs. SKE), simulated depth of the dihedral angle, direction of the gradient of velocity of the dots (concave vs. convex wedges), and condition of observation (stationary-observer vs. moving-observer conditions). Stimulus type was the between-subjects independent variable; all of the other variables were within subjects.

**Procedure.** The subjects were run individually in two sessions. Eight subjects were presented with velocity flow fields in which both the SKE component and the observer-relative component of motion parallax were present. In the first session, half of these subjects were run in the self-produced parallax condition; in the second session, the same subjects were presented with the externally produced parallax condition. For the other half of these subjects, the order of presentation was reversed. The remaining 8 subjects were presented with velocity flow fields in which only the SKE component of motion parallax was present. In the first session, half of these subjects were run in the stationary-observer condition; in the second session, the same subjects were run in the moving-observer condition. For the other half of these subjects, the order of presentation was reversed. The stimuli were presented in 3 blocks within each session, resulting in 48 trials. The trials of the first block of each session were treated as practice trials. The stimuli within each block were presented in random order for each subject.

Two judgments were requested for each trial. First, subjects were instructed to report the perceived sign of depth by saying either "convex" or "concave." Second, subjects estimated the depth magnitude of the stimuli by verbally instructing the experimenter on how to adjust the depth of the icon present in the upper part of the computer terminal. They were instructed to match the depth of the icon with the amount of depth they perceived in the dot displays.

After the instructions were given, the subjects placed an eye patch over their least preferred eye and positioned their heads on the chinrest. While the experiment was run, the experimental room was dark. No restriction was placed on viewing time. Each experimental session lasted about 30 min.



		DISPLAY	
		Object-relative	Object-relative + Common Component
			
OBSERVER	STATIONARY	SKE	MOTION PARALLAX
	MOVING	MOTION PARALLAX	SKE

Figure 7. Experimental design for Experiment 1.

## Results

A multivariate analysis of variance (MANOVA) was performed on two dependent variables (DVs): the perceived amount of depth (pooled across all within-subjects DVs) and the perceived depth-order relations (the proportion of responses in which faster gradients were associated with closer points). Stimulus type was the independent variable. With the use of Wilks's criterion, the combined DVs were significantly affected by the stimulus type,  $F(2, 13) = 10.91$ ,  $p < .01$ , with a degree of association between the independent variable and the combined DVs corresponding to  $\eta^2 = .63$ .

Univariate analyses of variance (ANOVAs) showed a significant effect for the DV corresponding to the perceived depth-order relations,  $F(1, 14) = 19.85$ ,  $p < .001$ ,  $\eta^2 = .59$ , but not a reliable difference in the perceptual performance related to motion parallax and SKE stimuli in terms of the perceived depth magnitudes,  $F(1, 14) = 3.15$ ,  $p > .05$ . The pooled within-group correlation between the two DVs was  $-.03$  (14 df).

We performed a second MANOVA, with condition of observation as the grouping variable. This analysis showed that the mean differences in the composite DV did not reliably discriminate between the two groups corresponding to the stationary-observer versus the moving-observer condition of observation,  $F(2, 29) = 0.15$ ,  $p > .05$ .

**Perceived depth.** The mean perceived depth for each simulated depth and for each stimulus is shown in Figure 8. A profile analysis was performed to examine more closely the effect of the within-subjects independent variable (simulated depth) on the amount of perceived depth. Stimulus type was the between-subjects independent variable. For the levels test, we found no reliable difference with Wilks's criterion for the perceptual performance associated with either motion parallax or SKE velocity flow fields when the scores were averaged over all simulated distances,  $F(1, 14) = 3.15$ ,  $p > .05$ ,  $\eta^2 = .18$ . Overall mean perceived depth for each stimulus was 2.34 cm for motion

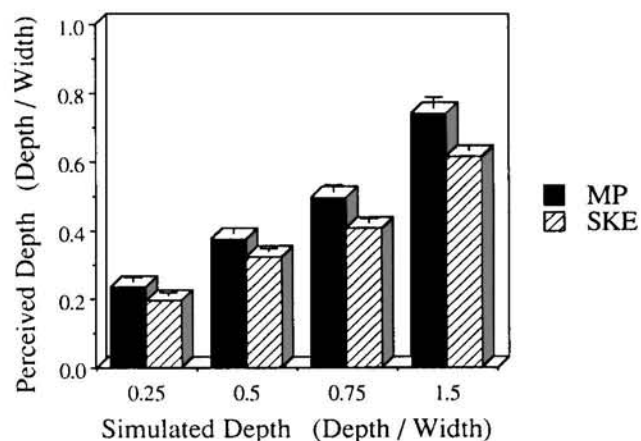


Figure 8. Experiment 1: Mean perceived depth plotted as a function of simulated depth. (The values are expressed as depth-to-width ratios. Vertical bars represent standard errors.)

parallax stimuli and 1.96 cm for SKE stimuli. When averaged across stimuli, however, the profile of the perceived depth magnitudes was found to deviate significantly from flatness,  $F(3, 12) = 63.62$ ,  $p < .0001$ ,  $\eta^2 = .94$ . As shown in Figure 8, the effect of the variable simulated depth was in the correct direction, indicating that an increase in the amount of simulated depth was associated with a corresponding increase of the perceived depth magnitudes. We found no significant deviation from parallelism of the profiles,  $F(3, 12) = 0.69$ ,  $p > .05$ ,  $\eta^2 = .15$ .

**Depth order.** We performed a four-way frequency analysis to develop a hierarchical logit model, using as data the frequencies of correct and incorrect judgments for each subject and for each combination of the variables analyzed. For the experimental conditions in which the velocity flow field in relation to the observer was determined only by the SKE component of motion parallax, the responses in which faster velocities were associated with closer points were treated as correct judgments. Variables analyzed were stimulus type, simulated depth, direction of the gradient of velocity, and condition of observation.

Stepwise selection by simple deletion of effects produced a model that included 1 two-way association (Simulated Depth  $\times$  Direction of the Gradient of Velocity) and the first-order effects of stimulus type, simulated depth, and direction of the gradient of velocity. The model had the following likelihood ratio:  $\chi^2(23, N = 512) = 9.86$ ,  $p = .99$ , indicating a good fit between observed frequencies and expected frequencies generated by the model. The largest standardized residual had a  $z$  value equal to  $-0.93$ , which is not statistically significant. The measures of strength of association, however, indicated that only about 20% of the variance of the DV was accounted for by this model (entropy = 0.18, concentration = 0.19). The variable stimulus type was significant,  $\chi^2(1, N = 512) = 28.60$ ,  $p = .0001$ . In the experimental conditions in which both the SKE and the observer-relative components of motion parallax were present, 92% of the sign-of-depth judgments were correct. In the conditions in which the velocity flow field in relation to the observer was determined by the SKE component of motion parallax only, faster velocities were judged to be closer 62% of the time. Also significant was the interaction between simulated depth and direction of the gradient of velocity,  $\chi^2(1, N = 512) = 14.84$ ,  $p < .01$ . This interaction indicated that the proportion of the responses coded as correct decreased with the increase of the amount of the simulated depth for the concave wedges, whereas the opposite was true for the concave wedges. None of the other first-order effects or higher order associations reached statistical significance.

## Discussion

The results of Experiment 1 indicate that in the motion parallax condition, perceived depth magnitudes were a monotonically increasing function of simulated depth, as it has often been reported in the literature. We obtained an equivalent positive monotonic increase of perceived depth magnitudes in the SKE condition as well. In this case,



however, the perceived depth magnitudes cannot be interpreted to be a function of the simulated depth because SKE displays do not simulate differential depth magnitudes but always geometrically specify an infinite distance. In the SKE condition, therefore, the monotonic increase of perceived depth must be attributed to the differential amounts of object-relative motions. Taken together, the results of the motion parallax and SKE conditions indicate that the differential perceived depth magnitudes were in both cases determined by object-relative motions only. The common observer-relative motions (in the motion parallax condition) did not influence the monocular derivation of depth magnitudes. Observer-relative transformations, instead, influenced only the near-far signing of perceived sequential depth.

These results are consistent with the hypothesis that the derivation of monocular depth from motion is mediated by a perceptual heuristic. Depth magnitudes were based solely on the object-relative motion component, although by itself these transformations do not define differential depths. Thus, the perceptual system must be relating differences in the magnitude of object-relative transformations to an inherent default or bias.

As to the judgments of the sign of depth, note that the significant main effect of stimulus type (motion parallax vs. SKE) did not replicate the findings of Braunstein and Andersen (1981). Braunstein and Andersen presented stationary observers with velocity flow fields produced by either SKE patterns (in their terminology, gradients of velocity with minimum dot speed of  $0^\circ/\text{s}$ ) or motion parallax gradients of velocity. No effect of minimum velocity was found in the analysis of subjects' discriminations between concave (center-far) and convex (center-near) angles. This result was taken as indicating an equivalence of the perceptual response to SKE and motion parallax velocity flow fields with regard to depth order. A methodological difference may explain the discrepancy between these two studies. Braunstein and Andersen (1981) used a within-subjects design, whereas in the present study we examined the variable of interest between subjects. If the purpose of the research is to investigate the difference of the perceptual response to SKE and motion parallax velocity flow fields, it is important to control for the effects of expectation from one stimulus type to the other by means of a between-subjects experimental design. H. Ono and Steinbach (1990) advanced the hypothesis that expectations play an important role in determining the perceptual outcome in situations in which the pattern of relative motion on the screen does not have a single interpretation. In experiments in which subjects had to report perceived depth and perceived motion of ambiguous motion parallax displays, different results have been obtained by using the same stimuli in a between-subjects and a within-subjects experimental design.

## Experiment 2

In Experiment 1, a horizontal wedge was used because it completely eliminates changing texture gradient information. Surfaces that intersect at a horizontal edge project the

same gradient of dot density on the picture plane at all horizontal displacements in relation to the observer. Experiment 2 used a vertical wedge and thereby introduced changing texture gradient information corresponding to the velocity gradient information. Imagine a vertically oriented wedge that is moving away from the line of sight to the left. The density of texture on the wedge's farthest surface will become compressed, whereas that on its near surface will decrease. Similarly, traditional SKE displays, consisting of contours, always present a gradient of relative contour proximity as the contours move back and forth. The purpose of Experiment 2 was to determine the effect of this further source of depth information.

## Method

**Subjects.** Eight University of Virginia students (4 women and 4 men) participated in this experiment. Two (graduate students) were familiar with depth-from-motion displays but were not involved in research on the phenomenon. The other 6 (4 undergraduates and 2 graduate students) had never seen these kinds of stimuli before. They were paid \$5 for their participation. None of them had participated in Experiment 1.

**Stimuli.** The stimuli were essentially the same as in Experiment 1, except that the simulated object was a wedge made up of two slanted planes meeting at a vertical ridge in relation to the observer. The relative motion between a vertical dihedral angle and an observer produced a continuous variation in slant of the simulated planes. This change in slant, which corresponded to different amounts of foreshortening in different points of the trajectory of the wedge, in turn produced a gradient of texture in stationary frames of the dot displays.

As in Experiment 1, two different stimuli were defined. The first stimulus type corresponded to the velocity flow field providing both object-relative and observer-relative transformations (motion parallax). The second stimulus type corresponded to the velocity flow field determined by object-relative transformations only (SKE). Because in Experiment 1 the variable condition of observation (static observer vs. moving observer) did not have a significant influence on the performance, in Experiment 2 all subjects were run in the stationary-observer condition.

**Apparatus.** The apparatus was the same as in Experiment 1.

**Design.** Three independent variables were examined in a mixed factorial design: stimulus type (motion parallax vs. SKE), simulated depth (the same parameters as in Experiment 1 were used), and direction of the gradient of velocity of the dots (concave vs. convex wedges). Stimulus type was the between-subjects variable; all other variables were within subjects.

**Procedure.** The subjects were run individually in one session. Four subjects were presented with velocity flow fields in which both the object-relative and observer-relative components of motion parallax were present. All of them were run in the stationary-observer, externally produced parallax condition. The remaining 4 subjects were presented with velocity flow fields in which only the SKE component of motion parallax was present. Otherwise, procedure and instructions were the same as in Experiment 1.

## Results

We performed a MANOVA on the DVs corresponding to the perceived amount of depth and the perceived depth-order relations, following the procedure used for the anal-



ysis of the data of Experiment 1. The composite score created by the combined DVs was significantly affected by the stimulus type,  $F(2, 5) = 10.13$ ,  $p < .05$ ,  $\eta^2 = .80$ .

Univariate ANOVAs showed that the type of the stimulus had a significant effect on the perceived depth-order relations,  $F(1, 6) = 24.05$ ,  $p < .01$ ,  $\eta^2 = .80$ , but not on the amount of perceived depth,  $F(1, 6) = 0.01$ ,  $p > .05$ . The pooled within-group correlation between the DVs was .08 (6 df).

**Perceived depth.** The mean perceived depth for each simulated depth and for each stimulus type is presented in Figure 9. We performed a profile analysis on the perceived amount of depth associated with the differential simulated distances, with stimulus type as the between-subjects independent variable. We found no reliable difference in the perceptual performance associated with motion parallax and SKE velocity flow fields,  $F(1, 6) = 0.01$ ,  $p > .05$ ,  $\eta^2 = 0$ . Overall mean perceived depth for each stimulus was 2.01 cm for motion parallax stimuli and 2.03 cm for the SKE stimuli. When averaged across stimuli, however, the profile of perceived depth magnitudes significantly deviated from flatness in the predicted direction,  $F(3, 4) = 40.48$ ,  $p < .01$ ,  $\eta^2 = .97$  (see Figure 9). No significant deviation from parallelism was found,  $F(3, 4) = 1.14$ ,  $p > .05$ ,  $\eta^2 = .46$ .

We performed a profile analysis to compare the perceived-depth judgments of Experiment 2 with the corresponding distribution of judgments in Experiment 1 (only the subset of judgments obtained in the stationary-observer condition was used). This analysis showed that the gradient of texture introduced in the dot displays of Experiment 2 did not significantly influence perceived depth,  $F(1, 14)$ ,  $p > .05$ ,  $\eta^2 = .04$ .

**Depth order.** We performed a three-way frequency analysis to develop a hierarchical logit model, using as data the frequencies of correct and incorrect judgments for each subject and for each combination of the variables analyzed. For the experimental conditions in which the velocity flow field in relation to the observer was determined

only by the SKE component of motion parallax, the responses in which faster velocities were associated with closer points were treated as correct judgments, following the procedure used in the analysis of the data of Experiment 1. Variables analyzed were stimulus type, simulated depth, and direction of the gradient of velocity.

Stepwise selection by simple deletion of effects produced a model that included only the first-order effect of stimulus type. The model had the following likelihood ratio:  $\chi^2(14, N = 256) = 9.52$ ,  $p = .78$ , indicating a good fit between observed frequencies and expected frequencies generated by the model. The largest standardized residual was not statistically significant ( $z = 1.79$ ). The measures of strength of association, however, indicated that reduction in uncertainty in prediction by the model was low (entropy = 0.10, concentration = 0.12). The variable stimulus type was significant,  $\chi^2(1, N = 256) = 7.34$ ,  $p = .01$ . In the motion parallax experimental condition, 89% of the sign-of-depth judgments were correct. In the experimental conditions in which the velocity flow field was determined only by the SKE component of motion parallax, 56% of judgments associated faster velocities with closer points. None of the other first-order effects and higher order associations reached statistical significance.

## Discussion

In Experiment 2, the specification of the three-dimensional shape of the simulated wedge in the motion parallax condition was strengthened by the introduction of changing texture gradient information that corresponded to velocity gradient information. The results obtained, however, indicate that this further source of three-dimensional information did not alter the pattern of findings obtained in Experiment 1. In Experiment 2, the monocular derivation of depth magnitudes in the motion parallax condition is fully accounted for by the perceptual response to the SKE. As in Experiment 1, the observer-relative transformations were effective only in specifying the order of perceived sequential depth.

With regard to the perceptual efficacy of changing texture gradient information for the specification of three-dimensional structure, the null finding obtained in the present study does not allow general conclusions to be drawn beyond the scope of the present experimental manipulation. The changing texture gradient information provided in the present experiment may not have been salient enough to influence the perceptual response in the depth estimation task. Be that as it may, it is interesting that neither the observer-relative transformations nor the changing texture gradient information influenced the perceived depth magnitudes in the motion parallax condition. Differential perceived depth magnitudes were determined solely by the differential amounts of the object-relative motions.

## Experiment 3

In Experiment 3, we compared the perception of depth from motion parallax and SKE patterns by using nested

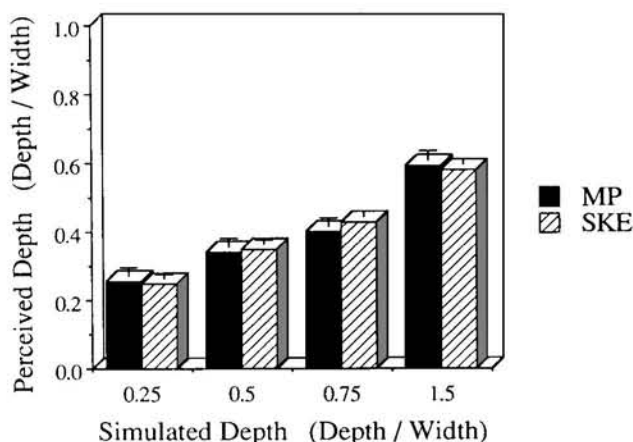


Figure 9. Experiment 2: Mean perceived depth plotted as a function of simulated depth. (The values are expressed as depth-to-width ratios. Vertical bars represent standard errors.)

contour displays similar to the traditional SKE cone. To preserve the same parameters of motion as in Experiments 1 and 2, SKE patterns were composed of a number of nested circles with only a horizontal component of motion (see Figure 6). The purpose of this experiment was to compare the perceived depth of SKE cones with that elicited by the random-dot displays of Experiments 1 and 2.

### Method

**Subjects.** Eight University of Virginia undergraduates (4 women and 4 men) participated in this experiment in partial fulfillment of a course requirement. None of them had participated in Experiment 1 or Experiment 2. Two additional subjects were excluded from the analysis because of failure to follow instructions.

**Stimuli.** As depicted in Figure 6, stimulus displays presented the relative motion between four nested circles and a dot. Each of the stimuli was located in a  $5.08 \times 5.08$ -cm square area. The diameter of the outermost circle matched the width of the screen window. Viewing distance was the same as in Experiments 1 and 2.

The motion parallax stimuli simulated the translation of a cone in relation to a stationary observer. The cone was defined by four equally spaced contours placed on its surface and by a dot in correspondence with its apex. Four different distances between the apex and the base of the cone were simulated with the same parameters as in Experiments 1 and 2 (0.25, 0.50, 0.75, and 1.50 in terms of the depth-to-width ratio). Both convex and concave cones were simulated.

All of the subjects were run in the stationary-observer condition. Two kinds of velocity flow fields were simulated. The first stimulus type corresponded to the velocity flow field produced by the relative motion between a moving cone and a stationary observer. In this case, both the SKE component and the observer-relative component of motion parallax were presented on the computer terminal. The second stimulus type was produced by the presence of only the SKE component of the previously described motion parallax velocity flow field. Because the observer's head was kept stationary, the velocity flow field presented in this condition was determined by an SKE cone with only an  $x$ -axis motion component.

An icon representing the side view of the simulated cone was present in the upper right corner of the computer terminal. The base of the icon was matched in size to the diameter of the simulated cone. The depth of the icon was adjustable from the keyboard.

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

**Design.** Three independent variables were examined in a mixed factorial design: stimulus type (motion parallax vs. SKE), simulated depth (the same parameters as in Experiments 1 and 2 were used), and direction of the gradient of velocity of the dots (concave vs. convex cones). Stimulus type was the between-subjects variable; all other variables were within subjects.

**Procedure.** Procedure and instructions were the same as in Experiment 2.

### Results

We performed a MANOVA on the DVs corresponding to the amount of perceived depth and the perceived depth-order relations, following the procedure used in the analysis

of the previous experiments. The mean difference between the composite score created by the combined DVs was statistically significant,  $F(2, 5) = 6.70$ ,  $p < .05$ ,  $\eta^2 = .73$ . None of the univariate ANOVAs, however, reached significance. The pooled within-group correlation between the DVs was .73 (6 df).

**Perceived depth.** The mean perceived depth for each simulated depth and for each stimulus type is shown in Figure 10. We performed a profile analysis on the perceived depth magnitudes corresponding to differential simulated distances, which showed no reliable difference in the perceptual performance associated with motion parallax and SKE velocity flow fields in the levels test,  $F(1, 6) = 2.24$ ,  $p > .05$ ,  $\eta^2 = .27$ . Overall mean perceived depth for each stimulus was 2.51 cm for the motion parallax stimuli and 2.06 cm for the SKE stimuli. When averaged across stimulus types, however, the profile of the perceived depth magnitudes significantly deviated from flatness in the correct direction,  $F(3, 4) = 89.96$ ,  $p < .001$ ,  $\eta^2 = .98$  (see Figure 10). We found no significant deviation from parallelism,  $F(3, 4) = 0.08$ ,  $p > .05$ ,  $\eta^2 = .36$ .

We performed a profile analysis to compare the perceived-depth judgments of Experiment 3 with the results obtained in Experiment 2. The variable of interest (random-dot vs. nested circles displays) did not reach significance,  $F(1, 14) = 1.63$ ,  $p > .05$ ,  $\eta^2 = .10$ , indicating that the perceived depth associated with the displays used did not significantly differ across Experiments 2 and 3.

**Depth order.** We performed a three-way frequency analysis on the frequencies of correct judgments for each subject and for each stimulus. For the SKE displays, the trials in which faster velocities had been associated with closer points were computed as correct judgments, as in the analysis of Experiments 1 and 2. Variables analyzed were stimulus type, simulated depth, and direction of the gradient of velocity.

Unexpectedly, this analysis showed that the perceptual response to the two stimulus types (motion parallax vs. SKE) did not differ significantly with regard to depth or-

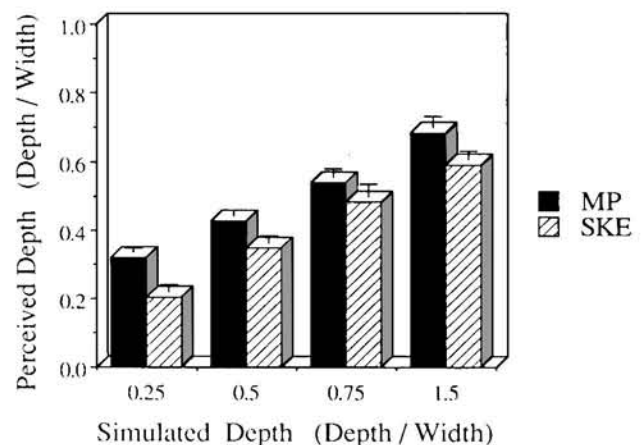


Figure 10. Experiment 3: Mean perceived depth plotted as function of simulated depth. (The values are expressed as depth-to-width ratios. Vertical bars represent standard errors.)

der,  $\chi^2(1, N = 256) = 1.11, p > .05$ . In the motion parallax condition, 61% of the sign-of-depth judgments were correct. In the condition in which the velocity flow field was determined only by the SKE component of motion parallax, 48% of judgments associated faster velocities with closer points. None of the other first-order effects or higher order interactions resulted in significance.

We performed a follow-up experiment to clarify the unexpected result that showed no significant difference between the distributions of the depth-order judgments for the motion parallax and SKE conditions. We advanced the hypothesis that the low accuracy of subjects' judgments in the case of the motion parallax velocity flow fields occurred because these displays were perceived as performing a rotation as well as a translation. A component of perceived rotation has been taken as evidence in the literature that an orthographic analysis, which does not allow the recovery of depth-order information, is applied by the perceptual system to the stimulus displays (Braunstein, 1988). Therefore, we expected incorrect judgments to be more likely associated with trials in which the perception of apparent rotation was reported. Five subjects were presented with the motion parallax velocity flow fields of Experiment 3. They were asked to report both depth order and amount of apparent rotation (no rotation, slight rotation, mostly rotation), following the procedure of Braunstein and Andersen (1981). Judgments were collected for 48 trials per subject. The results obtained were consistent with the hypothesis formulated. There was a high proportion of trials in which a strong component of rotation was perceived (.41). The proportion of incorrect judgments for these trials was .49 versus .25 for the trials in which no rotation or only slight rotation was reported.

### Discussion

The findings of the present experiment show that in the motion parallax condition, the observer-relative transformations did not contribute to the perceived depth magnitudes. Because the stimuli used in the SKE condition of Experiment 3 are equivalent to the traditional SKE displays produced with the turntable methodology (Proffitt et al., 1992), the results obtained provide the most direct support to the hypothesis that depth perception elicited by a motion parallax gradient of velocity can be accounted for solely by the perceptual response to the SKE.

In Experiment 3, the introduction of figural information did not modify the magnitude of perceived depth elicited by the random-dot displays of Experiment 2. Although we do not want to generalize from this result beyond the parameters used, we do want to stress that within the range of the present experimental manipulation, the only determinant of the monotonic increase in perceived depth magnitudes was the differential amount of object-relative motion.

### Comparison of Experiments 1–3

The mean perceived depth for each simulated depth and for each experiment is shown in Figure 11. We performed

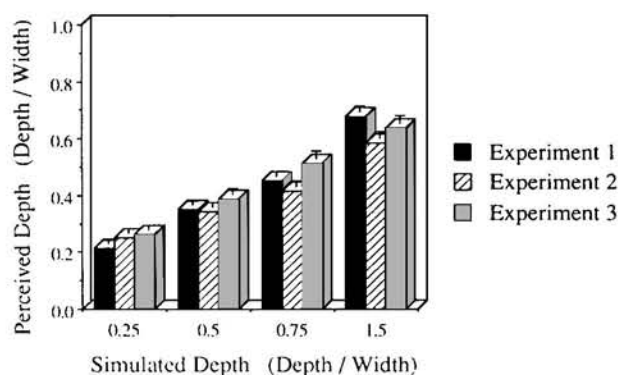


Figure 11. Mean perceived depth in Experiments 1–3 plotted as a function of simulated depth. (The values are expressed as depth-to-width ratios. Vertical bars represent standard errors.)

two discriminant function analyses, one for the data set corresponding to the motion parallax stimuli and one for the data set corresponding to the SKE stimuli. These analyses used two DVs, the perceived amount of depth and the perceived depth-order relations, as predictors of membership in three groups corresponding to each of the experiments described so far. The same procedure used for the previous MANOVAs was followed in these analyses.

### Results

Through Pillai's trace criterion, one discriminant function reached significance in the case of the motion parallax data set,  $F(4, 26) = 4.74, p < .01$ . This discriminant function accounted for 76% of the between-group variability, and it maximally separated the group corresponding to Experiment 3 from the other two groups. Centroids in reduced space for Experiments 1–3 were, respectively, 0.95, 0.89, and  $-2.80$ .

An analogous discriminant function analysis performed on the data set corresponding to the SKE stimuli did not reveal any reliable difference between the three groups corresponding to each of the previous experiments. Through Pillai's trace criterion, we obtained a nonsignificant result,  $F(4, 26) = 0.28, p > .05$ .

### Discussion

In the case of the SKE stimuli, the perceptual performance was not significantly influenced by the experimental manipulations that differentiated the three experiments considered here. The performance of the observers, in terms of both the perceived amount of depth and the perceived depth-order relations, was not reliably affected by the introduction of the gradient of texture used in Experiment 2. Furthermore, when the object-relative transformation had been equated, the perceptual performance associated with random-dot displays (Experiment 2) did not differ from the perceptual performance associated with displays endowed with figural information (Experiment 3).

On the other hand, the perceptual performance was significantly affected by the manipulations that differentiated

the present experiments in the case of the motion parallax stimuli. The significance of the discriminant function can be interpreted by considering the correlation matrix between the predictors and the composite score, which maximally separates the groups. The loading matrix showed a high loading for the DV corresponding to the perceived depth-order relations (.97) and a negligible loading for the DV corresponding to the perceived amount of depth (-.14). The dimension along which the group corresponding to Experiment 3 differs from the other two groups can thus be interpreted in terms of the DV corresponding to the perceived sign of depth, without any appreciable contribution of the DV corresponding to the perceived amount of depth. Therefore, we can conclude that in the case of the motion parallax stimuli as well, when the object-relative transformation had been equated, the introduction of either texture gradient information or figural information did not significantly affect the perceived depth magnitudes in the set of experiments considered here.

### Experiment 4

Experiment 4 was designed to test the hypothesis that perceived depth magnitudes are the product of a perceptual heuristic that derives depth magnitudes from both the projected size of the stimuli and the amount of the object-relative transformations. The stimuli were analogous to the displays used in Experiment 1. In the present experiment, however, the width of the simulated wedges was systematically varied. According to the compactness heuristics, we predicted that perceived depth should covary with the projected width of the stimuli.

### Method

**Subjects.** Forty University of Virginia students (20 women and 20 men) participated in this experiment in partial fulfillment of a course requirement. All of them were naive to the purpose of this experiment, and none had previously participated in any depth-from-motion experiments.

**Stimuli.** The stimuli were essentially the same as in Experiment 1 except that the width of the screen window assumed four different sizes (5.08, 3.81, 2.54, and 1.27 cm), whereas its length was kept constant (5.08 cm). As in Experiment 1, the width of the base of the simulated wedge matched the size of the screen window. Four different distances between the ridge and the base

of the wedge were simulated (1.90, 3.81, 5.71, 7.62 cm). Table 1 presents the simulated distances (in terms of the depth-to-width ratio) for the different screen-window sizes. As in the previous experiments, SKE stimuli were equivalent to the motion parallax stimuli except for the presence of the observer-relative motion component in the simulated gradient of velocity.

**Apparatus.** The apparatus was the same as in Experiments 1-3.

**Design.** Four independent variables were examined in a mixed factorial design: stimulus type (motion parallax vs. SKE), simulated depth of the dihedral angle, direction of the gradient of velocity of the dots (concave vs. convex wedges), and size of the screen window. Stimulus type was the between-subjects independent variable; all other variables were within subjects.

**Procedure.** Procedure and instructions were the same as in Experiments 2 and 3, except that three judgments were requested for each trial: the perceived amount of depth, the perceived sign of depth, and the perceived amount of rigidity. Perceived rigidity was defined on a scale ranging from 0 (*complete elasticity of the perceived object*) to 1 (*complete rigidity of the perceived object*). Subjects estimated the rigidity by verbally instructing the experimenter on how to adjust a cursor on a scale present in the upper part of the computer terminal.

### Results

By following the procedure used in the analysis of the previous experiments, we performed a MANOVA on three DVs: the amount of perceived depth, the perceived depth-order relations, and the perceived rigidity. The result of this analysis indicated that the composite score created by the combined DVs was significantly affected by the grouping variable (motion parallax vs. SKE),  $F(3, 36) = 7.89$ ,  $p < .001$ ,  $\eta^2 = .40$ .

Univariate ANOVAs showed that the variable stimulus type had a significant effect on the perceived sign of depth,  $F(1, 38) = 17.32$ ,  $p < .001$ ,  $\eta^2 = .31$ . But neither the amount of perceived depth— $F(1, 38) = 3.93$ ,  $p > .05$ ,  $\eta^2 = .03$ —nor the perceived rigidity— $F(1, 38) = 1.19$ ,  $p > .05$ ,  $\eta^2 = .09$ —was significantly affected by the grouping variable (motion parallax vs. SKE). The pooled within-group correlation between perceived depth and perceived depth-order relations was .08, the correlation between perceived depth and perceived rigidity was -.08, and the correlation between perceived depth-order relations and perceived rigidity was -.13 (38 df).

**Perceived depth.** The mean perceived depth for each simulated depth and for each size of the screen window is shown in Figures 12 and 13. We performed a profile analysis on the perceived amount of depth, with one between-subjects independent variable (stimulus type) and two within-subjects independent variables (simulated depth and size of the screen window). The levels test did not reveal any significant difference in the perceptual performance associated with motion parallax and SKE stimuli when the scores were averaged over the within-subjects variables,  $F(1, 38) = 1.19$ ,  $p > .05$ ,  $\eta^2 = .03$ . Overall mean perceived depth for each stimulus was 1.66 cm for the motion parallax stimuli and 1.89 cm for the SKE stimuli. The effect of the variable simulated depth was significant,  $F(3, 36) = 50.38$ ,  $p < .0001$ ,  $\eta^2 = .81$ , and so was the effect of the

Table 1  
*Simulated Distances for the Different Screen Window Sizes Expressed in Terms of Depth-to-Width Ratios*

Screen window	Simulated depth			
	1.90	3.81	5.71	7.62
5.08	0.37	0.75	1.12	1.50
3.81	0.50	1.00	1.50	2.00
2.54	0.75	1.50	2.25	3.00
1.27	1.50	3.00	4.50	6.00

*Note.* Screen window and simulated depth values are in centimeters.



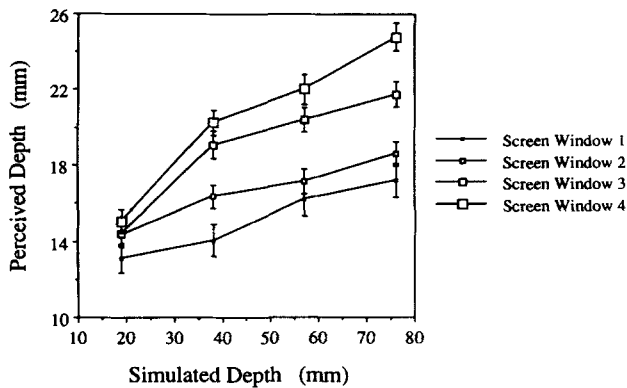


Figure 12. Experiment 4: Mean perceived depth plotted as a function of simulated depth. (Vertical bars represent standard errors. In this figure as in the following ones, the height of the screen windows are 1.27 cm for Screen Window 1, 2.54 cm for Screen Window 2, 3.81 cm for Screen Window 3, and 5.08 cm for Screen Window 4.)

size of the screen window,  $F(3, 36) = 14.47, p < .0001, \eta^2 = .55$ . These two variables, however, were found to interact significantly with one another,  $F(9, 30) = 4.80, p < .001, \eta^2 = .59$ .

**Perceived rigidity.** The mean perceived rigidity for each simulated depth and for each size of the screen window is shown in Figures 14 and 15. In a profile analysis, we examined the distribution of the rigidity judgments, with the same variables as the previous analysis: stimulus type, simulated depth, and size of the screen window. The levels test did not reveal a reliable difference in the performance related to the two stimulus types,  $F(1, 38) = 3.93, p > .05, \eta^2 = .09$ . The variable simulated depth reached significance,  $F(3, 36) = 11.78, p < .0001, \eta^2 = .49$ , and the effect of the variable size of the screen window was also significant,  $F(3, 36) = 8.90, p < .001, \eta^2 = .43$ . None of the interactions between the variables reached significance.

**Depth order.** We performed a four-way frequency analysis to develop a hierarchical logit model, using as data the frequencies of correct and incorrect judgments for each subject and for each combination of the variable analyzed. For the experimental conditions in which the velocity flow field in relation to the observer was determined only by the SKE component of motion parallax, we treated the responses in which faster velocities were associated with closer points as correct judgments, following the procedure used in the analysis of the data of Experiment 1–3. Variables analyzed were stimulus type, simulated depth, direction of the gradient of velocity, and size of the screen window.

Stepwise selection by simple deletion of effects produced a model that included the first-order effects of stimulus type and direction of the gradient of velocity, and the two-way association between the same two variables. The model had the following likelihood ratio:  $\chi^2(60, N = 2560) = 38.88, p = .98$ , indicating a very good fit between observed frequencies and expected frequencies generated by the model. The largest standardized residual had a  $z$  value

equal to 1.87, which is not statistically significant. The measures of strength of association, however, indicated that the DV was not well predicted by the variables used in this model (entropy = 0.04, concentration = 0.05). The variable stimulus type was significant,  $\chi^2(1, N = 2560) = 54.19, p = .0001$ . In the motion parallax experimental condition, 78% of the sign-of-depth judgments were correct. In the experimental conditions in which the velocity flow field was determined only by the SKE component of motion parallax, 59% of judgments associated faster velocities with closer points. The two-way association between stimulus type and direction of the gradient of velocity was also significant,  $\chi^2(1, N = 2560) = 13.80, p = .001$ . This interaction indicated that faster velocities were more likely to be associated with closer points in the case of convex stimuli than in the case of concave ones (82% vs. 73%) for the motion parallax displays, whereas the opposite was true for the SKE displays (54% vs. 64%). None of the other first-order effects or higher order associations reached statistical significance.

## Discussion

It has often been reported in the SKE literature that the amount of depth perceived in SKE displays depends on the size of the two-dimensional pattern used (Musatti, 1924; Zanforlin, 1988). In essence, SKE cones appear to be just slightly deeper than the diameter of the largest contour. The main result of Experiment 4 is the finding that a similar relation between perceived depth magnitudes and projective size holds for perceiving depth in motion parallax displays. In fact, the results of Experiment 4 indicate that an increase in the dimensions of the screen window was associated with an increase in perceived depth magnitudes, even though the simulated amount of depth was kept constant. Moreover, perceived depth magnitudes were never much greater than object width, even for the largest simulated depths. The maximum depth rating was about 1.4, corresponding to a simulated depth of 6 (in terms of the depth-to-width ratio). These findings are consistent with the hypothesis that perceived depth magnitudes are scaled in relation to the pro-

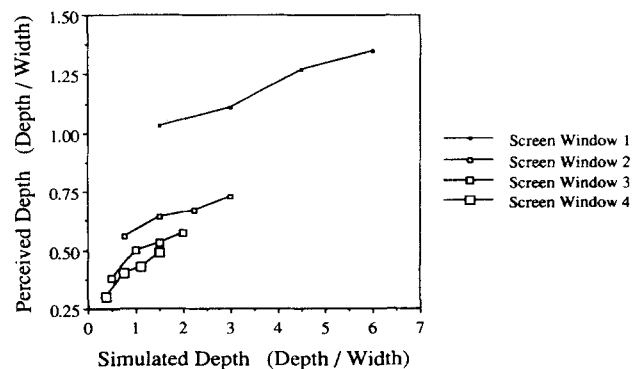


Figure 13. Experiment 4: Mean perceived depth plotted as a function of simulated depth. (The values are expressed as depth-to-width ratios.)

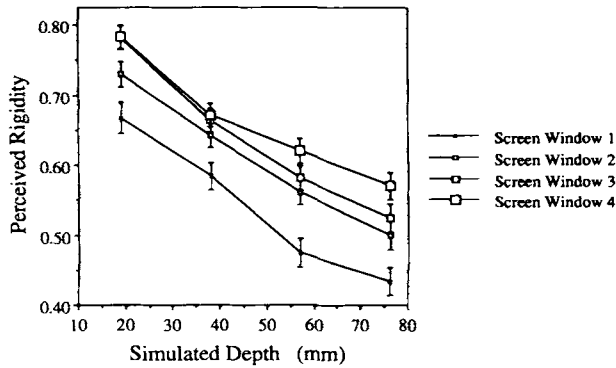


Figure 14. Experiment 4: Mean perceived rigidity plotted as a function of simulated depth. (Vertical bars represent standard errors.)

jected size of the stimulus and never greatly exceed unity in terms of the depth-to-width ratio.

For judgments of depth magnitudes, the interaction between simulated depth and size of the screen window can be interpreted by considering the different range of simulated depths spanned by the stimuli corresponding to each of the different screen windows' sizes (see Table 1). Small screen-window stimuli spanned a large range of simulated depth magnitudes, and correspondingly, perceived depth magnitudes approached ceiling sooner than in the large screen-window stimuli, which spanned a much smaller range.

The results for the rigidity ratings provide a converging source of evidence that the structural properties of the perceived objects are influenced by the projective size of the stimuli (see Figure 14). For each screen-window size, ratings of perceived rigidity decreased with increases of its object-relative motions. More important, however, the rigidity ratings covaried with the size of the screen windows. For any simulated depth, perceived rigidity decreased as projected size of the stimulus decreased; in other words, perceived rigidity decreased with the increase of the discrepancy between the amount of simulated depth and the projective size of the stimuli.

In conclusion, the results of Experiment 4 replicate the findings obtained in the previous experiments: No reliable

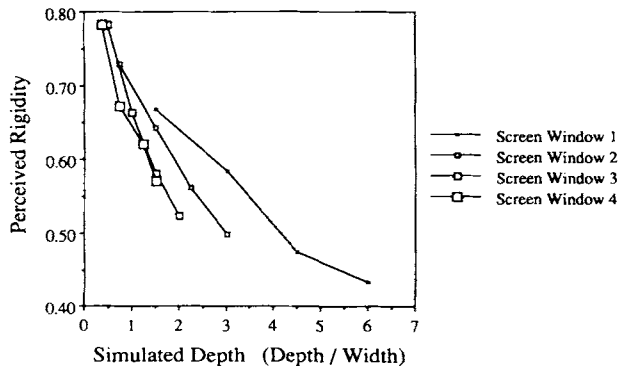


Figure 15. Experiment 4: Mean perceived rigidity plotted as a function of simulated depth. (The values are expressed as depth-to-width ratios.)

difference was found in the perceptual derivation of depth magnitudes from either motion parallax or SKE stimuli. The observer-relative component of motion did not influence perceived depth magnitudes. As in the previous experiments, however, we found a reliable difference between the perceptual performance associated with motion parallax and SKE stimuli with regard to the perceived sign of depth: Faster velocities were more likely to be associated with closer points in the case of motion parallax displays than in the case of SKE stimuli.

## General Discussion

In four experiments, we found that the perceptual derivation of depth magnitudes from motion parallax uses only a subset of the information available in the optic flow field, this being the object-relative component that constitutes the stimulus basis of the SKE. In particular, we found that the magnitude of depth perceived in motion parallax equals that seen in the SKE when their object-relative transformations are equated. This finding was obtained in Experiment 1, 2, and 4 with displays consisting of random-dot patterns. The same result was replicated in Experiment 3 by using SKE cones that Proffitt et al. (1992) have shown to be phenomenally equivalent to the traditional SKE displays that Musatti (1924) produced with the turntable methodology.

Unlike small rigid object rotations, in which the transformations specifying the angle of rotation are not sufficiently salient to be of much perceptual utility, in motion parallax the magnitude of the effective object rotation in relation to the observer is given robustly by the angle of object displacement (see Figure 5). Thus, the finding that observer-relative transformations are not used in the perceptual derivation of depth magnitudes from motion parallax is particularly surprising. This finding is even more striking when we consider that object-relative transformations do not geometrically specify differential object-relative depths. The presence of object-relative transformations in the absence of observer-relative motions is geometrically consistent with an infinite depth magnitude only.

The results of the present experiments in conjunction with those of Proffitt et al. (1992) have implications for the computational modeling of human vision. Formal analyses of the optic flow field have shown that given certain assumptions, projective transformations are sufficient to derive the three-dimensional structure of the distal objects (Koenderink, 1986; Longuet-Higgins & Prazdny, 1980; Ullman, 1979). At the same time, it has been shown that people are highly sensitive to projective displacements (Braunstein, Hoffman, & Pollick, 1990). Together, these findings suggest that the psychological process of deriving depth information from optic flow fields can be characterized in terms of a canonical geometrical analysis of the input to the perceptual system.

The present investigation, however, does not support such an interpretation. Our findings indicate that the perceptual process of deriving monocular depth from motion parallax cannot be described in terms of a canonical geometrical

analysis of the prevailing stimulus conditions, even when the stimulus information fully and robustly specifies the spatial properties of the distal objects and would therefore allow a geometrical derivation of depth magnitudes. Instead, we propose that the perceptual derivation of monocular depth from motion is better described as a heuristic process of which the SKE is symptomatic.

The monocular depth from motion parallax system uses object-relative transformations only, the magnitudes of which are combined with the size of the projected object's outline, thereby revealing a default regularity bias such as a compactness assumption. Consistent with this proposal are the findings of large depth underestimations that have been reported in the motion parallax literature. Figure 16 summarizes the results of studies that obtained measures of depth magnitudes perceived in motion parallax. Along with a summary of the results of the present studies, we have recoded here the data of others into depth-to-width ratios. As Figure 16 shows, accuracy of perceptual estimates has been found in studies that simulated objects with depth-to-width ratios in the neighborhood of unity and below. Underestimations of depth have been found when the depth of the simulated objects exceeded their width. Finally, with large magnitudes of depth, motion parallax has often been found to be an ineffective source of three-dimensional information, with observers in these situations reporting a streaming motion of the dots within the two-dimensional display (M. Ono et al., 1986).

With regard to the compactness assumption, we mean it to represent the minimal assumptive requirement needed to derive a depth magnitude in the absence of sufficient infor-

mation. In essence, the compactness assumption states that objects are likely to be about as deep as they are wide. To assume otherwise would be tantamount to assuming that objects have specific observer-relative orientations. For example, consider the two-dimensional projection of an elongated object. It is only in a small subset of possible observer-relative orientations—those in which the object is pointing toward the observer—that the compactness assumption is grossly violated. Because orientation-specific information is not provided in the projection of unfamiliar objects, orientation-specific assumptions are generally unwarranted. The compactness assumption therefore represents a minimal hypothesis in situations in which the object-relative transformations are the only source of information used by the perceptual system.

Our approach to understanding the perception of depth from monocular motion information relates everyday perceptual functioning to a particular illusion, the SKE. Now, some who promote an ecological approach to visual perception have disparaged the study of visual illusions as not providing useful insights into the normal psychological processes involved in everyday perception (Gibson, 1979). We take the research reported here as evidence that illusions can shed light on everyday perceptual functioning. Moreover, we believe that the inherent perceptual biases that are reflected in illusions need not necessarily be ecologically invalid. As currently construed, the compactness assumption makes ecological sense because the orientational specificity of objects in the world is relative to gravity but not to particular points of observation.

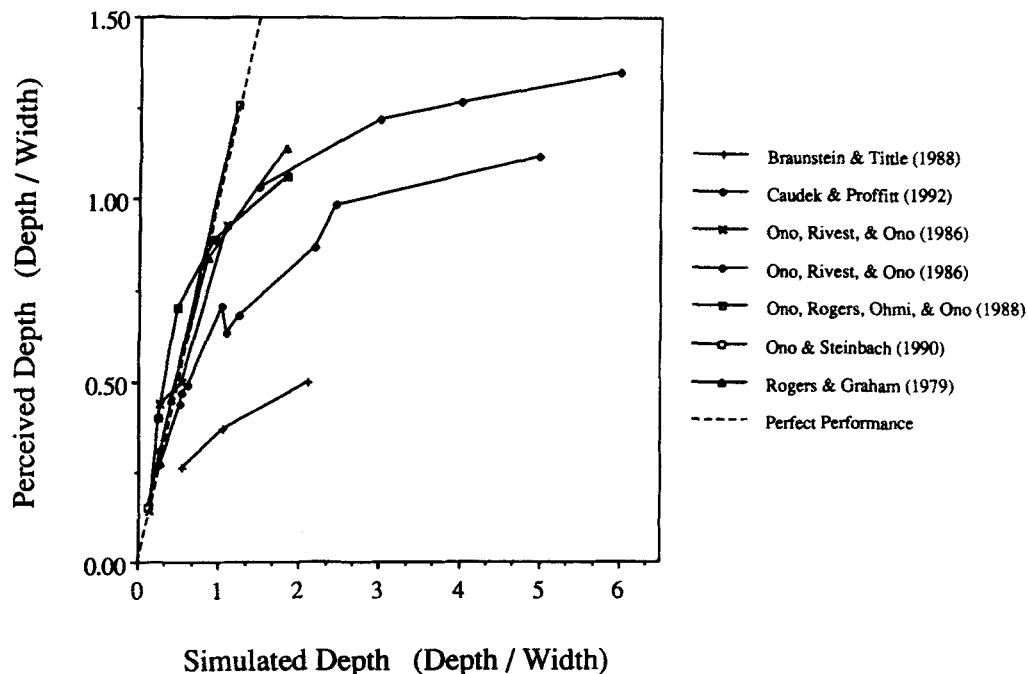


Figure 16. Mean perceived depth reported in recent investigations on the perceptual efficacy of motion parallax. (Perceived depth magnitudes are expressed as depth-to-width ratios. Values plotted are approximate in some cases because of the partial information provided.)



## Conclusion

The present investigation shows that a motion parallax velocity flow field can be decomposed into two components: observer-relative transformations and object-relative transformations. Both of these components are necessary to characterize adequately the spatial properties of distal objects. Our findings indicate, however, that the perceptual efficacy of motion parallax in determining depth magnitudes depends on only a subset of the available information. In motion parallax, observer-relative transformations contribute little or nothing to perceiving depth magnitudes. Their influence is primarily in specifying depth order.

The SKE patterns elicit depth perceptions because they preserve the object-relative component of a motion parallax velocity flow field. The SKE patterns and motion parallax gradients of velocity evoke the same amounts of depth when they exhibit the same object-relative transformations and are of the same projected size. Perceived depth in both situations is a function of these two variables.

Because in the SKE observer-relative transformations are not present, and in motion parallax they are not used, the perceptual derivation of depth from monocular motion information cannot be accounted for in terms of a canonical geometrical analysis of the stimulus information. Assumptions inherent to the perceptual system augment the perception of depth from monocular motion.

## References

- Braunstein, M. L. (1976). *Depth perception through motion*. New York: Academic Press.
- Braunstein, M. L. (1988). The empirical study of structure from motion. In W. N. Martin & J. K. Aggarwal (Eds.), *Motion understanding: Robot and human vision* (pp. 101–142). Higham, MA: Kluwer.
- Braunstein, M. L., & Andersen, G. (1981). Velocity gradients and relative depth perception. *Perception & Psychophysics*, 29, 145–155.
- Braunstein, M. L., Hoffman, D. D., & Pollick, F. E. (1990). Discriminating rigid from nonrigid motion: Minimum points and views. *Perception & Psychophysics*, 47, 205–214.
- Braunstein, M. L., & Tittle, J. S. (1988). The observer-relative velocity field as the basis for effective motion parallax. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 582–590.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Hillsdale, NJ: Erlbaum.
- Graham, C. H., Baker, K. E., Hecht, M., & Lloyd, V. V. (1948). Factors influencing thresholds for monocular movement parallax. *Journal of Experimental Psychology*, 38, 205–223.
- Hildreth, E. C. (1984). Computations underlying the measurement of visual motion. *Artificial Intelligence*, 23, 309–354.
- Koenderink, J. J. (1986). Optic flow. *Vision Research*, 26, 161–180.
- Loomis, J. M., & Eby, D. W. (1988). Perceiving structure from motion: Failure of shape constancy. In *Proceedings of the Second International Conference on Computer Vision* (pp. 383–391). Washington, DC: IEEE.
- Longuet-Higgins, H. C., & Prazdny, K. (1980). The interpretation of a moving retinal image. *Proceedings of the Royal Society of London, B*, 208, 385–387.
- Musatti, C. L. (1924). Sui fenomeni stereocinetici [On the stereokinetic phenomenon]. *Archivio Italiano di Psicologia*, 3, 105–120.
- Musatti, C. L. (1928). Sulla percezione di forma di figure oblique rispetto al piano frontale [Perception of shape of figures tilted in relation to the frontoparallel plane]. *Rivista di Psicologia*, 25, 1–14.
- Ono, H., & Comerford, J. (1977). Stereoscopic depth constancy. In W. Epstein (Ed.), *Stability and constancy in visual perception mechanisms and processes* (pp. 91–128). New York: Wiley.
- Ono, H., Rogers, B. J., Ohmi, M., & Ono, M. E. (1988). Dynamic occlusion and motion parallax in depth perception. *Perception*, 17, 255–266.
- Ono, H., & Steinbach, M. J. (1990). Monocular stereopsis with and without head movement. *Perception & Psychophysics*, 48, 179–187.
- Ono, M. E., Rivest, J., & Ono, H. (1986). Depth perception as a function of motion parallax and absolute distance information. *Journal of Experimental Psychology: Human Perception and Performance*, 12, 331–337.
- Proffitt, D. R., Rock, I., Hecht, H., & Schubert, J. (1992). The stereokinetic effect and its relations to the kinetic depth effect. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 3–21.
- Rogers, B., & Graham, M. (1979). Motion parallax as an independent cue for depth perception. *Perception*, 8, 125–134.
- Todd, J. T. (1982). Visual information about rigid and non-rigid motion: A geometrical analysis. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 238–252.
- Ullman, S. (1979). *The interpretation of visual motion*. Cambridge, MA: MIT Press.
- Wallach, H., Weisz, A., & Adams, P. (1956). Circles and derived figures in rotation. *American Journal of Psychology*, 69, 48–59.
- Zanforlin, M. (1988). The height of a stereokinetic cone: A quantitative determination of a 3-D effect from a 2-D moving patterns without a "rigidity assumption." *Psychological Research*, 50, 162–172.

Received April 23, 1991

Revision received February 18, 1992

Accepted February 25, 1992 ■