
Perceptual organization of apparent motion in the Ternus display

Zijiang J He

Department of Psychology, University of Louisville, Louisville, KY 40292, USA;
e-mail: z0he0002@athena.louisville.edu

Teng Leng Ooi

Department of Biomedical Sciences, Southern College of Optometry, 1245 Madison Avenue,
Memphis, TN 38104, USA; e-mail: tlooi@sco.edu

Received 26 August 1997, in revised form 8 March 1999

Abstract. A typical Ternus display has three sequentially presented frames, in which frame 1 consists of three motion tokens, frame 2 (blank) defines the interstimulus interval, and frame 3 has similar motion tokens with their relative positions shifted to the right. Interestingly, what appears to be a seemingly simple arrangement of stimuli can induce one of two distinct apparent-motion percepts in the observer. The first is an *element-motion* perception where the left-end token is seen to jump over its two neighboring tokens (inner tokens) to the right end of the display. The second is a *group-motion* perception where the entire display of the three tokens is seen to move to the right. How does the visual system choose between these two apparent-motion perceptions? It is hypothesized that the choice of motion perception is determined in part by the perceptual organization of the motion tokens. Specifically, a group-motion perception is experienced when a strong grouping tendency exists among the motion tokens belonging to the same frame. Conversely, an element-motion perception is experienced when a strong grouping tendency exists between the inner motion tokens in frames 1 and 3 (ie the two tokens that overlap in space between frames). We tested this hypothesis by varying the perceptual organization of the motion tokens. Both spatial (form similarity, 3-D proximity, common surface/common region, and occlusion) and temporal (motion priming) factors of perceptual organization were tested. We found that the apparent-motion percept of the Ternus display can be predictably affected, in a manner consistent with the perceptual organization hypothesis.

1 Introduction

When two stationary stimuli are presented sequentially at two locations, they are often seen as a single object moving from the first location to the second. This perception of apparent movement is of particular interest to vision researchers since the movement impression does not reflect the frank physical stimulation that impinged on the retina, but the consequence of the visual system's 'act' upon the external stimulus.

Of the various apparent motion paradigms, perhaps the one associated with the Ternus display is the most fascinating (Pantle and Picciano 1976; Ternus 1926), and is one from which we have learned a great deal regarding the underlying mechanism of apparent motion (Braddick 1980; Breitmeyer and Ritter 1986a, 1986b; Grossberg 1991; Grossberg and Rudd 1992; Pantle and Petersik 1980; Pantle and Picciano 1976; Petersik and Pantle 1979). A typical Ternus display consists of three frames that are presented sequentially, as illustrated in figure 1a. Frame 1 has three motion tokens (circles) and is displayed for a short duration. This is followed by frame 2 (blank) that consists of a variable duration or interstimulus interval (ISI). Then the third frame, which has similar motion tokens as the first frame but with their relative positions shifted to the right, is presented. Together, this sequential presentation results in the observer experiencing one of two possible types of apparent-movement impression (figure 1b). The first type is that of three tokens moving together as a group to the right (group motion); the second type is that of the first token on the left jumping over its neighbors to locate on the right, leaving its neighbors (the two tokens that overlap in space between frames 1 and 3,

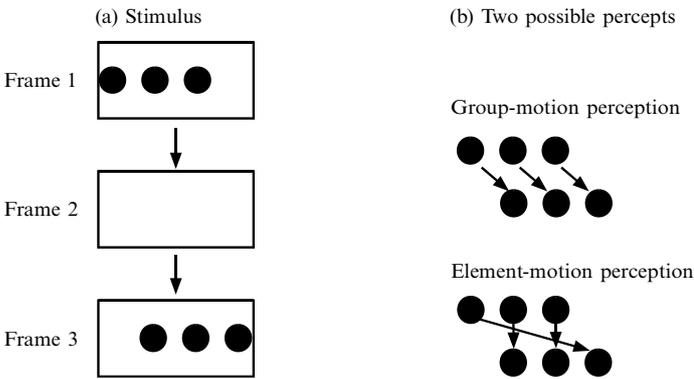


Figure 1. Illustration of the Ternus motion display. (a) The apparent-motion stimulus consists of three frames presented sequentially. Frame 1 consisting of three tokens is presented first, followed by frame 2 (blank), and frame 3. Frame 3 has three tokens—similar to frame 1—whose positions are shifted to the right relative to those in frame 1. (b) The two possible types of apparent-motion perception in the Ternus display. In ‘group-motion perception’, the three tokens are seen as a single unit moving to the right. In ‘element-motion perception’, the leftmost token is perceived to jump to the rightmost location while the second and third tokens (ie the inner tokens) flicker without moving.

which for simplicity will be referred to as the ‘inner tokens’) flickering but stationary (element motion).

It has been proposed that the motion perception in the Ternus display reflects the interaction between the lower-level (short range motion) and higher-level motion (long-range motion) processes (Braddick 1980; Pantle and Picciano 1976). The lower-level motion process detects whether there is a positional change in the two inner tokens (local motion) in going from frame 1 to frame 3. If no change/motion signal is detected in the two inner tokens, the higher-level motion process will assume that the inner tokens are flickering, but stationary objects. Thus the motion signal originates solely from the left end token in frame 1 that ‘jumps’ over to the right end in frame 3, producing the element-motion perception. On the other hand, when the lower-level motion process detects a local positional change in the inner tokens, the higher-level motion process will assume that the three tokens move as a whole (group) from the left in frame 1 to the right in frame 3, producing the group-motion perception (Braddick 1980; Pantle and Petersik 1980). This explanation has been supported by several empirical findings (Braddick 1980; Pantle and Picciano 1976; Pantle and Petersik 1980; Petersik and Pantle 1979). For example, it has been shown that when the inner tokens in frame 3 are shifted leftward or rightward with respect to the inner tokens in frame 1 by a magnitude larger than 16 min of arc, the observer is more likely to perceive group motion (Petersik and Pantle 1979). Pantle and Picciano (1976) also showed that the motion perception depends on the ISI: the shorter the ISI, the more element motion is seen. According to Breitmeyer and Ritter (1986a, 1986b), this is due to the fact that when the ISI is very short, pattern persistence prevents the visual system from producing a strong signal to indicate a positional change in the inner tokens.

Another way to understand how the visual system determines the motion perception of the Ternus display is to account for the perceptual organization factors governing its motion tokens (Ternus 1926). This is illustrated in figure 2, where the motion tokens in the Ternus display can group in two different ways. The first way is for the three motion tokens in each frame to group together as a whole (ie within frame grouping), to yield a group-motion perception. The second way is for the inner tokens in frame 1 (the two rightmost tokens) to group with the inner tokens in frame 3 (the two leftmost tokens) at the same location (ie across frame grouping). Thus, this leaves the outer

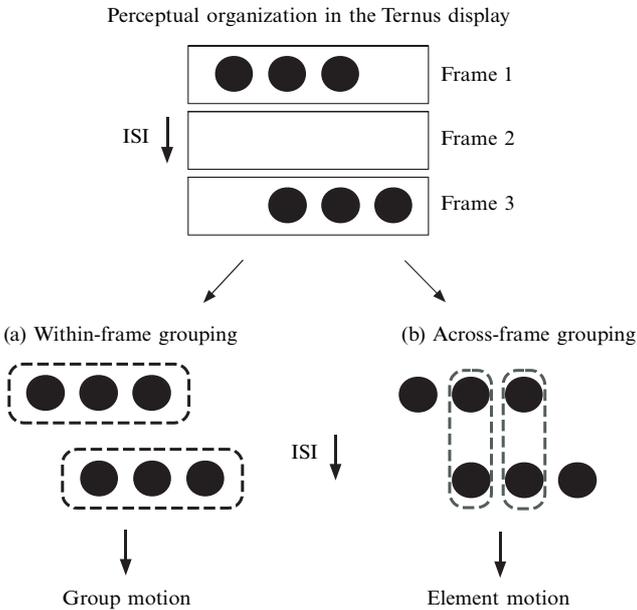


Figure 2. Schematic illustration of the perceptual organization hypothesis in the Ternus display. (a) Within-frame grouping. The grouping affinity among the three motion tokens in the same frame is strong when the visual system corresponds the three tokens in frame 1 as a ‘whole’ to the three tokens in frame 3 as a ‘whole’. This produces the group-motion perception. (b) Across-frame grouping. When the grouping affinity between the inner motion tokens of frame 1 (ie the two rightmost tokens) and the inner motion tokens of frame 3 (ie the two leftmost tokens) at corresponding locations is strong, these tokens are seen as stationary flickering objects. Meanwhile, the outer token at one end in frame 1 corresponds to the outer token at the other end in frame 3, to produce an apparent-movement perception of a token from one end to the other end (element motion).

token at the left end in frame 1 to correspond to the outer token at the right end in frame 3, leading to the apparent-movement percept of the token moving from the left to right (element motion). Invariably, the relative grouping strength between these two grouping tendencies determines the final motion perception.

Our objective here has been to examine the application of the perceptual organization hypothesis in the Ternus display. We manipulated various spatial and temporal perceptual grouping factors governing the motion tokens to explore how its motion perception is affected. Overall, our results reveal that the motion perception in the Ternus display can be reliably predicted by the perceptual organization of its motion tokens.

2 General method

2.1 Apparatus and stimuli

The Ternus display, appropriately modified to suit the purpose of each experimental condition, was presented on a computer monitor driven by a Commodore computer (A2000 or A3000). The motion stimuli used in experiments 1 and 6 were viewed directly by the observers, while those in experiments 2–5 were viewed through a pair of haploscopic prisms to create a stereoscopic Ternus motion display. The viewing distance was 100 cm.

The temporal sequence and duration of the apparent-motion stimuli employed in all the experiments were the same. Frames 1 and 3 were 200 ms each. The duration of frame 2 (blank) which defines the ISI was varied from 0 to 117 ms. The presentation sequence was frame 1–frame 2–frame 3–frame 2, and was repeated three times during a trial.

2.2 Observers

The experiments were conducted at both the University of Louisville and the Southern College of Optometry. All naive observers, whose ages ranged from 24–45 years, had normal or corrected-to-normal visual acuity and at least 40 s of arc of stereoacuity. Informed consent was obtained from each observer before commencing the experiments. The naive observers were given 300–500 practice trials to familiarize them with the Ternus display phenomenon before starting the proper data collection. The stimulus used for the practice trials consisted of a typical Ternus display on a homogenous background. The spatial dimensions of the motion tokens were the same as those in figure 11a. The frame duration was 200 ms and the ISI was varied from 0 to 117 ms.

3 Experiment 1. Form similarity

If the hypothesis that perceptual organization (within-frame versus across-frame grouping) determines the motion perception of the Ternus display is correct, we should expect that when the grouping affinity between the inner and outer tokens within a frame is reduced, the observer is less likely to perceive group motion. To provide support for this prediction, in our first experiment we varied the grouping affinity among the tokens by manipulating their form similarity with respect to the surround.

Figure 3 illustrates the two motion displays used. In condition (a) (anti-grouping), the Ternus display (the three middle horizontal circles in frames 1 and 3) is sandwiched by four pairs of white circular elements, two pairs stacked on top of one another above the Ternus display and two other pairs below. This arrangement leads to an almost equal tendency for the two inner motion tokens of the Ternus display to group either with the outer motion token, or with the additional white circular elements surrounding them to form a vertical rectangular array. To reduce the tendency for the Ternus display to group with the surrounding circular elements, condition (b) (grouping) has the contrast polarity of the surrounding circular elements reversed by making them black. In this way, the predominant grouping tendency would be among the three motion tokens. Accordingly, more group motion is expected to be seen in condition (b) (grouping) than in condition (a) (anti-grouping).

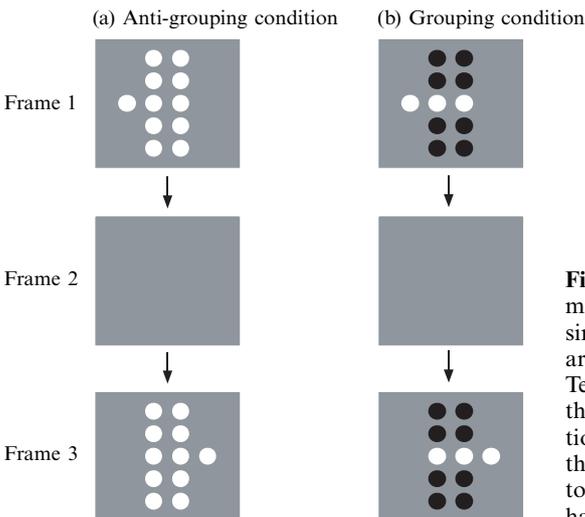


Figure 3. Experiment 1: The Ternus display modified to investigate the effect of form similarity. Two pairs of circular elements are placed above the motion tokens of the Ternus display and two other pairs below the display. (a) In the anti-grouping condition, the four pairs of circular elements have the same contrast polarity as the motion tokens. (b) In the grouping condition, they have opposite contrast polarity.

3.1 Stimulus and procedure

The diameter of the circular motion tokens and the surrounding circular elements was 34 min of arc. The luminances of the white and black elements were 130 cd m^{-2} and 0.9 cd m^{-2} , respectively. The elements were viewed upon a gray background of 34 cd m^{-2} . The horizontal inter-token distance (center to center) was 68 min of arc, while the vertical distance between the elements was 70 min of arc. The temporal parameters of the motion display have been described in section 2.

Before each trial, the observer was instructed to fixate the center of a gray computer screen. Thereafter, the observer self-initiated the trial by pressing the mouse button to present the motion tokens. The computer was programmed to randomly select the ISI for each trial. During the experiment, the observer would experience either a group-motion perception or an element-motion perception, to which he would respond to by pressing one of two keys on the computer keyboard according to his perception. Seven ISI durations, with 15 trials for each ISI, were tested on each observer. From these, the percentage of seeing group motion was obtained.

3.2 Results

Three naive observers (two males and one female) and one author were tested in both conditions (figures 3a and 3b) and their average data are shown in figure 4. Confirming our prediction, the observers saw more group motion in the grouping condition where the surrounding circular elements had opposite contrast polarity to the motion tokens (circles, figure 4), than in the anti-grouping condition where the surrounding circular elements had the same contrast polarity (squares, figure 4) (two-way within-subject ANOVA, $F_{1,3} = 13.19$, $p < 0.05$).

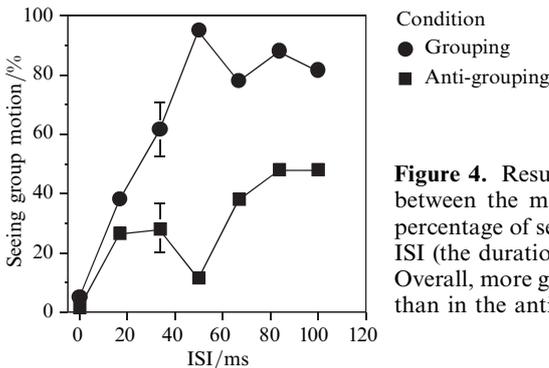


Figure 4. Results of experiment 1, where form similarity between the motion tokens was manipulated. The average percentage of seeing group motion is plotted as a function of ISI (the duration of frame 2) for the two conditions ($n = 4$). Overall, more group motion is seen in the grouping condition than in the anti-grouping condition.

4 Experiment 2. 3-D proximity

We now examined how a second grouping factor, proximity, affects the motion perception of the Ternus display, by employing a stereoscopic stimulus presentation (figure 5a—depth-separation condition). By free fusing the two left stereo images divergently or the two right stereo images convergently (figure 5a), one will perceive the two outer tokens as located in a depth plane behind the two inner tokens. This perceptual depth relationship is schematically displayed below the stereogram in figure 5a. Clearly, the perceived distance between the inner and outer tokens in this depth-separation display (figure 5a) is larger than that in the control Ternus display (figure 5b—iso-depth condition), where the tokens have the same depth. Thus, if the law of proximity applies to the Ternus display, a relatively weaker grouping tendency between the inner and outer tokens (within-frame grouping) in the depth-separation condition will result in the observer seeing less group motion. (Note: in the proper experiment, only three tokens were presented in each frame; what is shown as four tokens in the stereogram is there to depict all the possible spatial locations of the tokens in going from frame 1 to frame 3.)

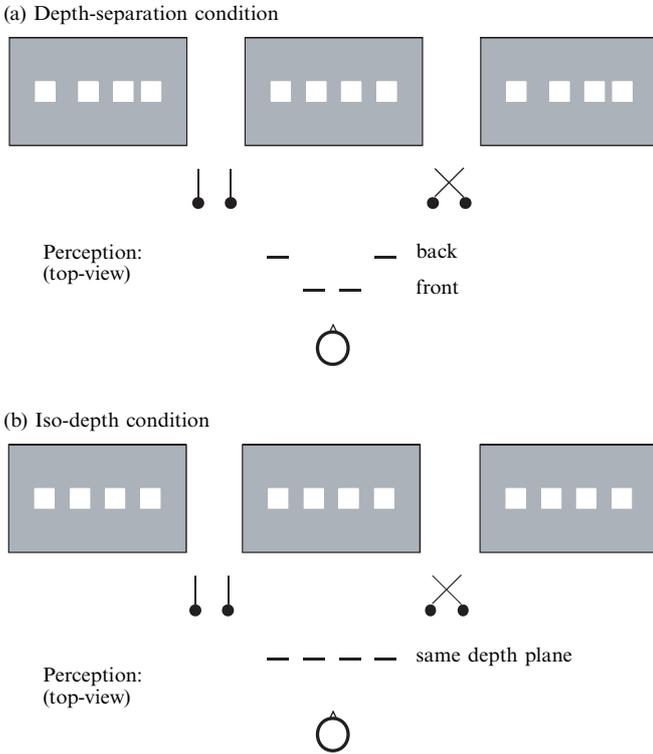


Figure 5. Experiment 2: The Ternus display modified to investigate the effect of 3-D proximity. Cross fusers should free fuse the middle and right stereo pair, while uncrossed fusers should free fuse the left and middle stereo pair. (a) Depth-separation condition. When fused, the two middle white motion tokens are seen in front of the two outer ones, as shown by the top-view cartoon. This arrangement reduces within-frame grouping tendency. (b) Iso-depth condition. When fused all the white motion tokens are seen in the same depth plane, increasing within-frame grouping tendency.

4.1 Stimulus and procedure

The white square motion tokens in figure 5 had a luminance of 148 cd m^{-2} and were presented against a gray background of 90 cd m^{-2} . Each token had a dimension of $33 \text{ min of arc} \times 33 \text{ min of arc}$. The horizontal inter-token distance (center to center) was about 66 min of arc for the two inner tokens in the depth-separation condition, and all the tokens in the iso-depth condition. The binocular disparity between the front and back tokens in the depth-separation condition was about 14 min of arc . The testing procedure was the same as in experiment 1.

4.2 Results and discussion

Figure 6 graphs the average results of five observers (three male and one female naive observers and one author). Overall, less group motion was reported in the depth-separation condition (squares) than in the iso-depth condition (circles) (two-way within-subjects ANOVA, $F_{1,4} = 12.627$, $p < 0.05$).

Interestingly, we found that although the pooled responses ($n = 5$) in figure 6 showed a statistically significant difference between the two conditions, there was also some individual differences among the five observers. While they all showed largely similar responses in the iso-depth condition, this was not so for the depth-separation condition. Unlike other observers who reported seeing over 85% more element motion for all ISIs tested in the depth-separation condition than in the iso-depth condition, one naive observer reported only a slight increase in element-motion perception.

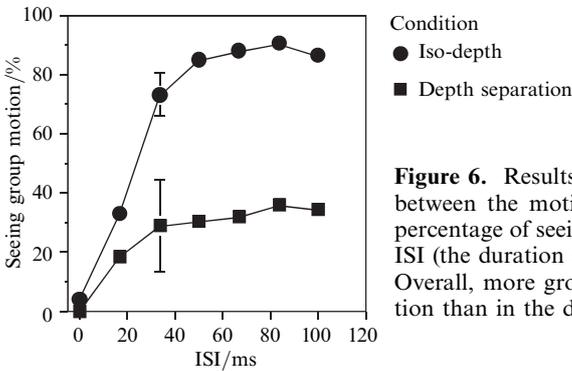


Figure 6. Results of experiment 2 where 3-D proximity between the motion tokens was manipulated. The average percentage of seeing group motion is plotted as a function of ISI (the duration of frame 2) for the two conditions ($n = 5$). Overall, more group motion is seen in the iso-depth condition than in the depth-separation condition.

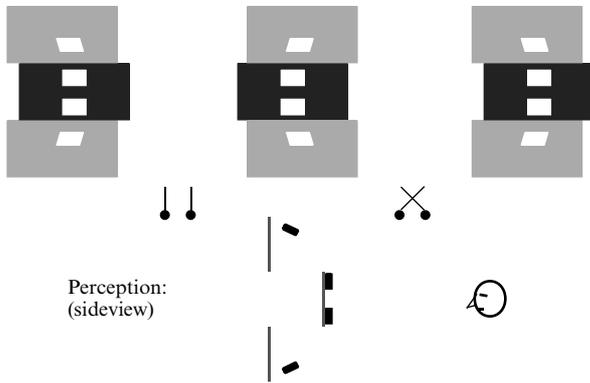
At this moment, we have not yet fully understood the basis of this individual difference. But one likely possibility that we are currently exploring is that this exceptional observer might require a longer temporal duration to perceive stereo depth. If this was the case, the ISIs and frame duration employed in the experiment above were probably too short for this observer to reliably experience that the depth separation between the inner and outer tokens was larger than their horizontal (center to center) distance in the depth-separation condition.

5 Experiment 3. Common surface

Experiment 2 revealed the effect of stereo depth (3-D proximity) on the motion perception of the Ternus display. This finding is consistent with previous reports of other types of 3-D apparent motion displays (eg Attneave and Block 1973; Green and Odom 1986). For example, Green and Odom demonstrated that in a 3-D competitive apparent-motion paradigm, observers had a bias to seeing less apparent motion between tokens that were located at different depths, than on a common depth plane. However, He and Nakayama (1994a) argued that it is the perceived surface separation, rather than the stereo depth difference between motion tokens per se, that causes observers to perceive less apparent motion across depth. To prove it, they compared two depth conditions with different surface attributes. In their surface-separation condition, two pairs of motion tokens were located in different depth planes and seen against a vertical background surface. Then in their common-surface condition, the background surface was slanted in depth so that it now supported the same two pairs of motion tokens. He and Nakayama found that their observers saw more apparent motion across depth in the common-surface condition, indicating that perceptual grouping affinity between the tokens is facilitated on a common background surface. This finding is also consistent with the proposal that grouping by common region is one of the tenets of perceptual organization (Palmer 1992).

The goal of experiment 3 was to investigate whether perceptual grouping affinity is increased for Ternus display stimuli on a common surface. Two different surface conditions were compared in the experiment, as shown in figure 7. By free fusing the top left and middle stereo images divergently, or the top middle and right ones convergently (figure 7a—surface-separation condition), one will see the two inner white motion tokens as located on a rectangular plane whose relative depth is in front of the depth location of the outer tokens at each end. Meanwhile in the bottom stereogram (figure 7b—common-surface condition), these same motion tokens are seen as located on a common convex background surface. We predict that this common background surface will induce a stronger perceptual grouping affinity among the motion tokens in the same frame (within-frame grouping), resulting in more group-motion perception.

(a) Surface-separation condition



(b) Common-surface condition

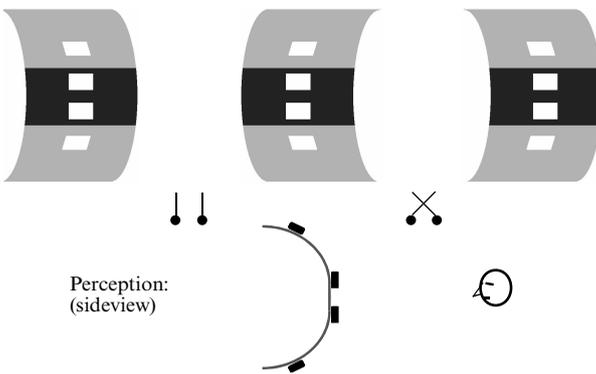


Figure 7. Experiment 3: The Ternus display modified to investigate the effect of common surface. Cross fusers should free fuse the middle and right stereo pair, while uncrossed fusers should free fuse the left and middle stereo pair. (a) Surface-separation condition. The two middle white squares (inner motion tokens) are located on a black rectangular plane in front of the outer white tokens at each end. (b) Common-surface condition. The depth relation between the motion tokens is the same as that in condition (a). However, now the tokens are all supported on a common curved, convex surface.

5.1 Stimulus and procedure

The physical properties of the motion tokens were similar in the two conditions (figure 7). In each condition, the two inner tokens (28.8 min of arc \times 16.8 min of arc) were oriented frontoparallel to the observer while the outer tokens (28.8 min of arc \times 11.95 min of arc) were slanted (7.5 min of arc of disparity). The inner tokens were perceived as closer to the observer than the two outer tokens (12.5 min of arc of disparity). The vertical separation between tokens was about 23.9 min of arc.

In the surface-separation condition (figure 7a), the tokens were seen against three frontoparallel background surfaces (160 min of arc \times 71.7 min of arc) that were filled with a small random square pattern (not shown). The middle background surface was seen as closer than the upper and lower ones (47.8 min of arc of disparity). In the common-surface condition (figure 7b), all the motion tokens were seen against a curved, convex surface that was also filled with a small random square pattern (not shown).

5.2 Results

The average results of four observers (two male naive observers and two authors) are shown in figure 8. Clearly, more group motion was seen in the common-surface condition (circles) than in the surface-separation condition (squares) ($F_{1,3} = 16.05$, $p < 0.05$). This finding confirms the prediction that motion tokens on a common surface will have a stronger within-frame grouping affinity, resulting in a perceptual bias towards group-motion perception. This finding also agrees with previous observations regarding the role of the common surface in apparent-motion perception (He and Nakayama 1994a).

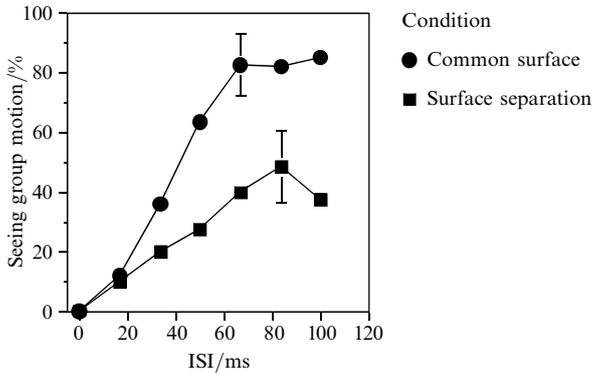


Figure 8. Results of experiment 3, where surface separation was manipulated. The average percentage of seeing group motion is plotted as a function of ISI (the duration of frame 2) for the two conditions ($n = 4$). Overall, more group motion is seen in the common-surface condition than in the surface-separation condition, indicating that when motion tokens are located on a continuous surface, within-frame grouping tendency is increased.

6 Experiment 4. Occlusion (amodal surface completion)

Various studies have shown that our visual system can spatially integrate visible segments of a partially occluded object (amodal surface completion, Kanizsa 1979; Michotte et al 1967). In fact, amodal surface completion has been found to play an important role in motion perception (He and Nakayama 1994b; Ramachandran and Anstis 1986; Shimojo and Nakayama 1990; Watamaniuk and McKee 1995; Yantis 1995). Here, we investigated if such spatial-integration ability can operate on motion tokens in the Ternus display to produce a group motion perception.

Our question and prediction above are reasonable, if one examines a likely real motion situation which resembles the Ternus display (figure 9). A black horizontal bar moving rightward is partially occluded by three white vertical rectangles. Owing to the occlusion by the white rectangles, the motion signal of the black horizontal bar is carried only by its two outer segments. Thus, for the observer to obtain a veridical perception of this real-world situation, ie a perception of a continuous black horizontal bar moving behind the white occluding rectangles, the visual system has to spatially integrate the motion signals from all visible segments of the black horizontal bar.

Motion behind occluders

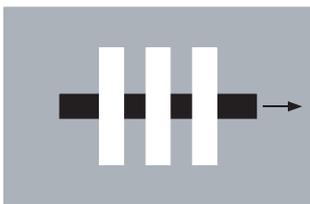


Figure 9. A black horizontal bar moves rightward behind three occluding vertical rectangles. For veridical perception of the moving bar, the motion signals carried by the two ends of the bar have to spatially integrate over the entire bar.

To show that amodal surface completion can affect motion perception of the Ternus display, we modified the Ternus display (occlusion condition, figure 10) such that it consisted of four horizontally aligned white squares (motion tokens) in juxtaposition to three vertical black rectangles. Because of the tendency of the visual system to perceptually complete the white square tokens, these tokens are often perceived as a white horizontal rectangle partially occluded by three vertical black rectangles (random square pattern not shown). Predictably, more group motion will be perceived since the tokens are perceptually grouped within each frame.

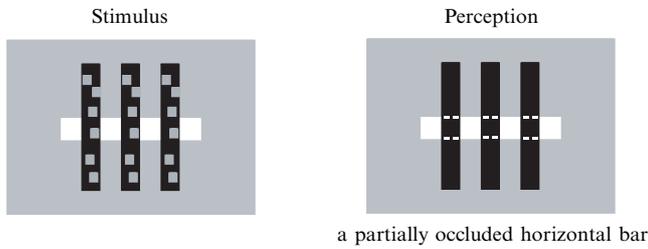


Figure 10. Experiment 4: The stimulus configuration of the test display. In this occlusion condition, three black vertical rectangular bars are inserted between the gaps separating the four motion tokens (white squares) of the Ternus display. The T-junctions formed between the white square tokens and black rectangular bars cause the white tokens to be seen as a continuous white horizontal bar that is partially occluded by the vertical black rectangular bars.

Two major factors have been suggested for causing amodal surface completion (Nakayama et al 1989). These are the presence of T-junctions and appropriate depth relationship. Thus, to prevent amodal surface completion, our first control experiment eliminated the presence of T-junctions by altering the shape of the motion tokens from squares to circles (figure 11a: eliminating T-junction). Note that in so doing, the circular edges of the motion tokens can no longer form T-junctions with their juxtaposed vertical black rectangles (Shimojo and Nakayama 1990). This reduces the grouping affinity between the tokens (figure 11a) and predictably will result in less group-motion perception, compared to the test condition (figure 10) where amodal surface completion occurs.

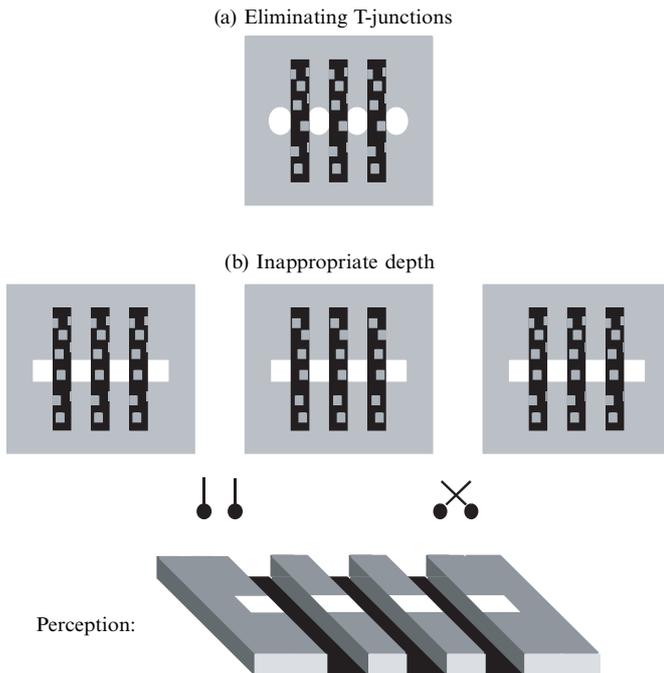


Figure 11. Illustration of the two control conditions used in experiment 4 to reduce amodal surface completion. (a) Eliminating T-junctions. By changing the shape of the motion tokens to circles, amodal surface completion (occlusion) is prevented since the motion tokens can no longer form T-junctions with the black vertical rectangles. (b) Inappropriate depth. By free fusing the left and middle stereo pair divergently, or the middle and right pair convergently, a gray background surface upon which four motion tokens (white squares) are located is perceived. However, the continuity of this surface is disrupted by gaps (black regions) that recede in depth inbetween the spaces separating the motion tokens, thus abolishing amodal surface completion between the motion tokens.

The second control experiment discouraged amodal surface completion by relying on the fact that amodal surface completion occurs only in the presence of appropriate stereoscopic depth information in the display (Nakayama et al 1989). Thus we designed a stereogram (figure 11b: inappropriate depth) in which the vertical black rectangles have uncrossed disparity (by virtue of the small red square pattern embedded within the black rectangles). By free fusing the left and middle stereo images divergently or the middle and right ones convergently, one will perceive the vertical black rectangles as recessed gaps separating the gray surface which supports the white square motion tokens. Notably, the presence of the recessed gaps between the white square tokens that are seen on the front plane prevents amodal surface completion from occurring, resulting in each white square token being perceived individually, rather than as part of an occluded horizontal rectangle. Consequently, less group motion will be observed in this control condition (figure 11b), compared to the test condition (figure 10).

6.1 Stimulus and procedure

The white square motion tokens in figures 10 and 11b were 33 min of arc \times 32 min of arc, while the white circular motion tokens in figure 11a were 30 min of arc in diameter. Each black region within which the small random red square pattern was embedded was 33 min of arc \times 183 min of arc. The horizontal separation between each black region was 33 min of arc. All images had zero disparity, except that in figure 11b, where the small random red square pattern that was embedded within the black vertical regions was given an uncrossed disparity of 11 min of arc. The testing procedure was the same as in experiment 1.

6.2 Results and discussion

The average data from four observers (two male and one female naive observers and one author) for the occlusion condition (figure 10), and two control conditions (figures 11a and 11b) are shown in figure 12. Consistent with prediction, more group motion was seen in the occlusion condition (open squares) where the motion tokens were seen as an occluded white horizontal rectangle. A two-way repeated-measures ANOVA reveals a significant main effect of the occlusion condition ($F_{2,6} = 8.3, p < 0.025$). Furthermore, Tukey's Studentized Range (HSD) test indicates a significant difference between the occlusion condition and the control condition with circular motion tokens ($p < 0.005$), and a significant difference between the two control conditions ($p < 0.005$).

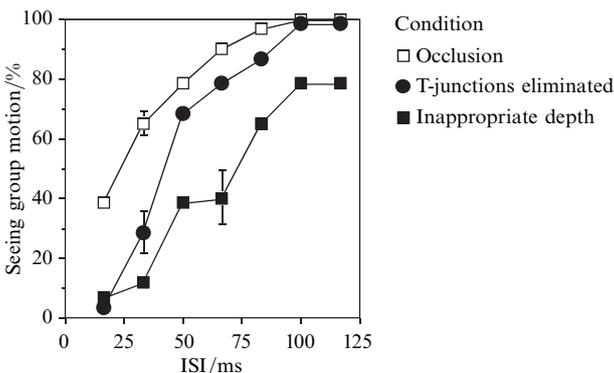


Figure 12. Results of experiment 4 where amodal surface completion cues were manipulated. The average percentage of seeing group motion is plotted as a function of ISI for the occlusion condition and two control conditions ($n = 4$). Overall, more group motion is seen in the occlusion condition than in the two control conditions, where amodal surface completion between the motion tokens was prevented. Curiously, eliminating amodal completion by manipulating the depth relationship in the control condition, rather than T-junctions, appears to be more effective in reducing group-motion perception.

Also consistent with the quantitative results above, our observers reported that they occasionally perceived the tokens as a single, occluded white horizontal rectangle moving left and right repeatedly behind three black rectangles in the occlusion condition, instead of simply as three individually moving tokens. Meanwhile in the control conditions the three motion tokens were always seen as separate entities above the background, regardless of whether they were moving in the element-motion or group-motion mode.

7 Experiment 5. What causes the difference between the two control conditions in experiment 4?

What factors could possibly contribute to the slightly different motion tendencies between the two control conditions in the previous experiment (figures 11a and 11b)? A first step to studying this question is to make the shape of the tokens the same in the two control conditions. This measure counters the argument that perhaps more element motion was experienced in the inappropriate-depth condition (figure 11b) because the square tokens were interpreted by the visual system as belonging to the vertical gray rectangles that supported them, as their vertical edges were perfectly aligned with the gray rectangles. Thus, in experiment 5, the square tokens of figure 11b were changed to circles.

7.1 Stimulus and procedures

The first condition for experiment 5 (figure 13a) was essentially the same as that used in experiment 4 (figure 11a). The second condition (figure 13b) was a modification of the inappropriate-depth condition in experiment 4 (figure 11b), in that the shape of the motion tokens was changed from squares to circles. All other aspects of the experiment remained the same.

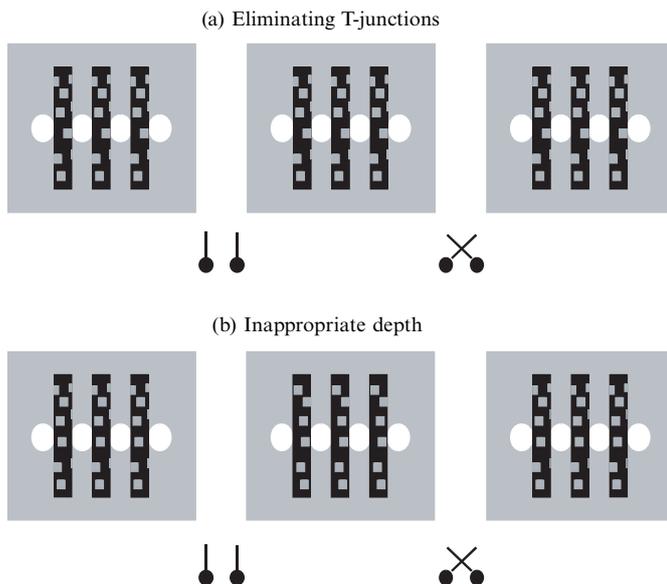


Figure 13. Experiment 5: Further investigation of the control conditions employed in experiment 4. Cross fusers should free fuse the middle and right stereo pair, while uncrossed fusers should free fuse the left and middle stereo pair. (a) Eliminating T-junctions: this condition is essentially the same as that in figure 11a. (b) Inappropriate depth: this condition is essentially the same as that in figure 11b, except that the shape of the motion tokens is altered from squares to circles to eliminate T-junctions.

7.2 Results and discussion

Four naive observers (two males and two females) who were not part of experiment 4, and one author who was, participated in experiment 5. The average results are shown

in figure 14. Although the percentage of seeing group motion is noticeably smaller in the second condition with inappropriate depth (squares), a two-way within-subjects ANOVA fails to reveal a significant difference in the perceived motion between the two conditions tested ($F_{1,4} = 4.269$, $p = 0.108$). Thus, this finding, that if the shape of the motion tokens in the two conditions is similar, this can reduce the difference in the apparent motion perception, helps explain why an opposite result was found in the two control conditions in experiment 4.

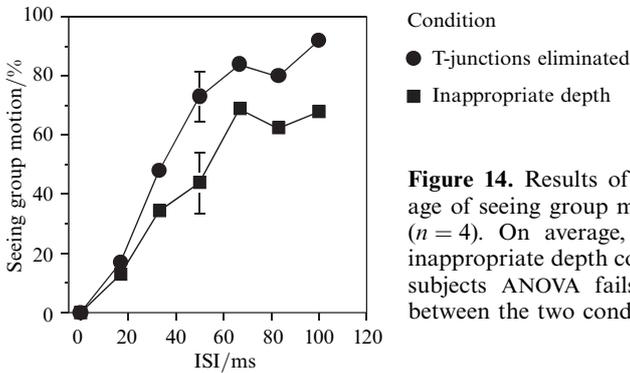


Figure 14. Results of experiment 5. The average percentage of seeing group motion is plotted as a function of ISI ($n = 4$). On average, less group motion is seen in the inappropriate depth condition. However, a two-way within-subjects ANOVA fails to reveal a significant difference between the two conditions ($F_{1,4} = 4.269$, $p = 0.108$).

Apart from the shape factor, another likely explanation for the inappropriate-depth condition in experiment 4 yielding more element-motion perception, can be related to the lack of a common background surface. Specifically, the tokens in the inappropriate-depth condition were located on a discontinuous surface—and as we have learned from our findings in experiments 2 and 3, a discontinuous background surface facilitates element-motion perception.

8 Experiment 6. Influencing perceptual grouping with priming

Visual perception is largely a product of bottom-up and top-down influences, in which memory plays a substantial role (Gregory 1970). Motion perception has been found to be experience-dependent (eg McKee and Welch 1985; Pinkus and Pantle 1997; Ramachandran and Anstis 1983). For example, Ramachandran and Anstis (1983) reported that in a typical 2×2 apparent motion display (von Schiller 1933), the instantaneously perceived motion direction is biased towards the perceived motion direction from a previous exposure.

Here we investigated how prior experience (priming) can affect perceptual grouping of motion tokens in the Ternus display. At issue is whether the tendency to see group or element motion can be modified when, prior to the Ternus display stimulation, a pair of flickering stimuli are briefly shown to the observer. Thus, in the priming condition (figure 15), two priming probes (white circular elements) are turned on and off twice (priming frames) before the standard Ternus display is presented. The function of the priming probes is to leave a 'no motion' memory trace, since they are merely seen as flickering, but stationary, objects. By carefully placing the priming probes to coincide with the subsequent locations of the two inner motion tokens of the Ternus display, one will be able to bias the visual system to interpret the two subsequent inner motion tokens as the same flickering probes as the priming probes at those locations. In this way, there is less tendency for the visual system to interpret the two subsequent inner tokens as being part of a triple-element group with the outer token (in the same frame). In other words, the role of the priming frames is to increase the across-frame grouping tendency and reduce the within-frame grouping tendency. Consequently, the visual system will be biased to perceiving element motion. To test this prediction, we compared the priming condition to a control condition that consisted of the standard Ternus display (figure 15).

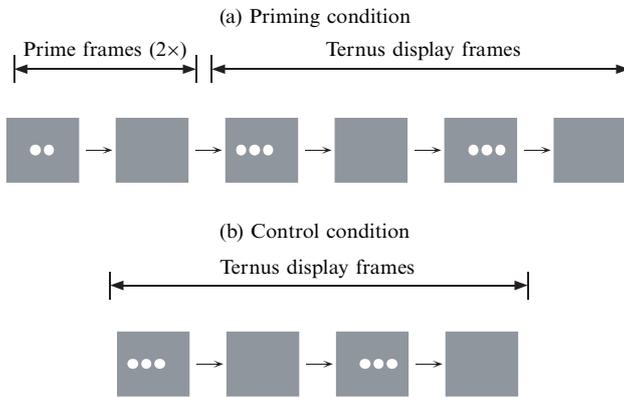


Figure 15. Experiment 6: The Ternus display modified to investigate the effect of priming. (a) Priming condition. This consists of a frame with two white circular elements and a second frame that is blank. These frames are presented twice before the Ternus display is shown to the observer. The priming elements are placed at locations corresponding to the locations of the two inner tokens of the subsequent Ternus display, and are perceived as flickering circles before the onset of the Ternus display. (b) Control condition. This consists of the typical Ternus display.

8.1 Stimulus and procedure

The motion displays used are illustrated in figures 15a and 15b. The stimulation duration for the priming frame and the Ternus display frame was 200 ms each. The diameter of the circular motion tokens was 34 min of arc. The luminance of these white motion tokens was 130 cd m^{-2} and they were placed against a 34 cd m^{-2} gray background.

During the experiment, the observer initiated a trial by pushing the mouse button. This led to the presentation of the Ternus display in the control condition (figure 15b). Meanwhile, in the priming condition (figure 15a), pushing the mouse button led to the display of the priming probes (twice) which was followed by the Ternus display. For both conditions, the observer's task was to report whether an element or a group motion was perceived.

8.2 Results and discussion

Figure 16 shows the average results of the observers (three male naive observers and one author). Clearly, less group motion was seen in the priming condition (inverted triangles) than in the control condition (circles) (two-way within-subjects ANOVA, $F_{2,6} = 10.129$, $p < 0.025$). This finding indicates that the 'no motion' memory trace of the priming probes successfully biased the perception of the subsequent Ternus display to an element-motion perception.

As mentioned in section 1, Braddick (1980) suggested that an element-motion percept is experienced when the lower-level motion process signals to the higher-level motion

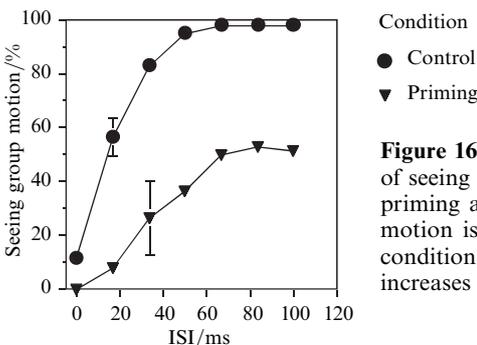


Figure 16. Results of experiment 6. The average percentage of seeing group motion is plotted as a function of ISI for the priming and control conditions ($n = 4$). Overall, more group motion is seen in the control condition than in the priming condition, suggesting that priming of the inner tokens increases across-frame grouping tendency.

process, determining the final motion percept, that the inner tokens have not moved. In this regard, it is plausible that our experiment reveals the memory-like characteristics of the motion system as the motion signal is mediated. However, further experiments are required to explore if the 'no motion' memory trace we just observed reflects the memory-like characteristics of the lower-level or higher-level motion process, or both.

9 General discussion

We tested here the hypothesis that the motion perception of the Ternus display (group versus element motion) depends on the perceptual organization of the motion tokens, both in space and in time. More specifically, when the grouping tendency among tokens in the same frame is strong, a group motion will be perceived. Conversely, when the grouping tendency between the inner tokens across frames is strong, an element motion will be perceived. This hypothesis is supported by our empirical findings, and provides a unique perspective to the observations made by previous investigators (eg Pantle and Petersik 1980; Petersik 1984). For example, Pantle and Petersik (1980) found that when inter-token distance decreases, the percentage of seeing group motion increases. Presumably, this is because reducing inter-token distance increases the grouping affinity among the tokens in the same frame.

Various researchers in the past have also studied how perceptual organization of motion elements in the first frame corresponds to the ones in the third frame (eg Kramer and Yantis 1997; Pantle and Petersik 1980; Ramachandran and Anstis 1983 1986). In fact, for a historical perspective, Ternus in his original paper wrote about 'phenomenological identity' of objects in apparent-motion displays. According to him, the visual appearance of an object can vary from moment to moment owing to its movement, illumination change, etc, even though the object's identity often remains unchanged. In the Ternus display, element-motion perception reflects a *retention of identity* of the inner tokens, while group-motion perception reflects an *exchange of identity* of the inner tokens (at same location). Following this line of thinking, 'across-frame grouping' factors or 'temporal grouping' effects (Kramer and Yantis 1997) act to increase the retention of identity of the inner tokens.

What then might be the ecological significance of across-frame grouping in the Ternus display? We speculate that it is perhaps related to an object's inertia in the physical world (Newton's first law), in which a stationary object remains at the same location until an extraneous force is applied to it. In fact, a similar phenomenon has also been noticed in a previous study by Ramachandran and Anstis (1986) in a modified von Schiller (1933) apparent-motion display. They reported that their observers had a bias to seeing an object in apparent motion remaining in the same motion course (direction) rather than changing its direction.

Admittedly, the perceptual-organization hypothesis focuses only on the consequence of grouping tendency among motion tokens, without specifying the mechanisms underlying the grouping mechanisms or how the various grouping factors are integrated (Ternus 1926). Several investigators in the past have speculated either explicitly or implicitly that the final choice of motion perception in the Ternus display is determined by a high-level motion mechanism (eg Braddick 1980; Grossberg and Rudd 1992; Pantle and Petersik 1980). If this is so, our current findings suggest that the final decision by the high-level motion mechanism includes accounting for the various grouping factors in the scene, as well as for the information it receives from the early-level motion mechanism.

Acknowledgements. This research was supported in part by a Sloan Research Fellowship from the Alfred P Sloan Foundation and a grant from the College of Arts and Sciences from the University of Louisville to ZJH, and a Faculty Research Grant from Southern College of Optometry to TLO. We thank the anonymous reviewers for their helpful comments and suggestions.

References

- Attneave F, Block G, 1973 "Apparent movement in tridimensional space" *Perception & Psychophysics* **13** 301–307
- Braddick O J, 1980 "Low-level and high-level processes in apparent motion" *Philosophical Transactions of the Royal Society of London, Series B* **290** 137–151
- Breitmeyer B G, Ritter A, 1986a "The role of visual pattern persistence in bistable stroboscopic motion" *Vision Research* **26** 1801–1806
- Breitmeyer B G, Ritter A, 1986b "Visual persistence and the effect of eccentric viewing, element size, and frame duration on bistable stroboscopic motion percepts" *Perception & Psychophysics* **39** 275–280
- Green M, Odom J V, 1986 "Correspondence matching in apparent motion: Evidence for three-dimensional spatial representation" *Science* **233** 1427–1429
- Gregory G L, 1970 *The Intelligent Eye* (New York: McGraw-Hill)
- Grossberg S, 1991 "Why do parallel cortical systems exist for the perception of static form and moving form?" *Perception & Psychophysics* **49** 117–141
- Grossberg S, Rudd M E, 1992 "Cortical dynamics of visual motion perception: Short-range and long-range apparent motion" *Psychological Review* **99** 78–121
- He J Z, Nakayama K, 1994a "Apparent motion determined by surface layout not by disparity or by 3-dimensional distance" *Nature (London)* **367** 173–175
- He J Z, Nakayama K, 1994b "Perceived surface shape determines correspondence strength in apparent motion" *Vision Research* **34** 2125–2136
- Kanizsa G, 1979 *Organization in Vision: Essays on Gestalt Perception* (New York: Praeger)
- Kramer P, Yantis S, 1997 "Perceptual grouping in space and time: Evidence from the Ternus display" *Perception & Psychophysics* **59** 87–99
- McKee S, Welch L, 1985 "Sequential recruitment in the discrimination of velocity" *Journal of the Optical Society of America A* **2** 243–251
- Michotte A, Thines G, Crabbe G, 1967 "Les compléments amodaux des structures perceptives" (Amodal completion of perceptual structures) *Studia Psychologica* (Louvain: Publications Universitaires de Louvain) [translated by T R Miles, E Miles, 1991, in *Michotte's Experimental Phenomenology of Perception* Eds G Thines, A Costall, G Butterworth (Hillsdale, NJ: Lawrence Erlbaum Associates) pp 140–167]
- Nakayama K, Shimojo S, Silverman G H, 1989 "Stereoscopic depth: Its relation to image segmentation, grouping, and the recognition of occluded objects" *Perception* **18** 55–68
- Palmer S, 1992 "Common region: A new principle of perceptual grouping" *Cognitive Psychology* **24** 436–447
- Pantle A J, Petersik J T, 1980 "Effects of spatial parameters on the perceptual organization of a bistable motion display" *Perception & Psychophysics* **27** 307–312
- Pantle A J, Picciano L, 1976 "A multistable movement display: Evidence for two separate motion systems in human vision" *Science* **193** 500–502
- Petersik J T, 1984 "The perceptual fate of letters in two kinds of apparent movement displays" *Perception & Psychophysics* **36** 146–150
- Petersik J T, Pantle A J, 1979 "Factors controlling the competing sensations produced by a bistable stroboscopic motion display" *Vision Research* **19** 143–154
- Pinkus A, Pantle A, 1997 "Probing visual motion signals with a priming paradigm" *Vision Research* **37** 541–552
- Ramachandran V S, Anstis S M, 1983 "Perceptual organization in moving pattern" *Nature (London)* **304** 529–531
- Ramachandran V S, Anstis S M, 1986 "Figure-ground segregation modulates apparent motion" *Vision Research* **26** 1969–1975
- Schiller P von, 1933 "Stroboskopische Alternativversuche" *Psychologische Forschung* **17** 179–214
- Shimojo S, Nakayama K, 1990 "Amodal representation of occluded surfaces: role of invisible stimuli in apparent motion correspondence" *Perception* **19** 285–299
- Ternus J, 1926 "The problem of phenomenal identity" English translation in *A Sourcebook of Gestalt Psychology* Ed. W D Ellis (London: Routledge & Kegan Paul)
- Watamaniuk S N J, McKee S P, 1995 "Seeing motion behind occluders" *Nature (London)* **377** 729–730
- Yantis S, 1995 "Perceived continuity of occluded visual objects" *Psychological Science* **6** 182–186