

Antiphase flicker induces depth segregation

AKIRA IWABUCHI

International Media Research Foundation, Tokyo, Japan

and

HIROSHI SHIMIZU

*Kanazawa Institute of Technology
and International Media Research Foundation, Tokyo, Japan*

We examined the influence of the temporal phase of flickering stimuli on perceptual organization. When two regions of a uniform random-dot field are flickered in temporal alternation with the same flicker rate, one of the regions appears to lie in front of the other. Within the range of temporal frequencies used in the present experiments, depth perception was maximal between 5 and 31.3 Hz. Which region of the two is perceived as lying in front is different from person to person and sometimes fluctuates within the same subject, but when two regions are of different sizes, the smaller region tends to be perceived in front for longer than the larger region. The depth segregation was not due to a luminance difference, because the average temporal luminance of the regions was kept equal. Strikingly, the illusory depth segregation is perceived even between two adjacent regions whose densities of dots, sizes, shapes, and flicker rates are identical. This result suggests that a difference of temporal phase between two flickering regions is crucial for this new depth perception.

Many studies have pointed out that depth cues in a visual scene are necessary for monocular depth perception, and they are classified into several classes, such as pictorial cues and kinetic cues (e.g., Coules, 1955; Gogel, 1964; Levine & Shefner, 1991; O'Shea, Blackburn, & Ono, 1994; Rogers, 1993).

Recent studies have added flicker to the list of monocular depth cues. In a series of experiments, Weisstein and colleagues showed that when some regions of a random-dot field are flickered and the others are not, the flickering regions appear to lie behind the nonflickering regions. Their work also indicated that the nonflickering region was perceived as figure, whereas the flickering one was perceived as ground (Meyer & Dougherty, 1987; Wong & Weisstein, 1984, 1985, 1987). This "flicker-induced ground" effect was optimal when the flicker rates were between 6 and 8 Hz. At low (1.4 Hz) and high (12.5 Hz) flicker rates, the depth segregation between the

flickering and nonflickering regions diminished (for review, see Weisstein, Maguire, & Brannan, 1992).

Weisstein and colleagues proposed a temporal-frequency channel theory to explain why flicker-induced depth segregation was perceived (e.g., Weisstein et al., 1992; Weisstein & Wong, 1986). According to the theory, a low-temporal-frequency channel carries the nonflickering stimulus and determines the figure, whereas a high-temporal-frequency channel carries the flickering stimulus and determines the ground. Furthermore, activities of these two temporal channels locate the nonflickering region in front and the flickering region behind in depth. Thus, the temporal-frequency content of suprathreshold textures in the adjacent regions reliably and systematically determines their figure-ground organization (Klymenko & Weisstein, 1986; Klymenko, Weisstein, Topolski, & Hsieh, 1989; Wong & Weisstein, 1984, 1985, 1987).

The theory that flickering signals are transmitted through a high-temporal-frequency channel specifying the ground region brought to mind the following question: If two regions are flickered under the same conditions (e.g., luminance, contrast, average spatial frequency, average temporal frequency) except for temporal phase (antiphase), how are the two flickering regions perceived? If signals with the same temporal frequency are transmitted through the same temporal channel, irrespective of temporal phase, signals from both flickering regions should be located in the same plane. In this case, no difference between these two regions would be perceived, and the two regions would seem to lie in the same depth plane. On the other hand, if temporal phase exerts a crucial effect on figure-ground organization or depth segregation, the two regions would be perceived as different.

This research was conducted as part of a project of the International Media Research Foundation, Tokyo, to whom the authors are grateful. This study was supported by Special Coordination Funds from the Science and Technology Agency of the Japanese Government. The authors would like to thank Edgar Körner and Ernst Pöppel for their stimulating discussions. The authors also wish to thank Robert P. O'Shea and the reviewers, for their help and suggestions in revising the manuscript, and Yuede Yang, for her suggestion of the boundary hypothesis and the parallax hypothesis in the Discussion. Thanks are also due Miyako Tasaki, Noriko Kashima, Hiroyoshi Kobiyama, Yoshiyuki Miwa, Kazumi Otsubo, Yasuhiro Kondoh, and Keigo Noda for their help and encouragement. Correspondence should be addressed to A. Iwabuchi, 3-5-1003 Ohyamacho, Itabashi-ku, Tokyo 173, Japan (e-mail: a-i@wa2.so-net.or.jp).

To clarify the role of temporal phase in flicker-induced figure-ground organization, we studied how two such flickering regions in antiphase are perceived. We found that depth segregation could be perceived between the two flickering regions with an optimal frequency of 10.4 Hz. The average temporal luminance of the two flickering regions was kept equal, and the depth was perceived even when the sizes of the two regions were identical; therefore, the depth segregation is not due to a luminance or size difference.

In the present paper, we will report this new depth phenomenon induced between two flickering regions in antiphase and suggest a role of temporal phase in figure-ground organization and monocular depth segregation in a visual scene.

EXPERIMENT 1

Method

Apparatus. All of our stimuli were generated under computer control on an AV tachistoscope (Iwatsu Isel, IS-701AB) whose frame rate was set at 4 msec/frame. The size of the screen was 30×18 cm (width \times length). The experiment room was darkened, and the display was installed into a cavity lined with black non-reflecting felt in order to eliminate extraneous stimulation from other parts of the room.

Stimuli. A random-dot-filled square ($6.0^\circ \times 6.0^\circ$) was divided into two regions: an inner square and an outer edge (Figure 1). One

of three sizes of the inner square ($1.5^\circ \times 1.5^\circ$, $3.0^\circ \times 3.0^\circ$, or $4.5^\circ \times 4.5^\circ$) was used, and the remaining region of the original square was used as the outer edge. Each dot subtended a $2.26' \times 2.26'$ of arc, and the dot density was 27.2%. There were no contours between the regions, so when the two regions were presented simultaneously without flicker, they were indistinguishable.

The two random-dot-filled regions were generated in advance and stored in two frames (Frame A and Frame B) of the tachistoscope. When Frames A and B alternated (i.e., A-B-A-B...), one cycle lasted 8 msec, for a temporal frequency of 125 Hz. To alter this frequency, we alternated two or more successive exposures of Frame A with the same number of exposures of Frame B. For example, two exposures (i.e., AA-BB-AA-BB...) yielded a cycle length of 16 msec and a temporal frequency of 62.5 Hz. In a similar manner, we produced temporal frequencies of 41.7, 31.3, 25, 20.8, 15.6, 10.4, 6.6, and 5 Hz (Figure 2). This range of temporal frequencies was selected from pilot studies showing that, at lower rates, apparent movement was perceived and, at higher rates, the regions were indistinguishable.

The average temporal luminance of the flickering stimuli was 11.4 cd/m^2 (each dot had a luminance of 83.8 cd/m^2 so that, averaged over space and time, a patch of the stimulus had a luminance of 11.4 cd/m^2), and background luminance was 0.1 cd/m^2 . The luminance of the stimuli was set low so that phosphor persistence was reduced to a minimum during flicker. The monitor had P4 phosphor with a persistence of 1.2 msec.

The stimuli were observed monocularly (with left or right eye) at a distance of 80 cm through a viewing window. The dots were luminous, not black as in the figures.

Subjects. Eight paid naive subjects (3 undergraduate students, 1 graduate student, and 4 office workers) and 1 of the authors (A.I.)

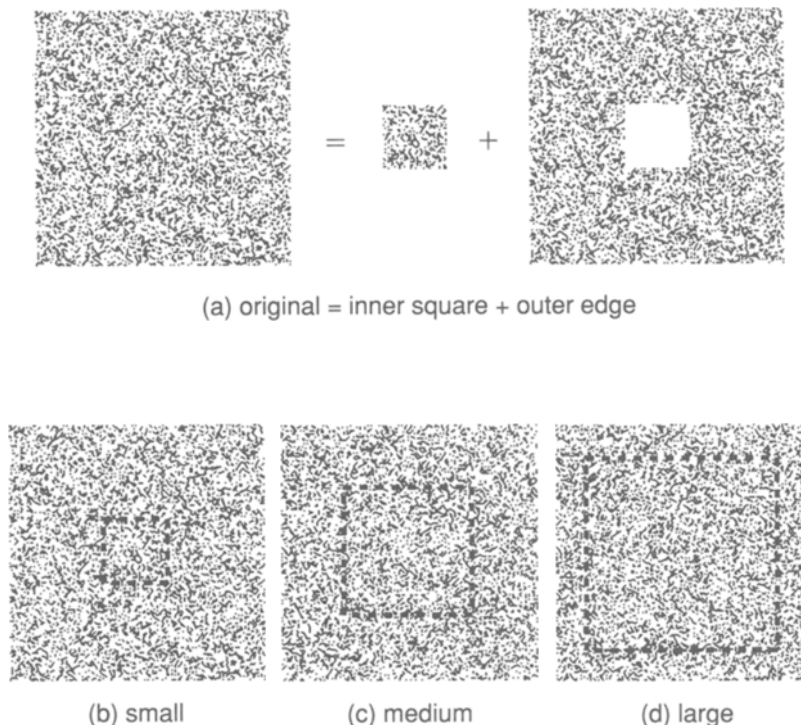


Figure 1. Stimulus configurations used in Experiment 1. A random-dot-filled square ($6.0^\circ \times 6.0^\circ$) was divided into two regions, an inner square and an outer edge (a). These two regions were alternated continuously. In each trial, one of three sizes of the inner square was used: (b) SMALL ($1.5^\circ \times 1.5^\circ$), (c) MEDIUM ($3.0^\circ \times 3.0^\circ$), or (d) LARGE ($4.5^\circ \times 4.5^\circ$). Each dot subtended a $2.26' \times 2.26'$ of arc, and the dot density was 27.2%. There were no contours on the pictures, and the border between the two regions was defined only by temporal phase difference. The dots were luminous, not black as in the figures.

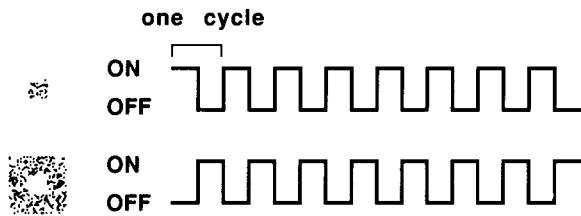


Figure 2. Schematic representation of flicker. The inner square and the outer edge were repeatedly presented in alternation. In each trial, 1 of 10 kinds of on-off square-wave flickering stimuli with a flicker rate of 125, 62.5, 41.7, 31.3, 25, 20.8, 15.6, 10.4, 6.6, or 5 Hz was used. Interstimulus interval was 0. Total presentation duration for a trial was 20 sec. Refresh rate of the monitor was 4 msec/frame.

participated in this experiment (7 males, and 2 females). They ranged in age from 21 to 32 years, with a mean age of 24.2 years. All subjects had normal or corrected-to-normal vision.

Procedure. The three size conditions, each with 10 temporal frequencies, generated 30 stimulus conditions. Each stimulus was presented once during the experiment, for a total of 30 trials for each subject. Sequences of presentation of the 30 trials were totally randomized. In each trial, the two regions were alternated continuously for 20 sec.

In preliminary experiments, the subjects were instructed to report what they saw during presentation of a total of 30 stimuli, and all ($N = 8$) reported spontaneously perceiving depth segregation between the two regions at examined temporal frequencies below

31.3 Hz (Figure 3). The subjects typically reported that the subjective appearance of the perceived depth was very similar to the impression of the depth obtained from a random-dot stereogram. The magnitude of the perceived depth was less than 1 mm, and the perceived depth was unstable: The region perceived to be in front was different from subject to subject, and sometimes the region in front and the region behind spontaneously alternated while the subject was viewing the display. In summary, at examined frequencies below 31.3 Hz, the subjects perceived depth segregation most of the time. In contrast, at temporal frequencies of 31.3 and 62.5 Hz, regional segregation without depth segregation predominated. At frequencies above 62.5 Hz, neither depth segregation nor regional segregation was perceived.

In Experiment 1, we investigated depth perception more quantitatively by measuring the cumulative times of four different experiences.

The subjects used a chinrest and viewed the stimuli monocularly. The fixation cross was presented in the center of the display at all times during the session; the subjects were instructed to fixate on this cross.

Before being tested, the subjects were given the following instructions:

Look at the fixation cross at the center of the display and answer which of the following is true: (1) the inner square looks like it is in front of the outer edge; (2) the inner square looks like it is behind the outer edge; (3) the inner square and the outer edge appear to lie in the same depth plane; and (4) other (it is difficult to judge the percept).

The subjects were instructed to respond by moving a joystick in front of them. Each subject underwent several practice trials before

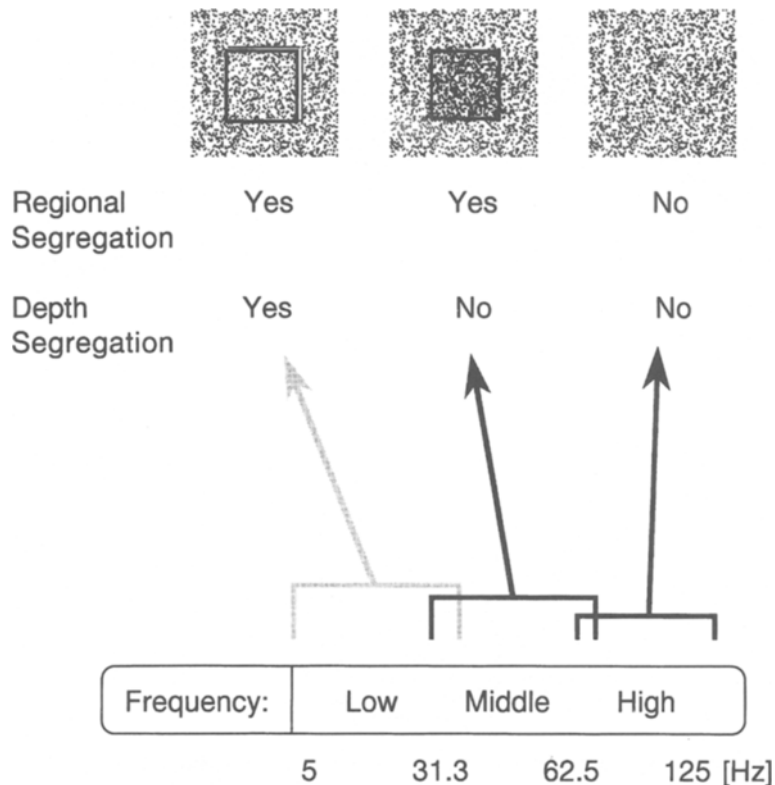


Figure 3. Schematic diagram of visual impression of the inner square/outer edge picture when the two regions flickered in antiphase. In the high-frequency range, regional segregation could hardly be perceived. In the middle range, regional segregation, but not depth segregation, was mainly perceived. In the low-frequency range, all subjects could perceive depth between the two regions.

the test trials. In the test sessions, a total of 30 trials were presented to each subject. The order of presentation was randomized. Each display was exposed for 20 sec; the interval between displays was 4 sec, during which time the screen was dark.

During the trials, the position of the joystick was recorded continuously. The subject was required to respond as soon as the display was turned on and to continue responding until the display turned off.

Results

We found that the percentage of response time (RT) that a region was seen as lying in front (or lying behind) was strongly dependent on the flicker rate. Although there were individual differences, all subjects perceived depth segregation. The average temporal luminance of the two flickering regions was kept equal, so depth segregation was not due to a luminance difference. Similar to in the preliminary experiment, the region perceived to be in front and that perceived to be behind alternated from time to time even for the same subject.

Figure 4 shows the mean percentage of RT during which depth segregation was perceived between the two flickering regions and the mean percentage of RT during which depth segregation was not perceived between the two regions.

The percentage of RT that the subjects answered "other" was 10.1%–34.7%, because most subjects took several seconds to judge the depth relation and to identify the percept by moving the joystick. At all of the examined frequencies below 31.3 Hz, the subjects perceived depth segregation most of the time. In contrast, at temporal fre-

quencies of 62.5 and 125 Hz, depth segregation was hardly perceived.

Figure 4 also shows that there was a relationship between the size of the inner square and which region was seen to be in front during the depth perception. When the inner square was small or medium, it was perceived as lying in front most of the time that depth was perceived. When the inner square was large, however, the percentage of RT in which it was perceived to be in front and the percentage of RT in which it was perceived to be behind were comparable. This suggests that smaller regions tend to appear in front of larger regions.

In Figure 5, replotted data from Figure 4 are shown, with the percent RTs of "perceived in front" and of "perceived behind" summed and presented as percent RT of "perceived depth." Mean percent RTs of perceived depth of three configurations (SMALL, MEDIUM, and LARGE) are shown in the upper-right panel of Figure 5. Note that the peak of depth perception was at 10.4 Hz.

The percent RT of perceived depth ("perceived in front" + "perceived behind") was analyzed. To fit psychometric functions to the three curves in the upper-left panel of Figure 5, we replotted the data of perceived depth to between 10.4 and 62.5 Hz (Figure 6a).

To test whether the psychometric functions are significantly different or not, we transformed the data as z-scores using the cumulative probability density function. The z-scores should portray the linear depth scale underlying the sigmoidal shape of the psychometric functions. Data having a percentage value of 0 have indeter-

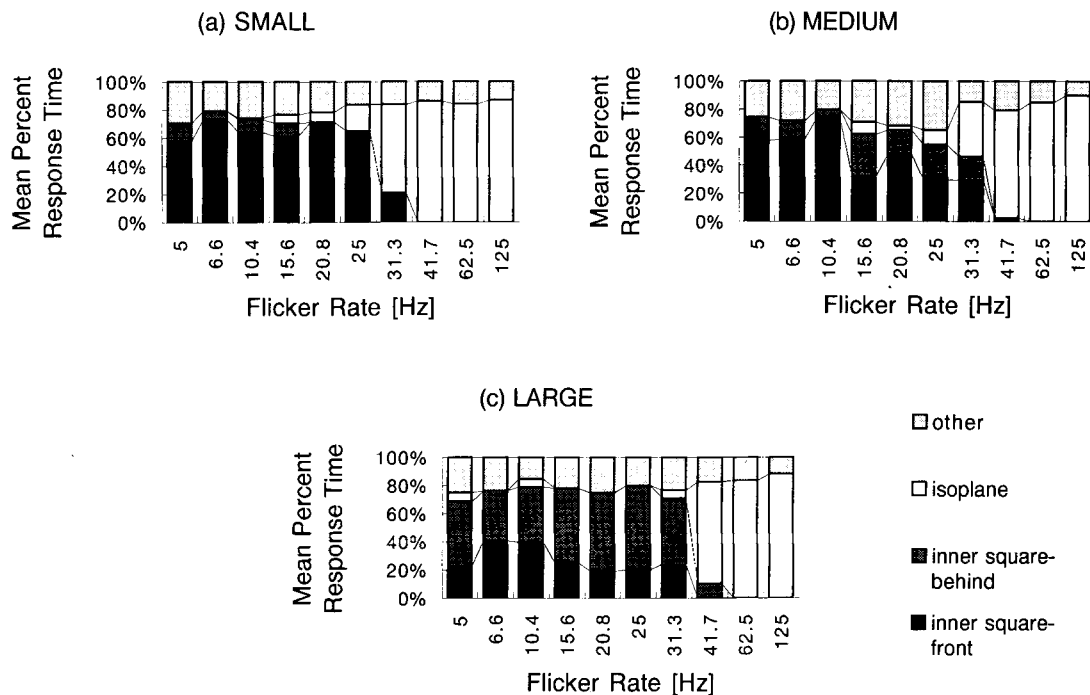


Figure 4. Results of Experiment 1. The mean percent of response times that a region was seen to be in front are plotted as a function of temporal frequency and size of the inner square for (a) SMALL, (b) MEDIUM, and (c) LARGE inner squares. Note that data are the means of 9 subjects.

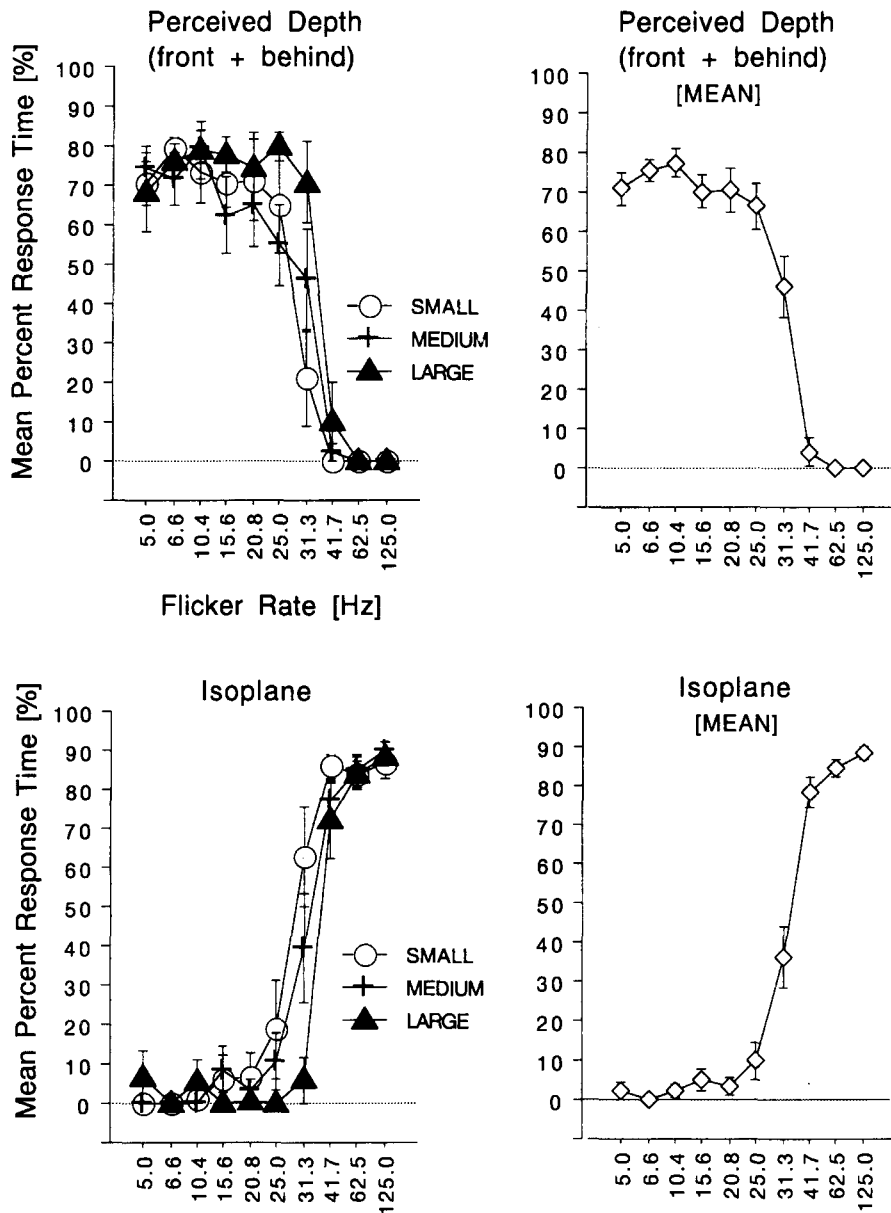


Figure 5. Replotted data from Figure 4. In the upper two panels, mean percent response times of “perceived depth” are plotted: time for each configuration and the average time are plotted in the upper-left panel and in the upper-right panel, respectively. In the lower two panels, mean percent response times of “isoplane” are plotted: time for each configuration and the average time are plotted in the lower-left panel and the lower-right panel, respectively. The vertical bars represent standard errors of the mean.

minable z-scores. For four such data, a z-score of -4.265 was calculated.

In Figure 6b, the best-fitting lines for the SMALL, MEDIUM, and LARGE inner squares are shown. We found a linear relationship between the mean percent RT of perceived depth and flicker rate for each group. The results of these linear regressions are given in Table 1. The 95% confidence intervals of the slopes for SMALL, MEDIUM, and LARGE overlap one another. If the slopes of the three lines have no statistical difference, then the size difference of

the inner square (SMALL, MEDIUM, or LARGE) had no significant effect on depth perception.

To test the significance of differences in the slopes of the three fitting lines, we conducted an analysis of variance (ANOVA). Deviations from a single regression line were not significant [$F(2,15) = 0.25$], indicating that the size difference of the inner square had no significant effect on depth perception. The regression equation for total z-scores (including SMALL, MEDIUM, and LARGE) is $y = 2.46 - 0.105x$, where y represents the mean percent

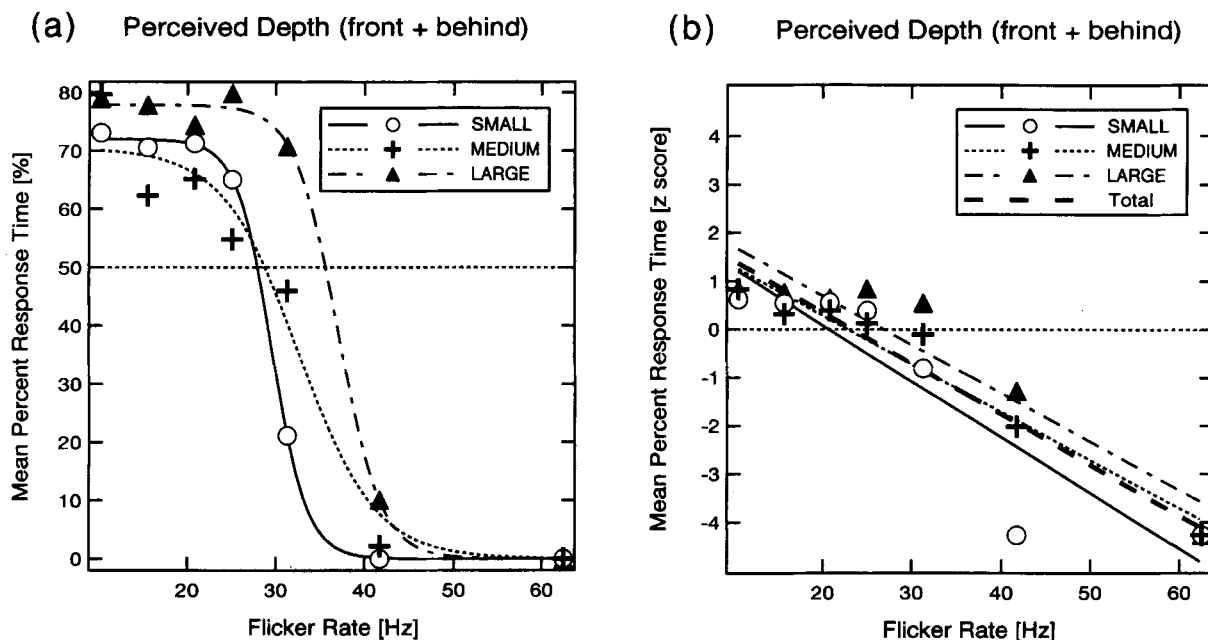


Figure 6. Replotted data from Figure 5. (a) Data of mean percent response times of “perceived depth” replotted to between 10.4 and 62.5 Hz. Psychometric functions were fitted to the data. (b) Data from Panel a transformed into z-scores and replotted. The linear regression coefficients of the best-fitting lines for each group are given in Table 1. The equations for the SMALL, MEDIUM, and LARGE inner squares are y (mean percent response time in z-score form) = $2.40 - 0.116x$ (flicker rate), $y = 2.28 - 0.100x$, and $y = 2.70 - 0.101x$, respectively. The equation for the total data is $y = 2.46 - 0.105x$ (see text).

RT in z-score form and x represents the flicker rate (see bold dashed line in Figure 6b). This equation accounts for 84.5% of the variance ($R^2 = .845$).

EXPERIMENT 2

Stimuli used in Experiment 1 had a center-surround organization, which may function as an occlusion cue to mediate depth segregation. In Experiment 2, by using another simple stimulus without a center-surround organization, we tested whether or not the depth effect could be perceived.

Method

Apparatus and Stimuli. The apparatus and stimuli were the same as those of Experiment 1, except that a picture of two adjacent rectangles (Figure 7), instead of the inner square/outer edge picture, was used as the display.

To produce regions, a random-dot-filled square ($6.0^\circ \times 6.0^\circ$) was divided vertically into two rectangular regions: a left rectangle and a right one. In each trial, one of three combinations of sizes

was used: (left rectangle; right rectangle) = ($1.5^\circ \times 6.0^\circ$; $4.5^\circ \times 6.0^\circ$), ($3.0^\circ \times 6.0^\circ$; $3.0^\circ \times 6.0^\circ$), or ($4.5^\circ \times 6.0^\circ$; $1.5^\circ \times 6.0^\circ$).

Subjects. Eight new paid naive subjects (5 undergraduate students and 3 graduate students) and 1 of the authors (A.I.) participated in this experiment (8 males, and 1 female). They ranged in age from 22 to 33 years, with a mean age of 24.6 years. All subjects had normal or corrected-to-normal vision.

Procedure. The experimental procedure was the same as that used in Experiment 1, except for the instructions.

Before being tested, the subjects were given the following instructions:

Look at the fixation cross at the center of the display and answer which of the following is true: (1) the right rectangle looks like it is in front of the left rectangle; (2) the left rectangle looks like it is in front of the right rectangle; (3) the right and left rectangles appear to lie in the same depth plane; and (4) other.

Results

The results were similar to those for the inner square/outer edge picture. Again, all subjects perceived depth segregation between the two regions at low temporal frequencies (Figure 8). This result indicates that the depth segregation in Experiment 1 can be perceived in another kind of picture. The percentage of RT that a region is seen to be in front (or behind) is strongly dependent on the flicker rate.

Figure 8 shows that the same region can be perceived as either in front or behind and that the smaller region was seen in front for longer than the larger region. There is, however, an asymmetry: The smaller region was perceived in front for longer when it was on the right (Figure 8c) than when it was on the left (Figure 8a). We sus-

Table 1

Linear Regression Coefficients for Relationships Between Mean Percent Response Time of Perceived Depth and Flicker Rate in Experiment 1

Inner Square	Intercept	Slope (B)	SE_B	R^2	β
SMALL	2.40	-.116	.024	.827	-.909†
MEDIUM	2.28	-.100	.011	.943	-.971†
LARGE	2.70	-.101	.017	.874	-.935†

Note— β = standardized coefficient of slope. † $p < .01$.

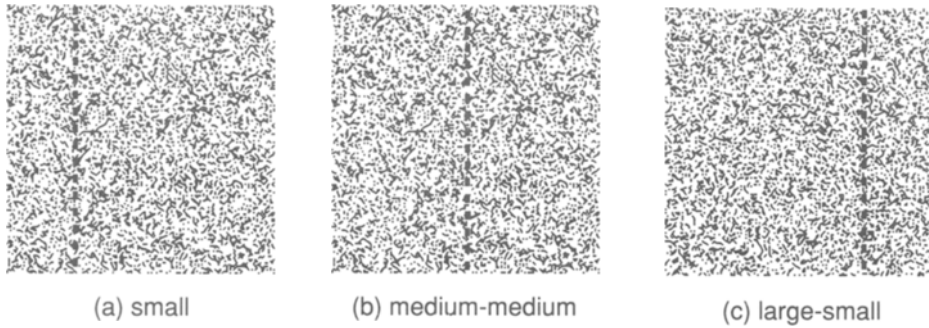


Figure 7. Stimulus configurations used in Experiment 2. A random-dot-filled square ($6.0^\circ \times 6.0^\circ$) was divided into two regions: left and right rectangles. In each trial, one of three configurations was used: (a) SMALL-LARGE (left, $1.5^\circ \times 6.0^\circ$; right, $4.5^\circ \times 6.0^\circ$), (b) MEDIUM-MEDIUM ($3.0^\circ \times 6.0^\circ$; $3.0^\circ \times 6.0^\circ$), or (c) LARGE-SMALL ($4.5^\circ \times 6.0^\circ$; $1.5^\circ \times 6.0^\circ$). Each dot subtended a $2.26' \times 2.26'$ of arc, and the dot density was 27.2%. There were no contours on the pictures, and the border between the two regions was defined only by temporal phase difference. The dots were luminous, not black as in the figures.

pect this may have been an artifact of the left-right direction of the monitor's raster.

In Figure 9, replotted data from Figure 8 are shown, with the percent RTs of "perceived in front" and of "perceived behind" summed and presented as percent RT of "perceived depth." Mean percent RT of perceived depth of three configurations (SMALL-LARGE, MEDIUM-MEDIUM, and LARGE-SMALL) is shown in the upper-right panel of Figure 9. Note that the sharp peak of depth perception was at 10.4 Hz.

As in Experiment 1, the percent RT of perceived depth ("perceived in front" + "perceived behind") was ana-

lyzed. To fit psychometric functions to the three curves in the upper-left panel of Figure 9, we replotted the data of perceived depth to between 10.4 and 62.5 Hz (Figure 10a). The z-scores were then calculated as in Experiment 1.

In Figure 10b, the best-fitting lines for the SMALL-LARGE, MEDIUM-MEDIUM, and LARGE-SMALL stimuli are shown. Again, we found a linear relationship between the mean percent RT of perceived depth and flicker rate for each group. The results of these linear regressions are given in Table 2. The standard errors of the slopes show that the slope for the LARGE-SMALL differs by more than

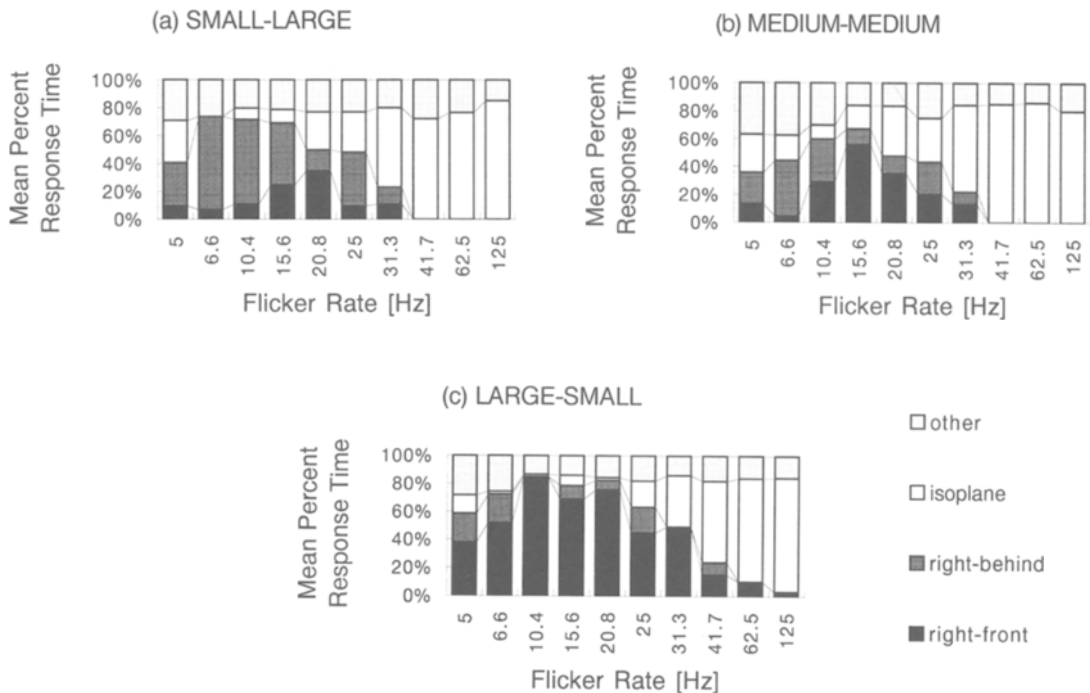


Figure 8. Results of Experiment 2. The mean percent response times that a region was seen to be in front are plotted as a function of temporal frequency and configuration for (a) SMALL-LARGE, (b) MEDIUM-MEDIUM, and (c) LARGE-SMALL configurations. Data are the means of 9 subjects.

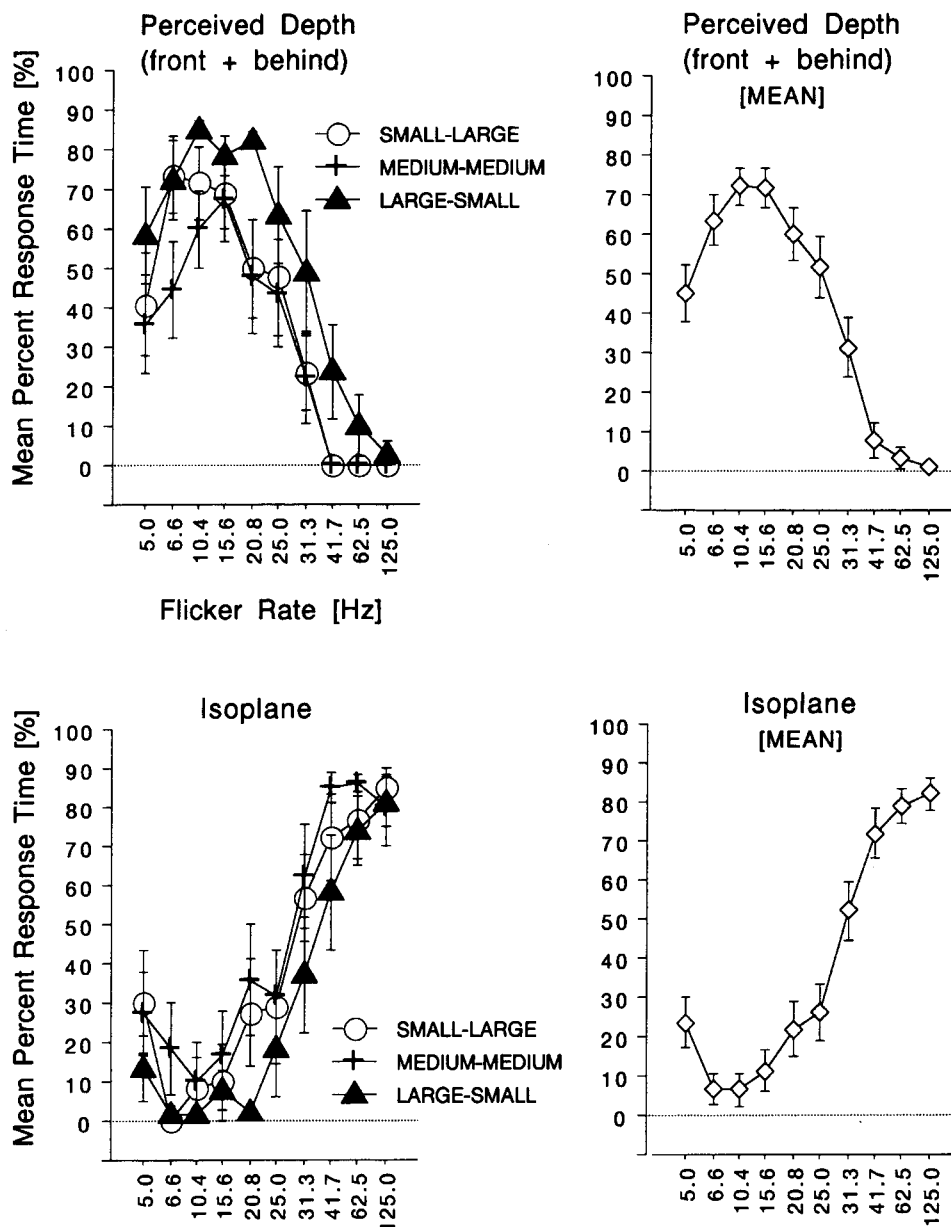


Figure 9. Replotted data from Figure 8. In the upper two panels, mean percent response times of “perceived depth” are plotted: time for each configuration and the average time are plotted in the upper-left panel and the upper-right panel, respectively. In the lower two panels, mean percent response times of “isoplane” are plotted: time for each configuration and the average time are plotted in the lower-left panel and the lower-right panel, respectively. The vertical bars represent standard errors of the mean.

the 95% confidence intervals of the slopes for the SMALL-LARGE and MEDIUM-MEDIUM, whereas the 95% confidence intervals of the slopes for the SMALL-LARGE and MEDIUM-MEDIUM overlap.

An ANOVA for testing the homogeneity of the three regression lines showed that the difference between the slopes of the regression lines for the three groups was significant [$F(2,15) = 3.92, p < .05$], indicating that the difference in stimulus configuration had a significant effect on depth perception. As mentioned above, this significance might have been due to the artifact.

As in Experiment 1, the depth segregation was not due to a luminance difference, since the average temporal luminance of the two flickering regions was kept equal. It should be noted that the depth segregation cannot be fully explained from the viewpoint of size difference or surrounding effect, because even in the case of the MEDIUM-MEDIUM pattern in Figure 7b, in which the regions are identical in size and form, all subjects perceived depth segregation, with no significant bias toward seeing one or the other region in front. In the case of the MEDIUM-MEDIUM pattern, the only difference between

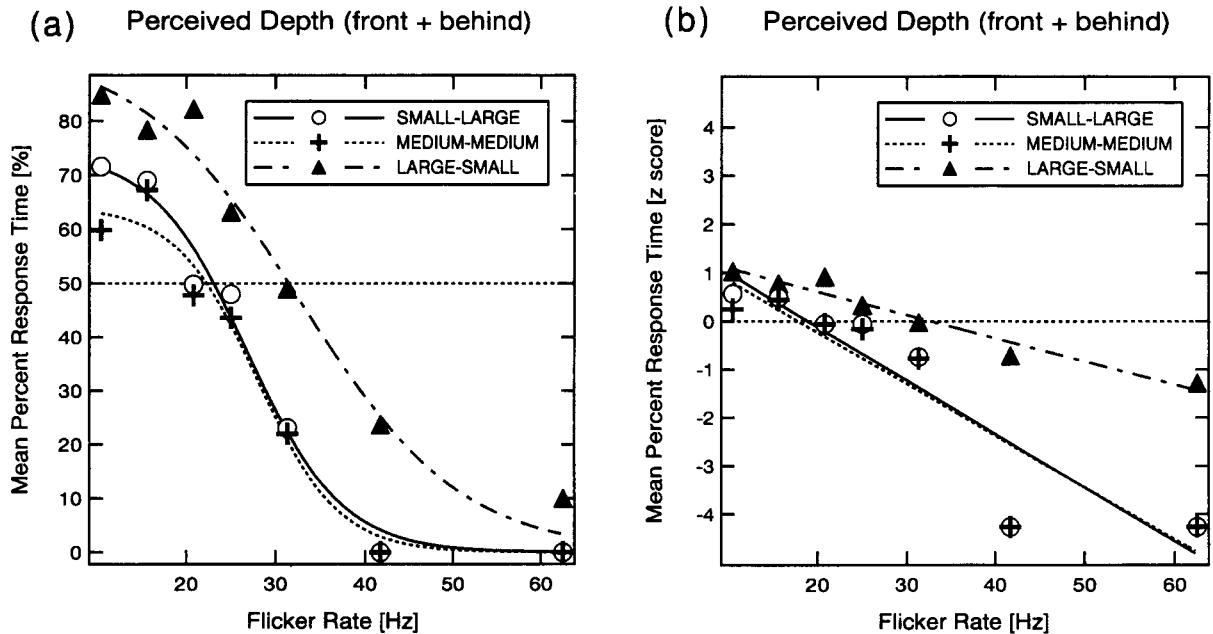


Figure 10. Replotted data from Figure 9. (a) Data of mean percent response times of “perceived depth” replotted to between 10.4 and 62.5 Hz. Psychometric functions were fitted to the data. (b) Data from Panel a transformed into z-scores and replotted. The linear regression coefficients of the best-fitting lines for each group are given in Table 2. The equations for the SMALL-LARGE, MEDIUM-MEDIUM, and LARGE-SMALL stimuli are $y = 2.10 - 0.111x$, $y = 1.91 - 0.107x$, and $y = 1.58 - 0.048x$, respectively.

the two regions was that of the temporal phase of flicker. This indicates that difference in the temporal phase is crucial for the depth perception.

EXPERIMENT 3

If the depth effect is affected only by the temporal phase, and not by the texture of the stimuli, the percepts of the stimuli would be similar, irrespective of dot density. The purpose of Experiment 3 was to study the effect of dot densities on the distinctness of the induced depth perception.

Method

Apparatus and Stimuli. The apparatus and stimuli were the same as those of Experiment 1, except that only one size of the inner square and outer edge was used.

To produce regions, a random-dot-filled square ($6.0^\circ \times 6.0^\circ$) was divided into two regions: an inner square and an outer edge. The size of the inner square was always $4.5^\circ \times 4.5^\circ$, which was the same size as that of the LARGE inner square in Experiment 1. For each trial, one of four different dot densities (dot density of 10.0%, 27.2%, 41.9%, or a solid picture) was used (Figure 11).

Table 2
Linear Regression Coefficients for Relationships Between Mean Percent Response Time of Perceived Depth and Flicker Rate in Experiment 2

Stimuli	Intercept	Slope (B)	SE _B	R ²	β
SMALL-LARGE	2.10	-.111	.021	.844	-.919†
MEDIUM-MEDIUM	1.91	-.107	.022	.831	-.912†
LARGE-SMALL	1.58	-.048	.005	.949	-.974†

Note—β = standardized coefficient of slope. †p < .01.

There were no contours on the regions, so when the two regions were presented simultaneously, it could not be noticed that the display consisted of two regions.

The four dot densities, each with 10 temporal frequencies (12.5, 62.5, 41.7, 31.3, 25, 20.8, 15.6, 10.4, 6.6, and 5 Hz) generated 40 stimulus conditions. Each stimulus was presented once during the experiment, for a total of 40 trials for each subject. Sequences of presentation of the 40 trials were totally randomized. In each trial, the two regions were alternated continuously for 20 sec.

The time average of luminance of the flickering stimuli was 4.2 cd/m² (dot density of 10.0%), 11.4 cd/m² (dot density of 27.2%), 20.1 cd/m² (dot density of 41.9%), or 62.6 cd/m² (the solid picture); background luminance was 0.1 cd/m².

Subjects. Eight new paid naive subjects (3 undergraduate students, 2 graduate students, and 3 office workers) and 1 of the authors (A.I.) participated in this experiment (2 males, and 7 females). They ranged in age from 21 to 32 years, with a mean age of 23.9 years. All subjects had normal or corrected-to-normal vision.

Procedure. The experimental procedure was the same as that used in Experiment 1, except for the total number of trials per subject (40 trials in Experiment 3; there were 30 trials in Experiment 1).

Results

Figure 12 shows the mean percent RT during which depth segregation was perceived between the two flickering regions and the mean percent RT during which depth segregation was not perceived between the two regions. Again, all subjects could perceive depth segregation, though there were individual differences.

When the picture was filled with random dots, the percepts of flickering regions seemed to be similar, irrespective of dot density (10.0%, 27.2%, or 41.9%). At examined frequencies below 31.3 Hz, the subjects perceived depth segregation most of the time. In contrast, at tem-

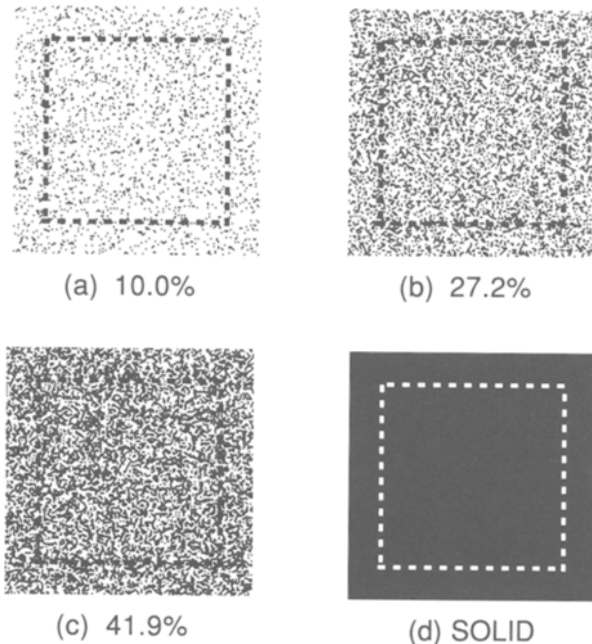


Figure 11. Stimulus configurations used in Experiment 3. A random-dot-filled square ($6.0^\circ \times 6.0^\circ$) was divided into two regions: an inner square and an outer edge. The size of the inner square was always $4.5^\circ \times 4.5^\circ$. In each trial, one of four different dot densities [dot density of (a) 10.0%, (b) 27.2%, (c) 41.9%, or (d) SOLID] was used. There were no contours on the pictures, and the border between the two regions was defined only by temporal phase difference. The dots were luminous, not black as in the figures.

poral frequencies of 62.5 and 125 Hz, depth segregation was weak (Figures 12a, 12b, 12c).

When the picture was solid, the percent RT of depth perception predominated at temporal frequencies of 41.7 Hz or below, and, against our expectation, the depth perception consumed a large amount of observation time even at temporal frequencies of 62.5 and 125 Hz (Figure 12d).

In Figure 13, replotted data from Figure 12 are shown, with percent RTs of "perceived in front" and of "perceived behind" summed and presented as percent RT of "perceived depth." Mean percent RT of perceived depth of four dot densities (10.0%, 27.2%, 41.9%, and SOLID) is shown in the upper-right panel of Figure 13. Again, the peak of depth perception was at 10.4 Hz.

As in Experiment 1, the percent RT of perceived depth ("perceived in front" + "perceived behind") was analyzed. To fit psychometric functions to the four curves in the upper-left panel of Figure 13, we replotted the data of perceived depth to between 10.4 and 62.5 Hz (Figure 14a). The *z*-scores were then calculated as in Experiment 1.

In Figure 14b, the best-fitting lines for the dot densities of 10.0%, 27.2%, and 41.9%, and for the solid stimuli are shown. Again, we found a linear relationship between the mean percent RT of perceived depth and flicker rate for each group. The results of these linear regressions are given in Table 3. The standard errors of the

slopes show that the slopes for dot densities of 10.0% and 27.2% and for the solid stimuli differ by more than each one's 95% confidence interval. Only the 95% confidence intervals of the slopes for dot densities of 27.2% and 41.9% overlap.

The results of an ANOVA testing the homogeneity of the four regression lines showed that the difference between the slopes of the regression lines for the four groups was significant [$F(3,20) = 19.00, p < .01$], indicating that the difference in the dot density of stimuli had a significant effect on depth perception.

DISCUSSION

General Properties

At first glance, our findings seem to indicate an effect similar to the flicker-induced depth segregation reported by Weisstein and colleagues (for a review, see Weisstein et al., 1992). However, there are certain differences between the previous findings and the present ones. Below, we summarize the differences.

Depth can be perceived between two regions with identical flicker rates. In most previous studies on flicker-induced depth segregation, the stimulus contained both a flickering and a nonflickering region, and depth segregation was perceived between the two regions, with the flickering region always stably perceived behind the nonflickering region (e.g., Wong & Weisstein, 1984, 1985, 1987). In a few studies, a similar depth perception was examined between two flickering regions with different flicker rates (Klymenko & Weisstein, 1989; Klymenko et al., 1989).

All of the above results suggest that the difference in temporal frequency is essential to the depth effect.

In this study, we found that depth segregation was perceived between two flickering regions with identical flicker rates, which suggests that temporal phase and temporal frequency played a crucial role in determining figure-ground organization (see also discussion below).

Optimal temporal frequencies for depth perception. Throughout the three experiments, the depth could be perceived for a long time between 5.0 and 31.3 Hz, and the optimal frequency for depth perception was 10.4 Hz.

We speculate on why the perceived depth became less distinct at higher flicker rates (beyond 31.3 Hz) as follows.

For depth segregation to be perceived, it is necessary that a boundary be formed between the two random dot regions, and the difference in temporal phase contributes to the boundary formation. As Y. Yang (personal communication, November 29, 1995) suggests, it is possible that, at higher flicker rates, the perceived boundary formed between the two flickering regions became less distinct, which impedes depth perception.

If this speculation is correct, it is likely that depth can be perceived for longer between two regions with high dot density than ones with low dot density, because the sharpness of the boundaries of flickering regions is increased as the random-dot density of flickering regions

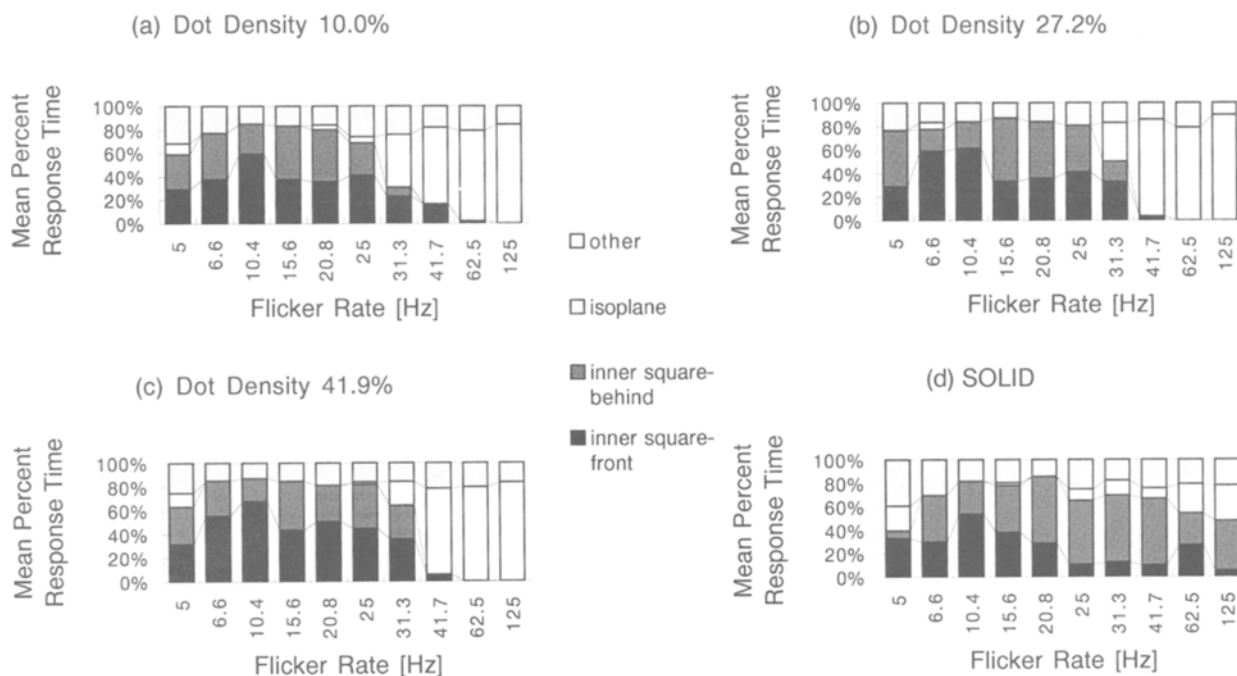


Figure 12. Results of Experiment 3. The mean percent response times that a region was seen to be in front are plotted as a function of temporal frequency and dot density for dot densities of (a) 10.0%, (b) 27.2%, (c) 41.9%, and (d) SOLID. Data are the means of 9 subjects.

increases, and this helps in the formation of the boundaries. The results of Experiment 3 support our hypothesis. At a high flicker rate (31.3 Hz, for instance, Figure 13), the subjects were better able to perceive depth as the random-dot density increased. Furthermore, when the two flickering regions were solid and had very sharp boundaries relative to the random-dot field, the subjects were better able to perceive depth than when the two regions were composed of random dots.

Most previous studies on depth segregation between flickering and nonflickering regions were conducted using square-wave on-off flicker, as in the present study; the maximal grounding of the flickering region behind a nonflickering region was found to occur for flicker rates of around 6–8 Hz, with a drop-off in the flicker-induced ground effect at lower and higher temporal frequencies (for contrast reversal flicker, see, e.g., Klymenko & Weisstein, 1989).

These tendencies were not dependent on the texture of the stimuli, on the specific configuration of the flickering and nonflickering areas, or on the flickering method. Among the stimuli used, for example, were flickering and nonflickering bars composed of random dots (temporal modulation 100%) at 1–12.5 Hz (Experiment 1 of Wong & Weisstein, 1984), flickering and nonflickering bars composed of random dots (temporal modulation 25%, 50%, 75%, or 100%) at 5–8.3 Hz (Experiment 2 of Wong & Weisstein, 1984), flickering and nonflickering bars made up of horizontal/vertical lines (temporal modulation 100%) at 3.6–12.5 Hz (Experiment 2 of Wong & Weisstein, 1985), flickering and nonflickering areas composed of concentric squares of random dots (tempo-

ral modulation 100%) at 3.6–12.5 Hz (Experiment 3 of Wong & Weisstein, 1985), flickering and nonflickering areas composed of the three segments of Rubin's faces/goblet reversible picture filled with random dots (temporal modulation 100%) at 1.4–12.5 Hz (Experiments 1A and 1B of Wong & Weisstein, 1987), and flickering and nonflickering areas consisting of subjective contours (temporal modulation 100%) at 0–25 Hz (Experiment 1 of Meyer & Dougherty, 1987).

In a few studies, however, no consistent results on the tuning of the flicker-induced ground effect with respect to temporal frequency were obtained. For example, using a pattern in which two regions containing sine-wave gratings underwent on-off flicker at 0–15 Hz, Klymenko et al. (1989) found that, under certain experimental conditions, the magnitude of the depth effect increased as the temporal frequency increased and reached its peak at 15 Hz.

Although the exact reason that the optimal temporal frequency for flicker-induced depth perception using on-off flicker in the previous studies (6–8 Hz) differs from that in the present study (10.4 Hz) is unknown, we speculate that it is due to a difference in the manner of presentation of the two regions. In our case, both regions were flickered, which tends to induce apparent motion from one region to the other at low flicker rates. It is possible that, at low flicker rates, including 6–8 Hz (which is the optimal frequency range for perceiving depth segregation between two flickering regions), apparent motion was perceived and the depth segregation was abolished, and optimal frequency for our case shifted to a higher temporal frequency (10.4 Hz).

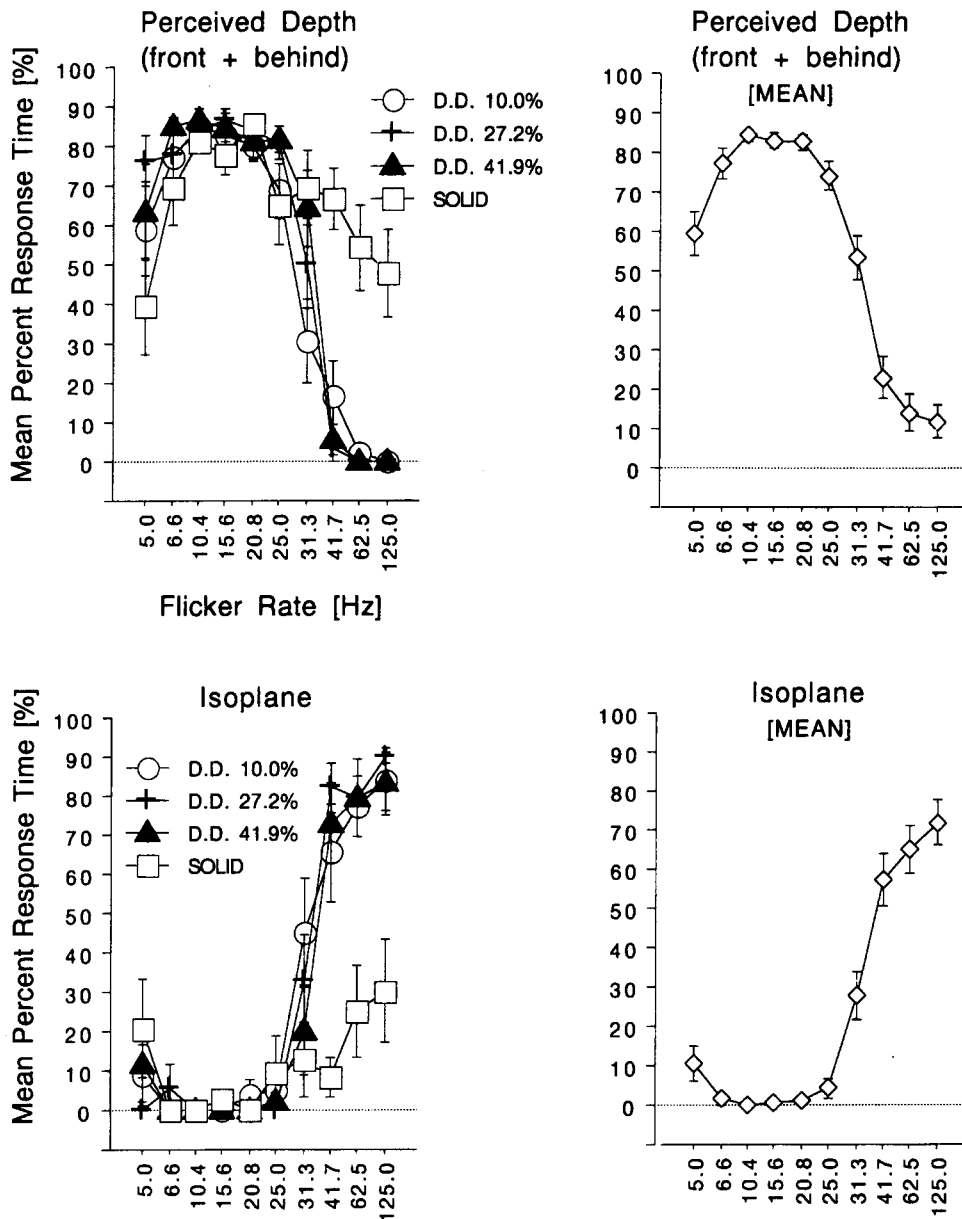


Figure 13. Replotted data from Figure 12. In the upper two panels, mean percent response times of “perceived depth” are plotted: time for each dot density and the average time are plotted in the upper-left panel and in the upper-right panel, respectively. In the lower two panels, mean percent response times of “isoplane” are plotted: time for each dot density and the average time are plotted in the lower-left panel and in the lower-right panel, respectively. The vertical bars represent standard errors of the mean.

Instability of perceived depth. The flicker-induced depth perception in previous studies seems to be highly stable, because the flickering region was always seen in front of the nonflickering region (e.g., Weisstein & Wong, 1986; Wong & Weisstein, 1984).

The spatiotemporal channel theory of Weisstein et al. (1992) explains the figure-ground organization in such previous studies fairly well, on the basis of the spatial frequency content of the stimuli as well as the temporal frequency content. The theory predicts that if regions

with relatively higher temporal frequency content are adjacent to regions containing relatively lower temporal frequency content, then the regions with the lower temporal frequency appear closer in depth than those with the higher temporal frequency, and that if regions with relatively higher spatial frequency content are adjacent to regions with relatively lower spatial frequency content, then the regions with the higher spatial frequency appear closer in depth than those with the lower spatial frequency.

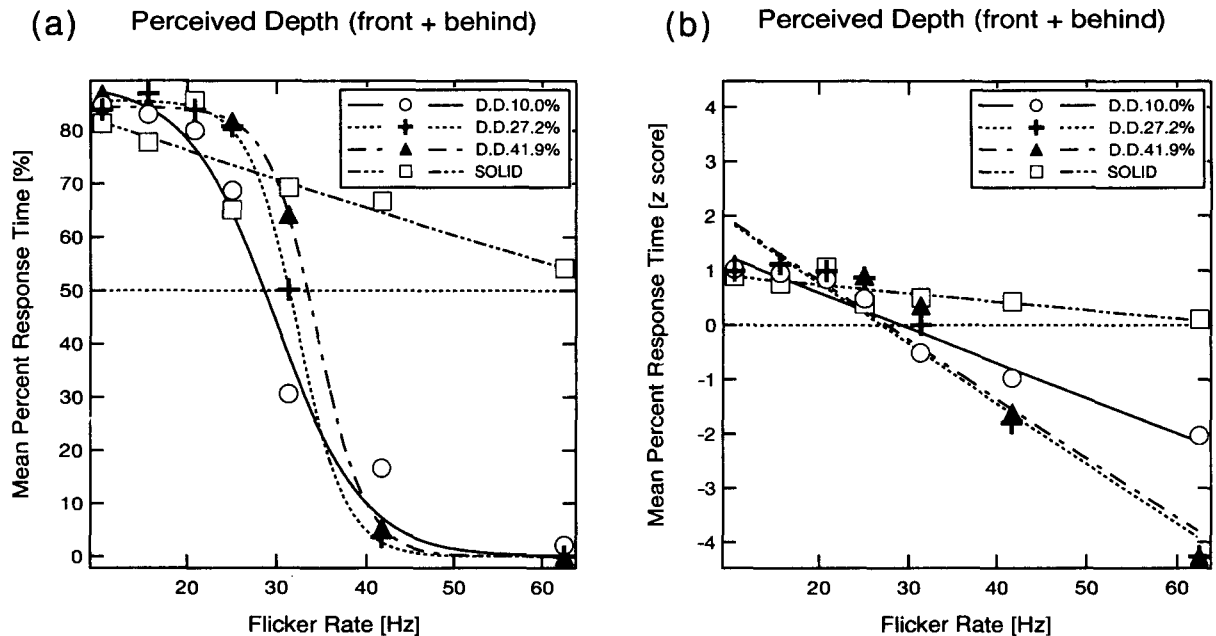


Figure 14. Replotted data from Figure 13. (a) Data of mean percent response times of "perceived depth" replotted to between 10.4 and 62.5 Hz. Psychometric functions were fitted to the data. (b) Data from Panel A transformed into z -scores and replotted. The linear regression coefficients of the best-fitting lines for each group are given in Table 3. The equations for the dot densities of 10.0%, 27.2%, and 41.9%, and for SOLID stimuli are $y = 1.88 - 0.065x$, $y = 2.98 - 0.111x$, $y = 3.01 - 0.109x$, and $y = 1.05 - 0.016x$, respectively. The abbreviation D.D. stands for dot density.

Roughly speaking, the pattern shown in our results indicating that smaller regions, on average, tend to be perceived in front for longer than larger regions is in accordance with their theory, because smaller regions have relatively higher spatial frequency content than larger regions. However, if the figure-ground organization can be completely determined by the spatial (and temporal) frequency content of stimuli, as the theory suggests, then the relative depths of figure and ground should be stable and unvarying. In other words, regions with relatively higher spatial frequency should be stable and should always appear to be in front. In the depth effect we found, however, the region perceived to be in front and that perceived to be behind alternated from time to time even for the same subject. This means that the smaller regions with relatively higher spatial frequency content sometimes switch from in front to behind.

Thus, it is difficult to explain fully the depth perception reported here by the theory of Weisstein et al. (1992) alone (see also discussion by Grossberg, 1994, p. 55).

Considering all of these differences, we conclude that the depth perception reported here cannot be attributed to previously reported factors.

Relationship With Other Monocular Depth Cues. It is known that pictorial depth cues are effective for inducing monocular depth perception (e.g., Levine & Shefner, 1991). For example, if there are two regions with different luminances at the same distance from an observer, the dimmer region is judged as being farther away than the brighter one (Coules, 1955). Our result cannot be accounted for in terms of luminance difference, since the average temporal luminances of the two flickering regions were identical. As another example, many investigators showed that relative size can be an effective cue of depth (e.g., Gogel, 1964). When two objects of different sizes are viewed from the same distance, the larger one is always thought to be closer to the subject. In our result, when depth was perceived between two regions, the smaller region tended to be judged more frequently as lying in front more frequently than as lying behind. Therefore, our results are not attributable to such size differences.

Motion parallax is also known to induce monocular depth perception (e.g., Braunstein, 1976; Rogers, 1993). It is generated when the observer moves with respect to a stationary world (self-produced parallax) or objects move with respect to a stationary observer (externally produced parallax). Using a stimulus display consisting of random-dot patterns, Rogers and Graham (1979) simulated the transformations of retinal image induced by externally produced parallax that would be produced while

Table 3
Linear Regression Coefficients for Relationships Between Mean Percent Response Time of Perceived Depth and Flicker Rate in Experiment 3

Dot Density	Intercept	Slope (B)	SE_B	R^2	β
10.0%	1.88	-.065	.006	.957	-.978†
27.2%	2.98	-.111	.013	.935	-.967†
41.9%	3.01	-.109	.014	.924	-.961†
SOLID	1.05	-.016	.005	.701	-.837*

Note— β = standardized coefficient of slope. * $p < .05$. † $p < .01$.

viewing a horizontal stripe lying in front of surrounding strips. In their experiment, the subject kept his/her head still while the experimenter moved the oscilloscope to and fro across the subject's line of sight. The experiment was set up so that the movement of the oscilloscope to the right (left) caused the center horizontal stripe of dots to move to the right (left) with respect to the surrounding random-dot strips. The subjects were able to perceive depth segregation between the center and the surrounding random-dot strips that moved in opposite directions.

There is a certain resemblance between the depth perception induced by Rogers and Graham's (1979) externally produced parallax and our flicker-induced depth perception. First, random-dot-filled regions were used in both experiments. In particular, the MEDIUM-MEDIUM stimulus used in our Experiment 2 (Figure 7b) can be regarded as a special case of externally produced parallax because alternating flicker of two adjacent MEDIUM stimuli is equivalent to repeated movement of one stimulus between two positions. Second, the subjects in both experiments reported similar visual impressions: During depth perception, no relative movement of the dots in the display was seen. In contrast, during apparent motion perception, depth was not perceived.

As Y. Yang (personal communication, November 29, 1995) suggests, the flicker-induced depth perception and the motion parallax-induced depth perception may share common mechanisms.

Role of Temporal Phase in Figure-Ground Organization Through Depth Segregation and Possible Mechanism Underlying Flicker-Induced Depth Perception

Weisstein et al. (1992) examined the relation between flicker-induced depth segregation and figure-ground organization using reversible figures, such as Rubin faces/vases, and found that flickering regions are predominantly perceived as background, whereas adjacent non-flickering regions are perceived as figures (Wong & Weisstein, 1984, 1985, 1987). On the basis of these results, they related flicker-induced depth perception and figure-ground organization and introduced a temporal frequency channel theory (e.g., Weisstein et al., 1992; Weisstein & Wong, 1986).

According to their theory, the region flickering at high frequency is coded by high-temporal-frequency channels to signal "a figure in front," whereas the nonflickering region (and region flickering at low frequency) is coded by low-temporal-frequency channels to signal "a ground behind." Furthermore, their theory emphasizes that figure and ground are processed in anatomically and functionally different pathways in the brain. It is widely accepted that there are two kinds of anatomically defined visual pathways in the brain: magnocellular (M) and parvocellular (P) pathways. Weisstein et al. (1992) related these anatomical pathways and psychophysically defined temporal frequency channels and concluded that M and P pathways could be identified with neural sig-

nals for "a ground behind" and "a figure in front" processing, respectively.

In the present study, we showed that depth segregation could be perceived between the two flickering regions with identical flicker frequencies but different temporal phases, which indicates that temporal phase was one of the critical factors of figure-ground organization. Although our results do not strongly conflict with Weisstein et al.'s theory of temporal frequency channels, additional explanation is needed because their theory does not fully explain our results. If Weisstein's notion that "figure in front" and "ground behind" are processed in anatomically different pathways is applicable to our results, then there must be at least two kinds of channel for each temporal frequency—for example, in-phase channels for figure processing and antiphase channels for ground processing. In our view, however, it seems unlikely to assume the existence of such a temporal phase-coding channel that is specialized for figure (or ground) processing for each temporal frequency, because our results showed that the same flickering region could be perceived as either figure (in front) or ground (behind) by the same subject. For this reason, we believe that the anatomical M (P) pathway is not specialized for ground (figure) processing; in other words, the same signal that is carried through the M (P) pathway can indicate either figure or ground.

Here, we consider possible mechanisms of flicker-induced depth segregation to explain how the temporal phase of the stimuli plays a role in depth segregation.

The theory of "binding by temporal coherence" is helpful in speculating on temporal phase-dependent mechanisms. According to this theory, synchronization of the cortical neurons underlies the processes by which visual features separated in a visual scene are combined to form perceptual groups (e.g., Crick & Koch, 1990; Eckhorn et al., 1988; Gray, König, Engel, & Singer, 1989; Grossberg, 1994; Keele, Cohen, Ivry, Liotti, & Yee, 1988; Shimizu & Yamaguchi, 1987).

We speculate that the synchronous mass activity of neurons is organized depending on the timing of presented visual stimuli. We hypothesize that information on random-dot stimuli from the in-phase flickering region is carried to the depth-related area of the cerebral cortex where dot information is grouped into perceptual units, such as rectangles, and each perceptual unit is represented by synchronous mass activity of the cortical neurons. Similarly, random-dot stimuli from the antiphase flickering region are represented by another burst of synchronous mass activity in the brain. Because a flickering stimulus induces phase-locked discharge of the cortical neurons (Bullier, Nowak, & Munk, 1994), it is plausible that two random-dot regions in antiphase flicker are represented by synchronous firing of two corresponding neuronal clusters in different temporal phases. We believe that the degree of perceived depth segregation is represented by the degree of phase difference between the two synchronously firing clusters.

In this work, we were concerned with a particular instance of the feature-binding problem—namely, the selective attribution of flickering signals to different depths in a visual scene—and we found that dots having the same temporal phase were assigned to the same depth plane.

REFERENCES

- BRAUNSTEIN, M. L. (1976). *Depth perception through motion*. New York: Academic Press.
- BULLIER, J., NOWAK, L. G., & MUNK, M. H. J. (1994). Temporal constraints in cortical processing. *European Journal of Neuroscience (Supplement)*, *7*, 127.
- COULES, J. (1955). Effect of photometric brightness on judgements on distance. *Journal of Experimental Psychology*, *50*, 19-25.
- CRICK, F., & KOCH, C. (1990). Towards a neurobiological theory of consciousness. *Seminars in Neuroscience*, *2*, 263-275.
- ECKHORN, R., BAUER, R., JORDAN, W., BROSCHE, M., KRUSE, W., MUNK, M., & REITBÖCK, H. J. (1988). Coherent oscillations: A mechanism for feature linking in the visual cortex? *Biological Cybernetics*, *60*, 121-130.
- GOGEL, W. C. (1964). Size cue to visually perceived distance. *Psychological Bulletin*, *62*, 217-235.
- GRAY, C. M., KÖNIG, P., ENGEL, A. K., & SINGER, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, *338*, 334-337.
- GROSSBERG, S. (1994). 3-D vision and figure-ground separation by visual cortex. *Perception & Psychophysics*, *55*, 48-120.
- KEELE, S. W., COHEN, A., IVRY, R., LIOTTI, M., & YEE, P. (1988). Tests of a temporal theory of attentional binding. *Journal of Experimental Psychology: Human Perception & Performance*, *14*, 444-452.
- KLYMENKO, V., & WEISSTEIN, N. (1986). Spatial frequency differences can determine figure-ground organization. *Journal of Experimental Psychology: Human Perception & Performance*, *12*, 324-330.
- KLYMENKO, V., & WEISSTEIN, N. (1989). Figure and ground in space and time: I. Temporal response surfaces of perceptual organization. *Perception*, *18*, 627-637.
- KLYMENKO, V., WEISSTEIN, N., TOPOLSKI, R., & HSIEH, C.-H. (1989). Spatial and temporal frequency in figure-ground organization. *Perception & Psychophysics*, *45*, 395-403.
- LEVINE, M. W., & SHEFNER, J. M. (1991). *Fundamentals of sensation and perception* (2nd ed.). Pacific Grove, CA: Brooks/Cole.
- MEYER, G. E., & DOUGHERTY, T. (1987). Effects of flicker-induced depth on chromatic subjective contours. *Journal of Experimental Psychology: Human Perception & Performance*, *13*, 353-360.
- O'SHEA, R. P., BLACKBURN, S. G., & ONO, H. (1994). Contrast as a depth cue. *Vision Research*, *34*, 1595-1604.
- ROGERS, B. (1993). Motion parallax and other dynamic cues for depth in humans. In F. A. Miles & J. Wallman (Eds.), *Visual motion and its role in the stabilization of gaze* (pp. 119-137). Amsterdam: Elsevier.
- ROGERS, B., & GRAHAM, M. (1979). Motion parallax as an independent cue for depth perception. *Perception*, *8*, 125-134.
- SHIMIZU, H., & YAMAGUCHI, Y. (1987). Synergetic computer and holonics—Information dynamics of a semantic computer. *Physica Scripta*, *36*, 970-985.
- WEISSTEIN, N., MAGUIRE, W., & BRANNAN, J. R. (1992). M and P pathways and the perception of figure and ground. In J. Brannan (Ed.), *Applications of parallel processing in vision* (Advances in Psychology, Vol. 86, pp. 137-166). Amsterdam: North-Holland.
- WEISSTEIN, N., & WONG, E. (1986). Figure-ground organization and the spatial and temporal responses of the visual system. In E. C. Schwab & H. C. Nusbaum (Eds.), *Pattern recognition by humans and machines: Visual perception* (Vol. 2, pp. 31-64). New York: Academic Press.
- WONG, E., & WEISSTEIN, N. (1984). Flicker induces depth: Spatial and temporal factors in the perceptual segregation of flickering and non-flickering regions in depth. *Perception & Psychophysics*, *35*, 229-236.
- WONG, E., & WEISSTEIN, N. (1985). A new visual illusion: Flickering fields are localized in a depth plane behind nonflickering fields. *Perception*, *14*, 13-17.
- WONG, E., & WEISSTEIN, N. (1987). The effects of flicker on the perception of figure and ground. *Perception & Psychophysics*, *41*, 440-448.

(Manuscript received May 31, 1996;
revision accepted for publication November 27, 1996.)