

Influence of flicker on perceived size and depth

R. J. MILLER and ROBERT PATTERSON
Washington State University, Pullman, Washington

Previous research (e.g., Wong & Weisstein, 1984a, 1985) has shown that flickering stimuli appear to be more distant than nonflickering stimuli at the same physical distance. Given this relation between flicker and perceived depth, inappropriate constancy scaling theories predict that flickering stimuli should be perceived as larger than nonflickering ones. In contrast, links between flicker and motion perception suggest that flickering stimuli should be perceived as smaller than nonflickering ones. Two experiments tested these contrasting predictions. In Experiment 1, 22 subjects compared flickering and nonflickering vertical lines and reported that the flickering stimulus appeared significantly smaller than the nonflickering one. In Experiment 2, 21 subjects reported that the stimuli used in Experiment 1 produced depth effects similar to those reported in previous experiments: flickering stimuli were perceived as more distant than nonflickering ones. The observed effect of flicker on perceived size was contrary to predictions from inappropriate constancy scaling theory, but consistent with views that motion and flicker are processed by the same pathway.

Several investigations have shown that flicker and depth perception are related. The usual paradigm involves a visual field divided into two or more areas flickering at different rates. Typically, flickering areas are perceived as more distant than nonflickering areas, and areas with higher flicker rates are perceived as more distant than areas of slower flicker (Klymenko & Weisstein, 1989; Klymenko, Weisstein, Topolski, & Hsieh, 1989; Wong & Weisstein, 1984a, 1984b, 1985, 1987). These effects are quite robust, remaining unaffected by variations in texture (e.g., Wong & Weisstein, 1985) or in the shapes of the flickering and nonflickering areas (e.g., Wong & Weisstein, 1984b, 1985, 1987), or by the presence or absence of contours surrounding the areas (Wong & Weisstein, 1987).

Several factors do influence the depth effect, however. In arrays comparing flickering and nonflickering areas, the effect is strongest when flicker is 6–8 Hz, weaker at higher or lower flicker rates (Wong & Weisstein, 1984a, 1984b, 1985, 1987). The sizes of the areas also can be significant. With vertical bars as stimuli, for example, Wong and Weisstein (1984a) found the effect to be strongest when bars were about 1° wide. The magnitude of perceived depth is highest at high modulation contrast, and when the ratio of the flickering area's average temporal (Talbot) luminance to the luminance of the nonflickering area is in the range of 0.75–1.0 (Wong & Weisstein, 1984a).

The possibility that flicker is a depth cue raises the question of whether flicker-induced depth produces the

same perceptual consequences as does depth perception induced by other cues. For example, are the effects of flicker-induced depth on size constancy similar to those produced by other depth cues? One answer is suggested by size constancy explanations for visual illusions. Probably the best known of these is the inappropriate constancy scaling theory of Gregory (1963, 1966, 1967, 1968, 1990), although other authors have reported variations (e.g., Leibowitz, Brislin, Perlmutter, & Hennessy, 1969; Tausch, 1954; for reviews, see Borresen, 1990; Coren & Girgus, 1978; Fisher, 1968; Robinson, 1972). According to this view, certain size illusions can be explained, in part, by depth inferences that lead to inappropriate scaling of the sizes of stimuli in the illusion arrays.

A common example is the Ponzo illusion, shown in Figure 1. Most observers perceive the top horizontal line as longer than the bottom one, although the two are actually equal in length. Advocates of inappropriate constancy scaling maintain that the radiating lines are processed as linear perspective, producing a perception of the top line as more distant than the bottom line. Since the two lines provide equal retinal images, and in the real world objects at different distances can only produce identical retinal images if the farther one is larger than the nearer (Euclid's law), the top line is inferred (and thence perceived) to be longer than the bottom line.

Support for inappropriate constancy scaling explanations of illusions is far from unanimous, and there are many competing explanations (e.g., Coren, Porac, Aks, & Morikawa, 1988; Gillam, 1980; Pressey & Epp, 1992). However, it is a parsimonious view for which there is considerable developmental, cross-cultural, and other empirical support (for reviews, see Borresen, 1990; Coren & Girgus, 1978; Fisher, 1968; Gregory, 1990; Over, 1968; Robinson, 1972). One might suspect, then, that flicker would produce illusions analogous to the

Partial support for this project came from funds provided for medical and biological research by the State of Washington Initiative Measure 171. The authors thank Michael Donnelly, Chad Lawrence, Scott Penningroth, and Jenny Cotterill for their technical assistance. Correspondence should be addressed to R. J. Miller, Department of Psychology, Washington State University, Pullman, WA 99164-4820.

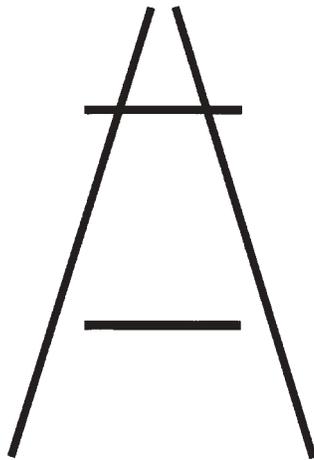


Figure 1. The classical Ponzo illusion. Most observers perceive the top line as longer than the bottom line, although the two are actually of the same physical length.

Ponzo: If a flickering stimulus is perceived as more distant than a nonflickering stimulus of the same physical size, the inappropriate constancy scaling approach suggests that the observer also will perceive the flickering stimulus as *larger* than the nonflickering stimulus.

An alternate line of reasoning would predict the opposite. There is a long tradition for the view that flicker and motion perception are related (Anstis, 1986). It has ties in psychology that extend at least as far back as Wertheimer's (1912) studies of stroboscopic motion, and, outside psychology, back to 18th- and 19th-century discoveries of the technical precursors of motion pictures (Macgowan, 1965). In modern terms, the relation is supported by the discovery that we are equipped with broadly tuned physiological mechanisms (i.e., channels or filters) that respond to both flicker and motion (Anstis, 1986; Buser & Imbert, 1987/1992; Green, 1981; Watson, 1986). For example, at the psychophysical level, it has been proposed that motion velocity is encoded from the relative activity in two broadly tuned temporal frequency channels (see, e.g., Thompson, 1982) or from the combination of outputs of several spatiotemporal frequency filters (Heeger, 1987). At the physiological level, Baker (1990) has shown that temporal frequency selectivity is systematically related to speed selectivity in cat cortical neurons. In addition, Schiller, Logothetis, and Charles (1990) have shown that lesions in the magnocellular pathway of primates impair both flicker and motion perception, implying that they share a common substrate.

Several investigators have demonstrated important relations between motion and size perception. One of the earliest, Ansbacher (1944), reported that a lighted arc on the edge of a rotating disk appeared to shrink as rotation speed increased. This effect, which has been confirmed in other experiments, occurs with linear as well as circular motion (Bhatia & Verghese, 1963; Marshall & Stanley, 1964; Stanley, 1970a, 1970b). Indeed, several investigators have shown that the relations between motion

variables (e.g., velocity) and size perception are quite regular and predictable (e.g., Bhatia & Verghese, 1964; Caelli, Hoffman, & Lindman, 1978). If flicker and motion are processed by the same pathway, one might expect analogous effects with flicker. For example, if one stimulus is flickering and another, equal-sized, stimulus is not, the Ansbacher and related effects would suggest that the flickering stimulus ought to be perceived as *smaller* than the nonflickering one.

In summary, the principal questions are the following: Will a flickering stimulus be perceived as different in size than an equal-sized nonflickering stimulus, and, if so, which one will be perceived as larger? Inappropriate constancy scaling theory suggests that the flickering stimulus should be perceived as the larger. Inference from the Ansbacher effect suggests the opposite. The purpose of the present study was to determine which prediction, if either, would be confirmed.

EXPERIMENT 1

Method

Subjects. Twenty-two subjects (13 females and 9 males) participated in Experiment 1. Their median age was 19.2 (range = 18–37) years. Those who required corrective lenses wore them during all tasks. All subjects had binocular and right-eye acuity (corrected or uncorrected) of at least 0.80. To minimize the risk of photoconvulsive epilepsy triggered by flickering stimuli, it was determined that no subject had any personal or immediate family history of seizures. All subjects gave informed consent to participate after the procedures had been fully explained. Each received credit toward a psychology course research requirement as compensation for participating.

Apparatus. The subject sat in a padded chair with an adjustable head and chinrest. Directly in front, perpendicular to the right-eye visual axis, was a back-projection screen, 94 × 122 cm, at a distance of 100 cm from the subject's eye. This viewing distance was selected to be consistent with earlier work with flicker as a depth cue (e.g., Wong & Weisstein, 1984a, 1985). Behind the screen were two Kodak Ektagraphic III ATS projectors with Kodak Ektanar *f*/3.5 lenses. Between each lens and the screen was a Lafayette Model 43011 electric shutter. The shutters were controlled by Hunter Decade Interval Model 111 Timers.

A small dot in the center of the screen, at the subject's eye level, marked the point that the subject fixated in order to begin each trial. Images of white vertical lines were projected onto the screen by the projectors. For each trial, two lines were visible—one to the left and one to the right of the fixation point. The inside edge of each line was 17.2 cm from the fixation point, with the center of each line at the height of the fixation point. Each line was 1.75 cm (1°) wide, the width Wong and Weisstein (1984a) found optimal for producing flicker-induced depth.

For any trial, one of the lines (the *standard*) was 23.33 cm (13.31°) long. The other (the *variable*) could be any of 13 lengths ranging from 21.93 cm (12.51°) through 24.73 cm (14.10°) in 0.23-cm steps. These lengths were selected after extensive pilot work with many combinations of lengths and widths. Given the assumption that the lines should be 1° in width to be consistent with previous research (as noted above), lengths were needed that accomplished several purposes. For one thing, the ratios of the lines' lengths to their widths should be sufficiently high to facilitate subjects' concentration on perceived differences in length alone, without their being distracted by width. There is, after all, no reason to expect that flicker should affect the perception of width any less than length. With the lines made sufficiently long, it was as-

sumed that instructions to judge the lines on the basis of length could be followed easily without any effects of flicker on width distracting subjects in their comparisons. Another consideration was the need for differences among variable line lengths to be sufficiently large so that they would be reliably discriminable, while being small enough to allow precision for the detection of even relatively small differences between flickering and nonflickering lines. Pilot work showed that the lengths described above satisfied all these criteria.

One of the two lines on each trial flickered at 6 Hz in an approximately square wave pattern. This rate was selected because it was within the range shown by Wong and Weisstein (1984a, 1984b, 1985, 1987) to be optimal for producing flicker-induced depth.

The luminance of the blank screen was 8 cd/m². The luminance of each line, without filters, was 2,000 cd/m². Neutral density (ND) filters could be inserted in front of the shutters. Five ND values were used: 0, 0.12, 0.30, 0.60, and 0.90. When in place, these filters produced stimulus line luminances of 2,000, 1,500, 1,000, 500, and 250 cd/m², respectively. When a line was flickering, the resulting time-average luminances produced by these filters were 1,000, 750, 500, 250, and 125 cd/m², respectively.

This apparatus made it possible to project two vertical lines for each trial. The subject could select either of the two lines by moving a joystick to the left or right, depending on which line he/she perceived as longer. For any trial, either line could be the standard or the variable, could be flickering or constantly illuminated, and could have any of the luminances described above.

Design. The principal independent variable was luminance ratio, defined as the ratio of the time-average luminance of the flickering stimulus to the luminance of the nonflickering stimulus. Five ratios defined the five levels of this variable: 0.25, 0.50, 1.0, 1.50, and 2.0. They were produced by projecting the nonflickering line at a luminance of 500 cd/m² and varying the luminance of the flickering line by using the filters described above.

Luminance ratio was varied because Wong and Weisstein (1984a) reported that it influenced the magnitude of perceived depth: depth was greatest with luminance ratios in the range of 0.75–1.0. It seemed possible that luminance ratio might also influence perceived size, especially if inappropriate constancy scaling were involved. In addition, brighter objects can appear closer than dimmer ones (Festinger, Coren, & Rivers, 1970; Graham, 1929; Ittelson, 1960; Oyama, 1960), and reducing their contrast causes some stimuli to appear more distant (Fry, Bridgman, & Ellerbrock, 1949; O'Shea, Blackburn, & Ono, 1993). Finally, there is the Brucke-Bartley effect, wherein flickering stimuli are brighter than nonflickering stimuli, even when their luminances are equated (Bartley, 1969; Riggs, 1972).

The dependent variable was the point of subjective equality (PSE), an estimate of which was provided by each series of trials. For each series, there were two critical line lengths, one for the standard line (always 23.33 cm), and one for the variable line perceived as subjectively equal to the standard line. The length of the latter was the PSE for that series of trials.

Procedure. After a brief introduction and a visual acuity test (conducted with a Bausch & Lomb Ortho-Rater), the subject sat at the apparatus and his/her position was adjusted to align the right eye with the fixation spot. An eyepatch was placed over the left eye, and a pair of stimulus lines, one flickering and the other not, was shown for 5 sec to provide an initial impression of the stimuli. Monocular viewing was incorporated in order to maintain consistency with earlier experiments (e.g., Klymenko & Weisstein, 1989; Klymenko et al., 1989; Wong & Weisstein, 1984a).

The subject was told that many pairs of lines resembling the ones just seen would be shown, and that he/she was to move the joystick, at the end of each 5-sec trial, to the left if the left line appeared longer or to the right if the right line appeared longer. The subject was told to return his/her gaze to the fixation spot before each trial. This use of a central fixation point to begin each trial

was adopted in order to be consistent with previous research (e.g., Klymenko & Weisstein, 1989; Wong & Weisstein, 1984a). However, subjects were assured that they could move their eyes in any manner they wished once the lines appeared on the screen. Next, 39 practice trials provided a variety of combinations of position (left or right line flickering), length, and luminance of the lines, to acquaint the subject with the full range of possible combinations.

Then, formal data collection began. The subjects were presented with 20 series of trials. For each series, 13 trials were presented, 1 for each of the 13 variable lines. Each trial consisted of a presentation of a pair of lines (the standard line and 1 of the 13 variables lines) for 5 sec. Each of the five luminance ratios was run as a block of four series of trials. The four series of trials corresponding to a given luminance ratio were run before changing to another luminance ratio.

Four series were required for each luminance ratio because of the inclusion of two control procedures. One was for left–right position. For two of the four series, the flickering stimulus was on the subject's left; for the other two series, the line on the subject's right flickered. The other control prevented confounding whether a line was flickering or not with its role as the standard or variable line. For two of the four series (one with the flickering stimulus on the right and one on the left), the standard line flickered and the variable line did not. For the other two series, only the variable line flickered.

The order of presentation of the five luminance ratios was randomized for each subject. Within each luminance ratio, the four series of trials were presented in random order with two restrictions: no order was used more than once for each subject, and each of the four series appeared at least once in each ordinal position for each subject. Within each of the four series, the 13 trials occurred in a different random order.

Results

For each subject, the point of subjective equality was determined for each of the 20 series of trials. Although the stimuli in each series of trials were presented to the subject in random order (thus resembling a method of constant stimuli), for scoring purposes the subject's responses to each series were rearranged in ascending order of variable line length and scored as a method of limits task (Engen, 1972). The point in the series was determined above which the variable line was perceived as shorter than the standard line and below which the variable was perceived as longer than the standard. The obtained value was recorded as the PSE for that series.

The 20 series of trials corresponded to the 20 cells of the $2 \times 2 \times 5$ within-subjects design. One independent variable was *flicker combination*: On half the trials, the standard line flickered and the variable line did not (Flicker Combination 1); on the other half, the variable line flickered and the standard line did not (Flicker Combination 2). The other independent variables were *position* of the standard line (left vs. right), and *luminance ratio* (0.25, 0.50, 1.00, 1.50, and 2.00). The mean PSE values for each of these cells are shown in Figure 2.

The effects of the three independent variables on PSE were tested in a $2 \times 2 \times 5$ (flicker combination \times position \times luminance ratio) analysis of variance (ANOVA). The effect of flicker combination was statistically significant [$F(1, 21) = 96.32, p < .001$]. None of the other main effects were significant, nor were any of the interactions. Although an examination of Figure 2 might lead

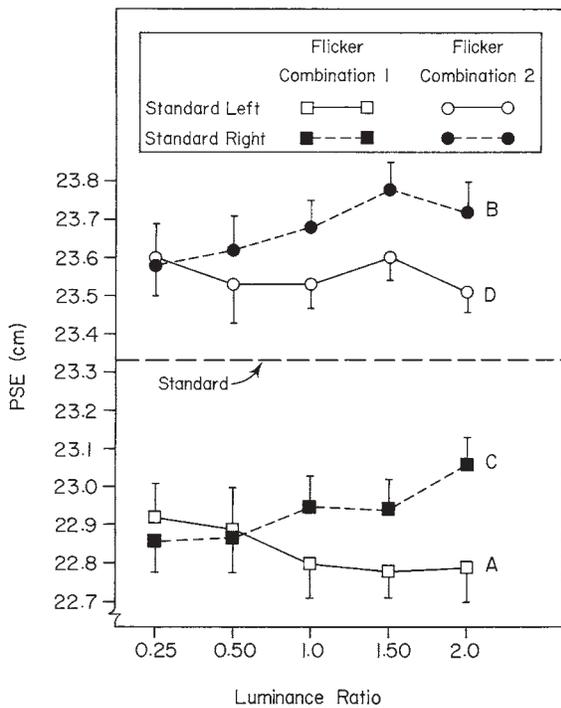


Figure 2. PSE data from Experiment 1. The 20 data points represent mean PSE values for all factorial combinations of flicker combination, position of the standard (left or right), and luminance ratio. Error bars correspond to standard errors of the mean. The horizontal dashed line represents the length of the standard line (23.33 cm).

one to suspect an interaction between position and luminance ratio, this interaction was not statistically significant, its probability exceeding .05.

The principal questions in this experiment were whether there was a significant illusion effect and, if so, whether the flickering line was perceived as longer or shorter than the nonflickering line. Presumably, if there had been no illusion, all PSE values would have approached 23.33 cm, the length of the standard line on all trials. To demonstrate an illusion effect, then, it becomes necessary to show that the PSE values departed significantly from 23.33. Since there were no significant effects of either luminance ratio or position, each subject's PSE values were collapsed across those variables.

The result was two PSEs for each subject, one for each of the two levels of flicker combination. For each subject, the PSE for Flicker Combination 1 was found by taking the mean of the 10 PSE values obtained under Flicker Combination 1. An analogous procedure was used to find the subject's PSE for Flicker Combination 2. For Flicker Combination 1, the overall mean PSE was 22.88 ($SD = 0.20$). Analysis showed this mean to be significantly different from 23.33 [$t(21) = -10.62, p < .001$, two-tailed]. For Flicker Combination 2, the overall mean PSE was 23.61 ($SD = 0.17$). This also was significantly different from 23.33 [$t(21) = 7.87, p < .001$, two-tailed].

To express these results in a different way, subjects underestimated the length of the flickering line by an average of 0.45 cm under Flicker Combination 1 and 0.28 cm under Flicker Combination 2; flickering lines were perceived as *shorter* than nonflickering lines. Although these illusion effects may seem rather small in absolute terms, it should be emphasized that they were phenomenally significant as well. When the apparatus was under construction, it was tested on a variety of observers in preliminary pilot work; these observers included the senior author and several graduate and undergraduate assistants. It was commonly observed that, when flickering and nonflickering lines were of the same real length, the flickering line clearly appeared smaller than the nonflickering line. This apparent difference was evident immediately; a 5-sec exposure of the lines was more than sufficient for the effect to be manifested.

The fact that the degree to which subjects underestimated the flickering line was different for Flicker Combination 1 than for Flicker Combination 2 raises the question of whether this difference was significant. To analyze this, the mean PSE values were again used. The absolute value of the difference between each of these two values and 23.33 was determined for each subject, yielding two difference scores, one representing the average amount by which the flickering line was underestimated under Flicker Combination 1 (overall mean = 0.45, $SD = 0.20$) and the other representing the equivalent value for Flicker Combination 2 (overall mean = 0.28, $SD = 0.17$). These two means were significantly different from one another [$t(21) = 6.71, p < .001$, two-tailed].

The preceding analyses all dealt with PSE values; all data were expressed in terms of the actual extent, in centimeters, to which the flickering line was underestimated. In some cases, however, one might be more interested in proportional effects. That is, if the length of a flickering line was underestimated (as compared with a nonflickering line), the amount of underestimation could be expressed as a proportion of the nonflickering line. Accordingly, each PSE was converted to a percent illusion value. For each series of trials, the two critical line lengths (the standard line, always 23.33 cm, and the PSE) were converted to percent illusion, $P = 100(N - F)/N$, where P = percent illusion, N = length of the nonflickering line, and F = length of the flickering line.

As would be expected, the majority (82%) of the percent illusion values were negative, indicating an underestimation of the length of the flickering line. The 20 mean percent illusion values that corresponded to the 20 cells of a $2 \times 2 \times 5$ (flicker combination \times position of the flickering line \times luminance ratio) design are represented in Figure 3. An ANOVA performed on these data showed that the effect of flicker combination was significant [$F(1,21) = 51.96, p < .001$]. Percent illusion was higher when the standard line flickered (Flicker Combination 1, overall mean illusion = -1.99%) than when the variable line flickered (Flicker Combination 2,

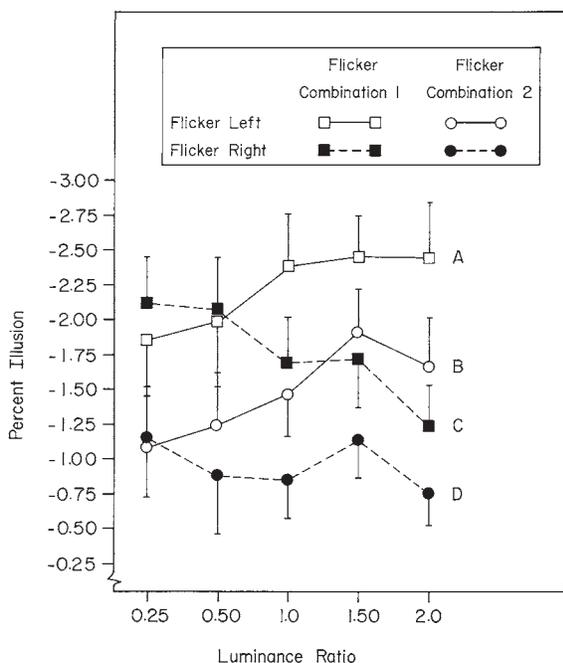


Figure 3. Percent illusion data from Experiment 1. The 20 data points represent mean percent illusion values for all factorial combinations of flicker combination, position of the flickering line (left or right), and luminance ratio. Error bars correspond to standard errors of the mean.

overall mean illusion = -1.21%). None of the other main effects or interactions were statistically significant.

To address the question of whether there was an overall significant illusion effect, percent illusion values were collapsed across position and luminance ratio, since these variables had no significant effect. For each subject, the mean of the 10 values for Flicker Combination 1 represented his/her percent illusion for that condition. The mean of the 10 values for Flicker Combination 2 represented his/her percent illusion for that condition. As indicated above, the overall means for the sample were -1.99% for Flicker Combination 1 and -1.21% for Flicker Combination 2. Each of these means was tested to determine whether it was significantly different from zero, since a complete lack of illusion would result in percent illusion values of zero. These analyses showed that the observed percent illusion values were significantly different from zero, both for Flicker Combination 1 [$t(21) = -10.62, p < .001$, two-tailed] and for Flicker Combination 2 [$t(21) = -7.86, p < .001$, two-tailed]. The finding that the percent illusion values were significantly different for Flicker Combination 1 than for Flicker Combination 2, already demonstrated with the previously described ANOVA, was confirmed with a t test analysis of the overall mean percent illusion values [$t(21) = 7.21, p < .001$, two-tailed].

Discussion

Experiment 1 showed that there was, in fact, an illusion effect: subjects underestimated the length of a flicker-

ing line as compared with a nonflickering line. This is the opposite of what would be predicted from an inappropriate constancy scaling explanation for size perception. Luminance ratio had no reliable effect on the illusion. The only independent variable that significantly influenced the size of the PSEs and the size of the illusion effect was flicker combination. The reason that flicker combination had a significant effect on PSE is relatively straightforward and easy to understand. Why flicker combination differentially affected the extent of the illusion, however, is more difficult to explain.

Examination of Figure 2 will show why flicker combination had a significant effect on the PSE values. Remember, the PSE values in this figure are based on a comparison of the standard line with the variable lines. In traditional size judgment tasks, the variable line serves as the scale, the "ruler" by which the perceived length of the standard is represented by the subject. The value of the PSE is the length of the variable line which the subject perceives as being equal to the length of the standard line. However, in the present experiment the standard line really had two lengths, at least perceptually, one on trials on which it flickered and another on trials when it did not flicker.

As we have shown, the effect of flickering a line is to reduce its perceived size as compared with when it is not flickering. On trials on which the standard line flickered (Flicker Combination 1, represented by Curves A and C in Figure 2), the fact that it flickered led to its being perceived as shorter than if it had not been flickering. Thus, the (nonflickering) variable line selected by the typical subject as matching the standard in length was relatively short, resulting in a PSE of less than 23.33 cm (the real length of the standard). On the other hand, on trials on which the standard did not flicker, it was perceived as longer than when it did flicker. Had the variable lines also not been flickering, one would expect that the variable line length chosen to match the standard would have been close to 23.33 cm and that the curve representing the PSE values would lie along the dashed line in Figure 2. However, the variable lines *did* flicker (leading them to be perceived as *shorter* than on the trials on which they were not flickering). Thus, the variable line length selected as equal to the standard was even longer than it presumably would have been if neither line had been flickering. As a result, the PSE values for this combination (Flicker Combination 2, represented by Curves B and D in Figure 2) were higher than the 23.33-cm length of the standard line. It is hardly surprising then that PSE values should show a significant effect of flicker combination.

More difficult to account for is the fact that the extent of the illusion effect varied, depending on flicker combination. There are two ways in which the effect of flicker on perceived length of stimulus lines might be made manifest. First, flicker might produce a constant *absolute* decrease in perceived length, independent of the real length of the line that is flickering. If that were the case, one would expect that the absolute amount by

which the flickering line was underestimated would be the same, regardless of flicker combination, and that the difference between all the PSEs shown in Figure 2 and the standard would be about the same (although the percent illusion for Flicker Combination 2 should be smaller than that for Flicker Combination 1). However, analysis showed that the PSEs for Flicker Combination 2 were significantly closer to the standard than were the PSEs for Flicker Combination 1. Thus, it is clear that flicker did not induce a constant absolute effect on perceived length.

Another possibility is that flicker produces a constant *proportional* decrease in perceived length, regardless of the real length of the flickering line. If that were the case, one would expect that there would be an effect of flicker combination on the difference between PSE and the standard, but not on percent illusion. If the effect of flicker is to reduce the perceived size of a line by a constant proportion as compared with when it is not flickering, the percent illusion scores should be about constant, regardless of whether the standard or the variable line flickers, and there would be no significant effect of flicker combination in the analysis of percent illusion scores. However, analysis showed that this effect was in fact significant. Thus, one must conclude that the effect of flicker on perceived length was not constantly proportionally related to the real length of the flickering line.

So what does account for the fact that the illusion effect varied with flicker combination? We can only speculate. It is possible that the effect of flicker on perceived length is some combination of absolute and proportional effects. Or perhaps having one stimulus line both flickering and varying in length from trial to trial (as was the case in Flicker Combination 2) forced closer attention to it, resulting in less of a tendency to underestimate its length relative to that of the nonflickering, standard line. In any case, this effect did not interact with any other variable. The most important finding remains: Significant illusion effects occurred in all conditions, regardless of whether the variable or the standard line flickered.

EXPERIMENT 2

The finding of significant underestimation of the flickering line in Experiment 1 raises at least two possibilities. One is that inappropriate constancy scaling does not account for flicker-mediated size perception. The other is that the assumption that the flickering stimulus appeared to be more distant than the nonflickering one is invalid. There were, after all, several differences between the stimuli used in Experiment 1 and those of previous researchers (e.g., Klymenko & Weisstein, 1989; Klymenko et al., 1989; Wong & Weisstein, 1984a, 1984b, 1985, 1987). Those researchers displayed stimuli on a computer CRT, whereas our stimuli were back projected onto a translucent screen. Our stimuli were comparable in width to those in previous experiments, but were longer. In many cases, those researchers used a height-to-width ratio of vertical bars of about 2:1, whereas we

used a ratio of about 13:1. Those researchers typically displayed bars in sets of four (each sharing a common boundary with the next), whereas we displayed only two stimuli (separated laterally by considerable space). Finally, in many cases, those researchers used textured stimuli containing dots or other patterns, whereas we used stimuli with no texture.

As mentioned earlier, flicker-induced depth is very robust for variations in stimulus shape, size, texture, and so forth, but it could be argued that the differences between Experiment 1 and previous experiments were great enough that they somehow eliminated the depth effect. It is possible that the observed illusion was inconsistent with inappropriate constancy scaling predictions because the necessary depth effect was not produced. Experiment 2 was designed to answer the question of whether the stimuli used in Experiment 1 in fact produced depth effects consistent with those in previous research.

The luminance ratios used in Experiment 1 were included in Experiment 2. In addition, since many of the stimuli used in earlier experiments have been textured, it was decided to compare textured and nontextured stimuli in Experiment 2. Wong and Weisstein (1985) found that texture essentially makes no difference to the flicker-induced depth effect, and we sought to determine whether this could be confirmed with the stimuli used in the present experiments.

Method

Subjects. Subjects for Experiment 2 were 12 males and 9 females who had not participated in Experiment 1. Their age range was 18–24 years, with a mean of 20.2 years. They met the same acuity and seizure-free-history criteria as had been applied in Experiment 1, and followed similar procedures for giving informed consent and receiving compensation.

Apparatus. With two differences, the apparatus was identical to that in Experiment 1. First, the stimulus lines were different. They were still projected in pairs, one of them flickering at about 6 Hz, as in Experiment 1. In Experiment 2, however, only two pairs of lines were used. For each pair, both lines always had the same length (23.33 cm). For half the trials, both lines were the same untextured stimuli as were used in Experiment 1. For the other half of the trials, both lines contained a pattern made up of random dots. The dots were 2.0–3.5 mm in diameter, were distributed randomly across the surface of each line, and numbered about 54 dots per line.

Thus it was possible to project two vertical lines for each trial, as in Experiment 1. For any trial, the lines could be textured or untextured. Either the left or the right line could be flickering, and the flickering line could have any of the five time-average luminances used in Experiment 1. As in Experiment 1, the nonflickering line always had a luminance of 500 cd/m².

Second, the dependent variable was different. Whereas in Experiment 1, subjects were asked to choose which of the lines in each pair was longer, in Experiment 2 they were asked (through use of the same joystick as in Experiment 1) to indicate which appeared *closer*. Furthermore, they could directly represent the amount of depth that they perceived by using a second joystick.

At various intervals (as described below) a dim red vertical line (0.1 × 3.0 cm) appeared at the center of the screen, 23 cm below the fixation spot. Slightly to the left of this line was a red spot, approximately 0.1 cm in diameter. The red line and spot were pro-

jected from behind the screen by a Metrologic 1.0-mW helium-neon laser (see Figure 4). The red spot was reflected from a front-surface mirror, mounted on a 0.075-rpm reversible-polarity dc motor, activation of which resulted in lateral movement of the projected spot at about 0.75 cm/sec. The second joystick was connected to the motor, making it possible for the subject to move the red spot to the left or right, thus varying the distance between the red spot and the stationary red line. The subject adjusted this distance to correspond to the amount of depth he/she perceived in each pair of stimulus lines. As shown below, the red line and spot were never visible concurrently with any of the stimulus line pairs.

Design. A 2 (textured vs. nontextured stimuli) \times 2 (left line flickering vs. right line flickering) \times 5 (luminance ratio) factorial within-subjects design was used in Experiment 2. Quantification of the dependent variable was based on which line the subject perceived as closer on each trial and how much closer he/she perceived it to be. The amount of perceived depth was assessed with the laser system. This quantity was expressed as a negative value if the flickering line was perceived as closer than the nonflickering line, and as positive if the nonflickering line was perceived as the closer one. On such a scale, zero indicates *no perceived depth* (i.e., both lines perceived as being at the same distance).

Procedure. The first steps were identical to Experiment 1. After an introduction and acuity screening, a pair of stimulus lines

(nontextured) was shown for 5 sec. The subject was told that other, similar pairs of lines would appear, and that for each 5-sec trial one of the lines would be flickering. His/her task for each trial was to decide which line in that pair appeared closer.

The subject was instructed to perform two tasks following each trial. The first was to indicate which of the lines in the pair appeared to be the closer by moving the first joystick to the left or right. The second task involved the laser-projected red spot and line that would appear on the screen as soon as he/she had completed the first task. The subject was instructed to adjust the distance between the red spot and line to show how much closer the stimulus line he/she had selected appeared to be (i.e., represent the amount of perceived depth by producing an equivalent distance between the red spot and red line). The subject was told to return his/her gaze to the fixation spot before each trial. Ten practice trials were then presented. These provided a variety of combinations of position (left or right) and luminance of the flickering line, with textured as well as nontextured pairs, to acquaint the subject with the range of possible combinations.

Then, 40 trials were conducted. These were run as five blocks of 8 trials each, one block for each of the luminance ratios. Order of presentation of the five luminance ratios was randomized for each subject. The 8 trials within each block were actually two consecutive presentations of a set of 4 trials. Of these 4 trials, 2 had textured lines (1 with the left line flickering, the other with the right line flickering), and the remaining 2 had nontextured lines (also 1 trial with the left line flickering, the other with the right line flickering). The 4 trials were presented in random order, and then the same 4 trials were presented again in a new random order. In scoring, the first 4 trials were treated as practice trials and discarded. The last 4 trials for each luminance ratio provided the formal data.

Results

For each trial, the magnitude of perceived depth (in centimeters), as indicated by the subject with the laser system, was recorded. If the flickering line was perceived as more distant than the nonflickering line, this value was recorded as positive. If the nonflickering line was perceived as more distant, the value was recorded as negative. The 20 trials that provided data were analyzed with a 2 (textured vs. nontextured) \times 2 (left line flickering vs. right line flickering) \times 5 (luminance ratio) repeated measures ANOVA. The means for all the cells in that analysis are shown in Table 1. None of the independent variables or any of the interactions were statistically significant.

The grand mean of all subjects' responses was 1.09. On the average, subjects perceived the nonflickering line as being 1.09 cm closer than the flickering line. Since the ANOVA showed no significant effects, the mean for each subject across all 20 conditions was determined. The resulting set of 21 values was then analyzed to determine whether they were significantly different from zero. This analysis resulted in a $t(20)$ of 4.36 ($p < .001$, two-tailed).

The results of Experiment 2 thus showed that subjects perceived the flickering line as more distant than the nonflickering line. This finding was consistent with previous studies, as cited in our introduction. None of the three independent variables (luminance ratio, the presence or absence of texture, or left or right position of the flickering line) had any statistically significant effect on this relation between flicker and perceived depth.

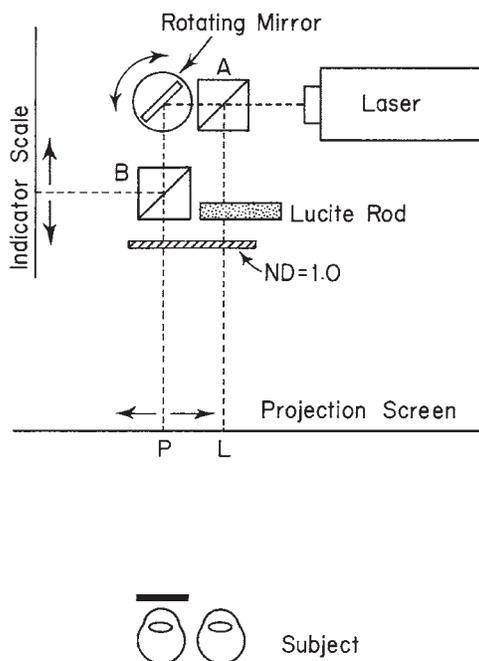


Figure 4. The laser device (viewed from above) used by subjects to indicate the amount of depth perceived in each line pair. The laser beam was split by the beam splitting cube at "A" into two beams. One of these was transmitted through a 0.95-cm-diameter Lucite rod which spread the beam vertically, resulting in a red line projected onto the screen at "L." The remaining beam was reflected from a front-surface mirror that rotated when the subject activated its motor with a joystick. This beam was again split (at "B"), one of the resulting beams being projected as a spot at "P," the other being projected onto an indicator scale visible only to the experimenter. When the mirror rotated, the spot at "P" moved laterally left or right (depending on the direction in which the subject moved the joystick). At the same time, another spot moved across the indicator scale, from which the experimenter could read the distance between "P" and "L."

Table 1
Perceived Depth Data (in Centimeters) for All Conditions
of Experiment 2

Texture	Luminance Ratio									
	0.25		0.50		1.00		1.50		2.00	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
	Left Line Flickered									
Present	1.35	1.98	0.75	2.64	1.03	2.49	1.23	3.09	0.30	2.59
Absent	0.60	2.95	1.29	1.77	1.43	2.81	-0.28	2.87	0.70	2.98
	Right Line Flickered									
Present	1.62	2.85	1.32	2.57	1.03	2.64	0.73	2.54	1.26	3.10
Absent	1.99	2.60	1.91	2.17	2.08	2.40	0.63	2.75	0.83	2.11

GENERAL DISCUSSION

The principal finding of this study is that of Experiment 1: subjects significantly underestimated the length of a flickering stimulus relative to a nonflickering stimulus of the same physical length. This illusion cannot be explained in terms of inappropriate constancy scaling, because the observed illusion was in the opposite direction to what would be predicted by this approach. If a flickering line is perceived as more distant than a nonflickering one (as confirmed by previous research and the results of Experiment 2), inappropriate constancy scaling would predict that a flickering line would be perceived as longer than a nonflickering line of equal length. This prediction was disconfirmed by Experiment 1.

In fairness, it must be acknowledged that the stimuli used in these experiments may have provided a test of inappropriate constancy scaling that was somewhat weaker than proponents of that position might desire. Although Experiment 2 did show that the nonflickering stimulus appeared to be closer than the flickering stimulus, observations made during pilot work and spontaneous comments from subjects showed that this depth effect did not have the phenomenal impact that the size effect of Experiment 1 had, in spite of its statistical significance. The projection screen was only 1 m from the eyes of the subjects, and it did not fill the entire visual field. Thus, there were a number of cues provided by the texture of the screen, and its edges, that supplied flatness information. Nevertheless, the fact remains that some depth, in the predicted direction, was observed, and the size effects obtained in Experiment 1 were directly opposite to what one would expect from inappropriate constancy scaling. Even if one argues that because of weak depth information there was no opportunity for inappropriate constancy scaling, one still has to explain the fact that there was an illusion. Clearly, inappropriate constancy scaling cannot account for it.

An alternative is the mirror image of the inappropriate constancy scaling argument. Perhaps it is not perceived depth that leads to perceived size, but rather perceived size that leads to perceived depth. It may be that the presence of flicker in one of two lines of equal physical length leads that line to be represented, somewhere in

the perceptual sequence, as shorter than the nonflickering line, much as if it had a smaller retinal image. Since in the real world objects of equal physical size but different retinal image size lie at different distances from the viewer, the subject perceives the flickering line as more distant than the nonflickering line. The advantage of this approach is that it makes the findings of Experiment 1 and Experiment 2 consistent with one another. The disadvantage is that it opens the same chicken-egg problem that has plagued size-depth scaling explanations of illusions for years, as in the case of the well-known moon illusion (see, e.g., Coren, 1989; Kaufman & Rock, 1962, 1989; McCready, 1986; Reed, 1989). Which comes first, size perception or depth perception?

In the present case it probably does not matter, because it is unlikely that the explanation described in the previous paragraph applies. The explanation is based on the assumption that two stimuli whose depths are being perceived must be known, at some level, to have the same physical size. However, past experiments have shown that of two stimuli, one flickering and the other not, the one that is flickering will be perceived as more distant even when the two stimuli are obviously of very different physical size (and shape). Wong and Weisstein (1987), for example, used a form of Rubin's reversible faces/goblet picture. On some trials, the inner "goblet" flickered, and on other trials, the outer "faces" flickered. The flickering area was always perceived as farther than the nonflickering area. This result occurred even in the absence of contours defining the different areas. Similar results were found when the stimulus was a diagonal band across a square; the stimulus was perceived as a diagonal band in front of a square background or as two salient triangles standing out in front of a square background, depending on whether the diagonal band or the triangles flickered. Wong and Weisstein (1985) found similar results when they used concentric squares, with either the inner or the outer square being perceived as closer, depending on which one was flickering. It seems clear from these findings that flicker directly produces the perceived depth effects, unmediated by size perceptions.

Even if the preceding explanation for perceived depth were correct, however, it would still be predicated on the idea that flicker affects size perception. It would leave open the explanation of how flicker exerts this effect. Neither inappropriate constancy scaling as originally described, nor its mirror-image alternative discussed above, provides an explanation for flicker-mediated effects on size perception. However, while the size illusion shown in Experiment 1 is difficult to explain by appealing to relations between perceived size and perceived distance, it is consistent with expectations based on motion processing, as discussed in the introduction. That is, we propose that we have revealed a flicker analogue of the Ansbacher effect.

Just as a rapidly moving stimulus appears to be smaller than a slower one (the Ansbacher effect), a flickering stimulus appears to be smaller than a nonflickering one. This

observation is what one would expect if flicker and motion were mediated by a common pathway (e.g., magnocellular processing stream). The on-off flicker of our stimuli would produce broadband energy across a wide range of spatiotemporal frequencies. This effect is similar in concept to the on-off flicker of sine wave gratings that may be decomposed into two oppositely moving gratings plus a stationary pattern (Sekuler, Pantle, & Levinson, 1978). It is reasonable to suppose that motion processes would be engaged by our flickering stimulus, albeit with no overall net energy in one direction or the other.

Finally, our results suggest that perceived size may be influenced by activation of more than one pathway. That is, there may be one pathway that involves depth perception (and concomitant constancy scaling), and another, different pathway separate from size constancy. An extension of this idea assumes that size constancy is not operative when the flicker/motion pathway is activated, an interpretation consistent with a study reported by Rogowitz (1984), who demonstrated that size constancy deteriorates when stimuli are illuminated stroboscopically. Rogowitz proposed that size constancy is processed only by sustained channels, and that under transient stimulation size constancy breaks down. Interestingly, the effects shown in the present Experiment 1 occurred at a flicker rate of 6 Hz, a value similar to the peak sensitivity of transient mechanisms (see, e.g., Kulikowski & Tolhurst, 1973). Thus, the size illusion produced by flicker may be mediated by a direct effect of transient system operation rather than by an indirect effect of flicker operating as a depth cue in the service of perceptual constancy.

REFERENCES

- ANSBACHER, H. L. (1944). Distortion in the perception of real movement. *Journal of Experimental Psychology*, **34**, 1-23.
- ANSTIS, S. (1986). Motion perception in the frontal plane. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance: Vol. 1. Sensory processes and perception* (pp. 16.1-16.27). New York: Wiley.
- BAKER, C. L. (1990). Spatial- and temporal-frequency selectivity as a basis for velocity preference in cat striate cortex neurons. *Visual Neuroscience*, **4**, 101-114.
- BARTLEY, S. H. (1969). *Principles of perception* (2nd ed.). New York: Harper & Row.
- BHATIA, B., & VERGHESE, C. A. (1963). Constancy of the visibility of a moving object viewed from different distances with the eyes fixed. *Journal of the Optical Society of America*, **53**, 283-286.
- BHATIA, B., & VERGHESE, C. A. (1964). Threshold size of a moving object as a function of its speed. *Journal of the Optical Society of America*, **54**, 948-950.
- BORRESEN, C. R. (1990). Illusion reversal rate as a function of subjective depth. *Journal of General Psychology*, **117**, 39-46.
- BUSER, P., & IMBERT, M. (1992). *Vision* (R. H. Ray, Trans.). Cambridge, MA: MIT Press. (Original work published 1987)
- CAELLI, T., HOFFMAN, W., & LINDMAN, H. (1978). Subjective Lorentz transformations and the perception of motion. *Journal of the Optical Society of America*, **68**, 402-410.
- COREN, S. (1989). The many moon illusions: An integration through analysis. In M. Hershenson (Ed.), *The moon illusion* (pp. 351-370). Hillsdale, NJ: Erlbaum.
- COREN, S., & GIRGUS, J. S. (1978). *Seeing is deceiving: The psychology of visual illusions*. Hillsdale, NJ: Erlbaum.
- COREN, S., PORAC, C., AKS, D. J., & MORIKAWA, K. (1988). A method to assess the relative contribution of lateral inhibition to the magnitude of visual-geometric illusions. *Perception & Psychophysics*, **43**, 551-558.
- ENGEN, T. (1972). Psychophysics: I. Discrimination and detection. In J. W. Kling & L. A. Riggs (Eds.), *Woodworth and Schlosberg's Experimental psychology: Vol. 1. Sensation and perception* (3rd ed., pp. 11-46). New York: Holt, Rinehart & Winston.
- FESTINGER, L., COREN, S., & RIVERS, G. (1970). The effects of attention on brightness contrast and assimilation. *American Journal of Psychology*, **83**, 189-207.
- FISHER, G. H. (1968). An experimental and theoretical appraisal of the inappropriate size-depth theories of illusions. *British Journal of Psychology*, **59**, 373-383.
- FRY, G. A., BRIDGMAN, C. S., & ELLERBROCK, V. J. (1949). The effects of atmospheric scattering on binocular depth perception. *American Journal of Optometry & Archives of the American Academy of Optometry*, **26**, 9-15.
- GILLAM, B. (1980, January). Geometrical illusions. *Scientific American*, **242**(1), 102-111.
- GRAHAM, C. H. (1929). Area, color, and brightness difference in a reversible configuration. *Journal of General Psychology*, **2**, 470-481.
- GREEN, M. (1981). Psychophysical relationships among mechanisms sensitive to pattern flicker and motion. *Vision Research*, **21**, 971-984.
- GREGORY, R. L. (1963). Distortion of visual space as inappropriate constancy scaling. *Nature*, **199**, 678-680.
- GREGORY, R. L. (1966). Visual illusions. In B. M. Foss (Ed.), *New horizons in psychology* (pp. 68-96). Baltimore: Penguin Books.
- GREGORY, R. L. (1967). Comments on the inappropriate constancy scaling theory of the illusions and its implications. *Quarterly Journal of Experimental Psychology*, **19**, 219-223.
- GREGORY, R. L. (1968, November). Visual illusions. *Scientific American*, **219**(5), 66-76.
- GREGORY, R. L. (1990). *Eye and brain* (4th ed.). Princeton, NJ: Princeton University Press.
- HEEGER, D. J. (1987). Model for the extraction of image flow. *Journal of the Optical Society of America A*, **4**, 1455-1471.
- ITTELSON, W. H. (1960). *Visual space perception*. New York: Springer-Verlag.
- KAUFMAN, L., & ROCK, I. (1962, June 15). The moon illusion, I. *Science*, **136**, 953-961.
- KAUFMAN, L., & ROCK, I. (1989). The moon illusion thirty years later. In M. Hershenson (Ed.), *The moon illusion* (pp. 193-234). Hillsdale, NJ: Erlbaum.
- 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.
- KULIKOWSKI, J. J., & TOLHURST, D. J. (1973). Psychophysical evidence for sustained and transient detectors in human vision. *Journal of Physiology*, **232**, 149-162.
- LEIBOWITZ, H., BRISLIN, R., PERLMUTTER, L., & HENNESSY, R. (1969, November 28). Ponzos perspective illusion as a manifestation of space perception. *Science*, **166**, 1174-1176.
- MACGOWAN, K. (1965). *Behind the screen: The history and techniques of the motion picture*. New York: Dell.
- MARSHALL, A. J., & STANLEY, G. (1964). The apparent length of light and dark arcs seen peripherally in rotary motion. *Australian Journal of Psychology*, **16**, 120-128.
- MCCREADY, D. (1986). Moon illusions redescribed. *Perception & Psychophysics*, **39**, 64-72.
- O'SHEA, R. P., BLACKBURN, S. G., & ONO, H. (1993, May). *Aerial perspective, contrast, and depth perception*. Paper presented at the meeting of the Association for Research in Vision and Ophthalmology, Sarasota, FL. Abstract published in *Investigative Ophthalmology & Visual Science*, **34**, 1185.
- OVER, R. (1968). Explanations of geometrical illusions. *Psychological Bulletin*, **70**, 545-562.
- OYAMA, T. (1960). Figure-ground dominance as a function of sector angle, brightness, hue, and orientation. *Journal of Experimental Psychology*, **60**, 299-305.

- PRESSEY, A. W., & EPP, D. (1992). Spatial attention in Ponzo-like patterns. *Perception & Psychophysics*, **52**, 211-221.
- REED, C. F. (1989). Terrestrial and celestial passage. In M. Hershen-son (Ed.), *The moon illusion* (pp. 267-280). Hillsdale, NJ: Erlbaum.
- RIGGS, L. A. (1972). Vision. In J. W. Kling & L. A. Riggs (Eds.), *Woodworth and Schlosberg's Experimental psychology: Vol. 1. Sensation and perception* (3rd ed., pp. 273-314). New York: Holt, Rinehart & Winston.
- ROBINSON, J. O. (1972). *The psychology of visual illusion*. London: Hutchinson.
- ROGOWITZ, B. E. (1984). The breakdown of size constancy under stroboscopic illumination. In L. Spillman & B. R. Wooten (Eds.), *Sensory experience, adaptation, and perception* (pp. 201-213). Hillsdale, NJ: Erlbaum.
- SCHILLER, P. H., LOGOTHETIS, N. K., & CHARLES, E. R. (1990). Role of the color-opponent and broad-band channels in vision. *Visual Neuroscience*, **5**, 321-346.
- SEKULER, R., PANTLE, A., & LEVINSON, E. (1978). Physiological basis of motion perception. In R. Held, H. W. Leibowitz, & H.-L. Teuber (Eds.), *Handbook of sensory physiology: Vol. 8. Perception* (pp. 67-96). New York: Springer-Verlag.
- STANLEY, G. (1970a). Static visual noise and the Ansbacher effect. *Quarterly Journal of Experimental Psychology*, **22**, 43-48.
- STANLEY, G. (1970b). Varying amount of static visual noise and the Ansbacher effect. *Psychonomic Science*, **20**, 343-344.
- TAUSCH, R. (1954). Optische Täuschungen als artifizielle Effekte der Gestaltungsprozesse von Größen- und Formenkonstanz in der natürlichen Raumwahrnehmung [Optical illusions as artifacts of the mechanisms of size and shape constancy in normal space perception]. *Psychologische Forschung*, **24**, 299-348.
- THOMPSON, P. (1982). Perceived rate of movement depends on contrast. *Vision Research*, **2**, 377-380.
- WATSON, A. B. (1986). Temporal sensitivity. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance: Vol. 1. Sensory processes and perception* (pp. 6.1-6.43). New York: Wiley.
- Wertheimer, M. (1912). Experimentelle Studien über das Sehen von Bewegung [Experimental studies of the seeing of movement]. *Zeitschrift für Psychologie*, **61**, 161-265.
- WONG, E., & WEISSTEIN, N. (1984a). 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. (1984b, May). *Flickering regions of a reversible picture are seen as backgrounds and non-flickering regions are seen as figures*. Paper presented at the meeting of the Association for Research in Vision and Ophthalmology, Sarasota, FL. Abstract published in *Investigative Ophthalmology & Visual Science*, **25**, 199.
- 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 January 12, 1994;
revision accepted for publication November 21, 1994.)