Evidence for disparity change as the primary stimulus for stereoscopic processing

BARBARA GILLAM and THOMAS FLAGG
Schnurmacher Institute for Vision Research, State University of New York, College of Optometry
New York, New York

and

DAVID FINLAY
University of Newcastle, Newcastle, NSW, Australia

Subjects were able to respond to a lens-induced stereoscopic slant more quickly and more accurately when it was imposed on only part of a surface rather than on the whole surface. This shows that the presence of a stereoscopic boundary, where disparity is discontinuous, increases the efficiency of stereoscopic processing. This finding is not consistent with many current models of stereopsis.

Recent models of stereoscopic depth perception (Grimson, 1980; Marr & Poggio, 1979; Mayhew & Frisby, 1980; Nelson, 1975) have been principally concerned with the problem of correspondence—the process by which the appropriate components of each eye's image are put together. It is generally assumed that once this match has been made the difference in retinal coordinates or "binocular disparity" of each pair of matched elements is known to the visual system. Depth is assumed to then be assigned throughout the field according to the disparities of all the matched elements. These models work best on surfaces, where configurations are similar in the two eyes and where disparity is homogeneous. They tend to have more difficulty (Mayhew & Frisby, 1980; Nelson, 1975) with the resolution of the boundaries of surfaces, where there are marked dissimilarities between corresponding configurations in the two eyes: specifically, this is true for configurations spanning surfaces at different depths where the disparities are discontinuous. In the present study, the possibility is explored that disparity discontinuities, such as those that occur at surface boundaries, may not be the most difficult parts of the images to resolve stereoscopically, but actually may facilitate the initiation and resolution of stereoscopic depth. Resolving disparity at boundaries would be an efficient strategy for the stereoscopic system, since disparities on the surface would then be redundant in resolving planar surfaces. Disparity processing across the surface would, however, be necessary to resolve undulations. Our data suggest two stages for stereopsis: a fast "boundary mode" based on local disparity discontinuities and a slower "surface

mode" based perhaps on absolute disparity or perhaps on the integration of small disparity variations.

The possible importance of depth boundaries in stereoscopic depth perception was suggested to us by investigations of the long latency associated with "aniseikonia," as reported by Ames (1946) and investigated by Seagrim (1967). Aniseikonia, or the state of "unequal images," is achieved by magnifying one eye's image. It is usually imposed on real scenes by placing over one eye an afocal cylindrical magnifier with the desired axis. If the magnification is in the horizontal meridian, a gradient of horizontal disparity across the field is created, which is consistent with a slant around a vertical axis. The degree of slant is related to the magnification by the following equation (Ogle, 1950):

$$\tan \theta = \frac{(M-1)}{2M} \cdot \frac{y}{a}, \qquad (1)$$

where $\tan \theta$ is the angle of slant in degrees, M is the magnification in percent, y is the observation distance, and a is half the interocular distance.

The long latency in seeing any depth effect when such a magnification is introduced is usually attributed to the fact that any monocular depth cues present are not consistent with the stereoscopic information for slant. It is known that when the stimulus has rectilinear features—thus providing perspective information—the stereoscopic slant, when it does occur, is usually less than geometrically predicted (Gillam, 1968).

The attenuated stereoscopic response associated with aniseikonia is ideal as a baseline for investigating factors that facilitate stereopsis or enhance the use made of the stereoscopic information present. Such a facilitation would be shown by either a decrease in the latency of the depth resolution or an increase in the degree of depth, or both. We observed informally that magnifying only part of the

binocular field seemed to greatly increase the response to the magnification. A slanted surface seemed to emerge instantly against the unslanted surface represented by the unmagnified part of the field. This suggested an important role for disparity discontinuities in resolving surface slant. The present paper describes an attempt to investigate this effect experimentally, using as the stimulus a regularly patterned frontal plane surface that provided perspective information as to its physical slant (see Figure 2). Aniseikonia imposed on the surface by means of lenses produced a disparity gradient consistent with slant away from the frontal plane around a central vertical axis. The use of the slant response has a unique advantage for investigating the role of stereoscopic boundaries in that it allows measurement of the stereoscopic effect without a boundary. This is, of course, not possible for frontal plane surfaces, although they are presumably processed in a similar way.

Figure 1 shows schematically the left- and right-eye views when a gradient of horizontal disparity is introduced over a whole field (1a) by magnifying all of the right eye's view and over part of the field (1b and 1c) by magnifying part of the right eye's view. In Figure 1b, the boundary of the slanted surface is a stereoscopic one with disparity discontinuities, whereas in Figure 1c it is not, since the unslanted part of the field is given only monocularly. (In Figure 1b, note the discrepancy in the configurations of the corresponding parts of the left and right images at the boundary where the disparity is discontinuous.) Since the disparities carrying the slant in the upper part of the images are identical in all three conditions, differences in the magnitude of stereoscopic slant perceived, and the speed with which it is perceived, can be attributed to the presence or absence of a boundary.
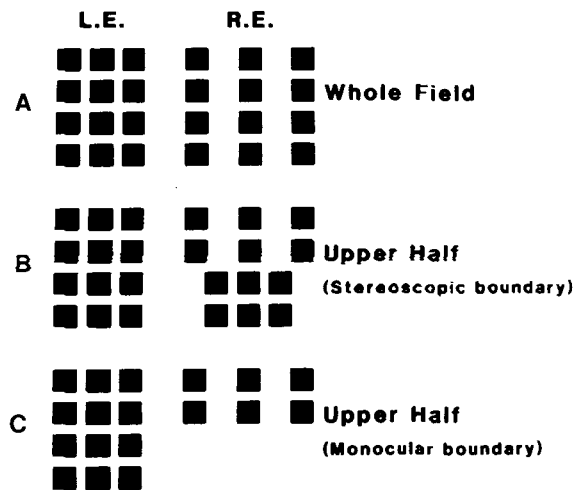

Figure 1. Schematic illustration of the three types of conditions used in the experiment. (a) Horizontal magnification of entire right eye's image (analogous to our whole-field condition). (b) Magnification of upper half of right-eye image with lower half visible but unmagnified (analogous to our stereoscopic boundary condition). (c) Magnification of upper half of right-eye image with lower half occluded for that eye (analogous to our monocular boundary condition).

METHOD

The three stereoscopic conditions are illustrated using a schematic stimulus in Figure 1. The analogues of these conditions were created by means of afocal horizontal magnifying lenses (Ogle, 1950) cut horizontally along their lower edges. By adjusting the height of the lens, a disparity gradient could be imposed on the whole field (Figure 1a), or on the upper half only (Figure 1b), with the lower half remaining stereoscopically in the frontal plane. The latter was called the "stereoscopic boundary" condition. Figure 1c illustrates a control condition in which the upper half was seen binocularly through the lens but the lower half was occluded for the right eye. This condition resembled the stereoscopic boundary condition in that it was consistent with a stereoscopically slanted plane seen against a frontal plane with a monocularly discernible boundary between them. However, in the control condition (known as the monocular boundary condition), the boundary between the depth planes was characterized by a change from stereoscopic to monocular information, rather than by disparity discontinuities.

The actual stimulus was a vertical grid pattern made up of lines of irregularly spaced dots drawn in India ink on a sheet of glass as illustrated in Figure 2. The glass was seen against a lightbox of luminance 1700 cd/m². Both the vertical and horizontal lines of the grid pattern were spaced 3.5° apart at the viewing distance used (56 cm). Dot spacing along the lines averaged two per degree of visual angle. The subject's view of the grid was restricted horizontally to 36° and vertically to 15°. Apertures restricting the view were placed close to the subject's eyes and were too out of focus to be salient or to constitute a stereoscopic boundary.

Lens magnifications were 5% and 8%, introducing disparity gradients equivalent to stereoscopic slants of 22° and 31°, respectively, at the observation distance used. Two naive subjects (an emmetrope, J.W., and a corrected myope, T.F.) and the senior author (an emmetrope) were used in the experiment.

The conditions for which measurements of degree and latency of slant were required (whole-field condition, stereoscopic boundary upper half, stereoscopic boundary lower half, and monocular boundary upper half) were each presented in combination with the two lens magnifications. Although there was no stimulus to slant for the stereoscopic boundary lower half field, its apparent slant was also measured in order to assess stereoscopic depth contrast, in the sense of repulsion of the two fields from each other. This would result in a deviation of the lower field from its true frontal plane position. J.W. was run before we introduced this condition, so data are reported on lower half slant for T.F. and B.G. only. Each of the eight conditions was presented twice in random order on each session, for a total of four sessions for J.W., five for T.F., and seven for B.G. We used Gillam's (1968) method of measuring slant, in which trained subjects used a remote control switch to rotate a monocularly viewed (therefore undistorted) comparison stimulus (a Meccano pulley wheel of 7.5-cm diameter—shown in Figure 2—which provides good monocular slant information) until it matched the apparent slant induced by the lenses. Training consisted of setting the wheel to match a real slanted surface (random-dot pattern) in which perspective and stereopsis were harmonious. The Meccano wheel was always set initially in the frontal plane, and the subjects were instructed to move it as necessary to keep it parallel at all times to the apparent slant of the designated surface. The importance of immediately registering any change in apparent slant was emphasized. Each trial lasted 1 min. The Meccano wheel was placed underneath and 7.5 cm in front of the grid pattern, and was seen against the same back-illuminated glass on which the pattern was drawn. The wheel was also illuminated from the front. A potentiometer attached to the vertical shaft of the comparison stimulus produced a resistance proportional to the degree of rotation of the shaft. A PDP-11/40 computer converted this resistance to a voltage and sampled it every second for the length of a trial (60 sec). The voltage was calibrated with respect to the angle of turn at the beginning of each experimental session. The
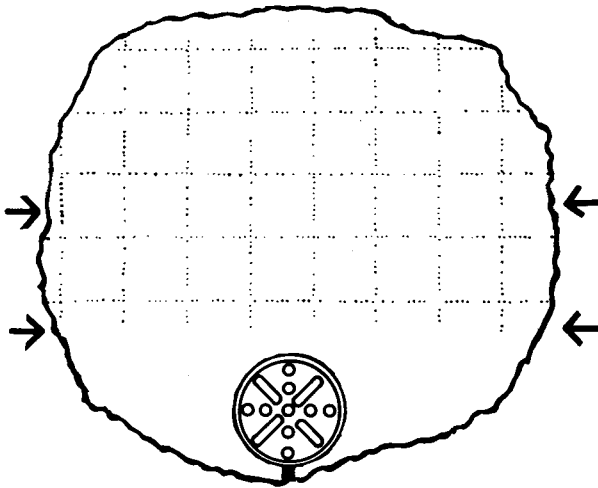
Figure 2. A scale drawing of the subject's view of actual stimulus used in the experiment. The upper arrows indicate where the bottom half of the lens intersected the field in the stereoscopic boundary condition, and the lower arrows show where it intersected the field in the whole-field condition. The lower arrows also show where the monocular occluder intersected the field for both the stereoscopic boundary and whole-field conditions. For the monocular boundary condition, the monocular occluder intersected the field at the upper arrows. The monocularly viewed comparison wheel was actually 7.5 cm in front of the glass on which the grid pattern was drawn (see text). Dotted lines were used to ensure that the stereoscopic process had adequate stimulation.

computer produced a graph showing the position of the comparison stimulus throughout the 1-min trial.

Latency was defined as the time taken for a subject to reach half the final setting when the final setting was at least 5°. Despite the latter restriction, all latencies plotted are based on at least six measurements per point. The use of half the final setting gives a much less arbitrary measure of latency than the final setting itself. The subjects typically make small adjustments to their final setting after quite long intervals, which might add considerably to the latency for little perceptual effect. However, since subjects usually proceed fairly smoothly to the region of their final setting, the time to reach half the final setting is not much affected by such adjustments at the end. For example, a fine adjustment by 2° to a slant setting of 30° might add as much as 10 sec to the latency were time to final setting used as the criterion, whereas if time to reach half the final setting were used instead, it would add only the time to traverse 1° in the middle, fast-changing region of the setting of the comparison stimulus (perhaps 500 msec).

RESULTS

For all three subjects the magnitude of perceived slant (final setting of the comparison stimulus) was much greater for the condition with a stereoscopic boundary than for the other two conditions.[1] A separate analysis of variance was run for each subject. In all three cases, the main effect of viewing condition was highly significant [$F(2,26)$ = 95.65, $F(2,18)$ = 18.51, and $F(2,14)$ = 12.81 for B.G., T.F., and J.W., respectively, $p < .001$ in all cases]. Data are shown in Figure 3. The bars represent standard errors of the mean. B.G.'s data demonstrated an even greater enhancing effect of the stereoscopic boundary than did J.W.'s and T.F.'s, because her

stereoscopic response to the control condition was even more attenuated than those of the other two subjects. A contrast effect is demonstrated by the fact that the lower half in the stereoscopic boundary condition appeared slanted in a direction opposite to the slant of the upper half. It is unlikely that this was an artifact, since there was no systematic bias in the same two subjects' settings in the practice trials. The contrast effect was small compared with the effect of the stereoscopic boundary on the degree of slant of the upper half field and is therefore unlikely to account for it.[2]

Latencies are shown for the same four conditions for two subjects, J.W. and B.G., in Figure 4. (T.F. found it difficult to make settings without bracketing, and his latencies therefore were not easily measurable by our criterion.) Separate analyses of variance were carried out for J.W. [$F(2,12)$ = 10.35, $p < .005$] and B.G. [$F(2,20)$ = 87.6, $p < .001$]. Stereoscopic slant of the upper half in the stereoscopic boundary condition occurred approximately two to four times as quickly as in the monocular boundary or the whole field conditions. The actual difference was even greater, since we did not measure, and eliminated from the analysis, latencies greater than 1 min. These all occurred for the whole-field and monocular boundary conditions. The large difference in favor of the stereoscopic boundary condition also occurred despite the fact that slants were greater in this condition and tracking them might therefore have been expected to take more time. The subjects reported no difficulty in fusing the patterns during the latency period. The delay was only in perceiving the stereoscopic slant.

DISCUSSION

There was a considerable increase in magnitude of stereoscopic slant (in the direction of geometric prediction) and a decrease in the latency of seeing the slant when the slanted surface was bordered by disparity discontinuities. In accordance with our preliminary observations, the degree of slant was found to be close to geometric prediction when the stereoscopic boundary was present and markedly attenuated (in accordance with previous findings on aniseikonia) when it was not. This is in line with the idea that slant was resolved at the boundary, since surface perspective could play little role there. On the other hand, when slant resolution is necessarily distributed over the surface, as in the whole-field conditions, conflicting surface perspective enters into the resolution of slant and attenuates the stereoscopic effect.

Our results indicate that stereopsis is not, at least initially, achieved by responding to the absolute disparity value of each part of the field. Large differences in disparity for adjacent elements seem to be critical. This supports the findings of Gogel (1956, 1972) and Gulick and Lawson (1976), that the depth response is to relative disparity and depends on local depth differences. Our results indicate an important role for such local depth differences in surface perception, namely, that the resolution of local depth at the boundary of a surface can extend to de-
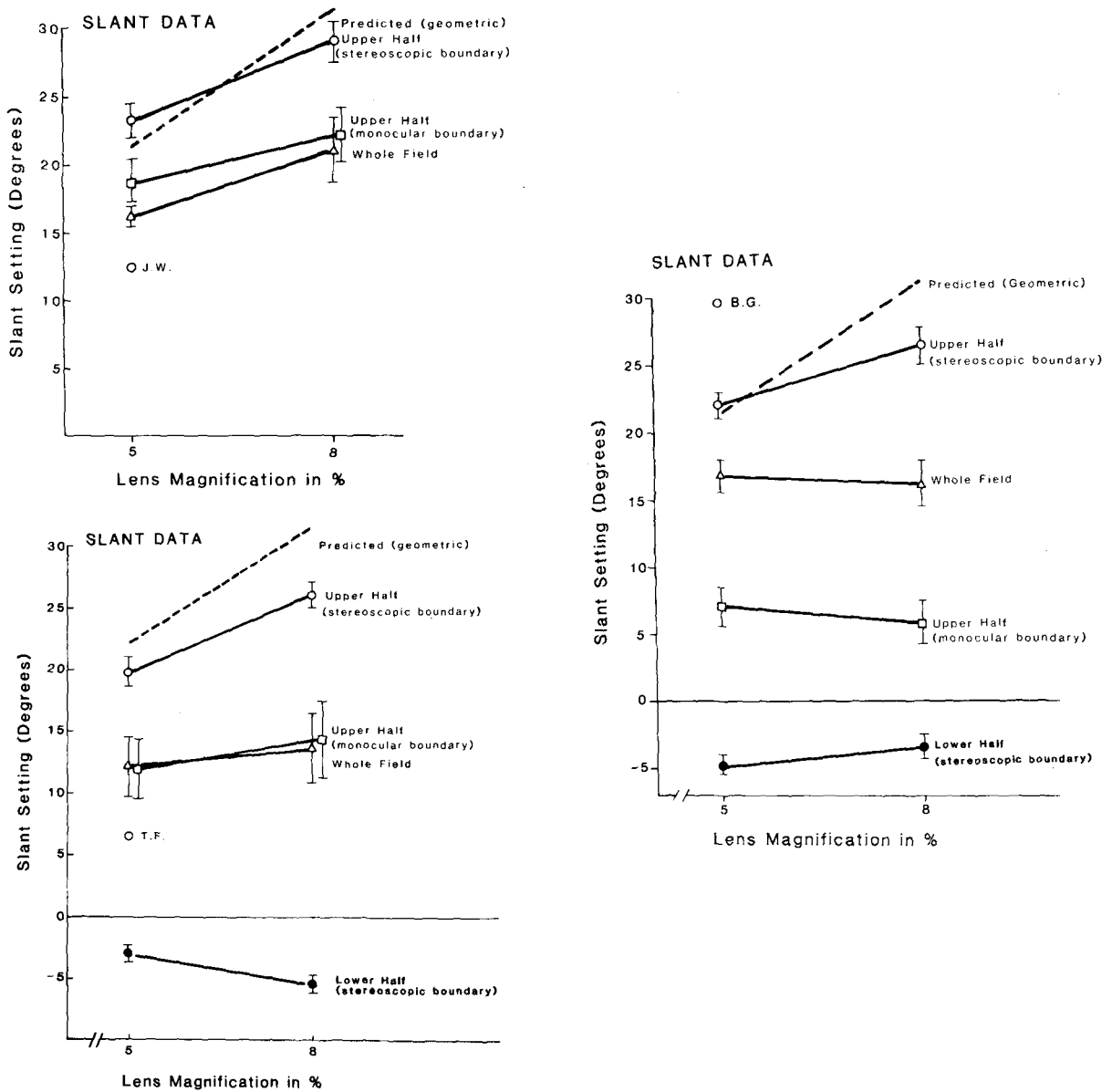
Figure 3. Slant data for three subjects. J.W. did not do lower half measurements.

tails across the surface. This could be done either by extrapolation from the boundary resolution or by the facilitation of response to the surface disparities. This result in the slant domain is clearly related to the finding of a depth analogue to the Craik-O'Brien-Cornsweet effect (Anstis & Rogers, 1978).

The importance of relative disparity in the initial stages of stereoscopic processing, as shown by our results, indicates that the older way of defining disparity as a *difference* in angles subtended at the two eyes by two points in space (Ogle, 1950) may be closer to reflecting how the stereoscopic system works than the currently popular definition in terms of a difference in the retinal coordinates of a single element.[3] The former definition also has the advantage of providing a disparity measure that remains

invariant over convergence change and therefore is more suitable for encoding. (The difference remains the same regardless of what the absolute disparities are.)

Another way of looking at our findings is to emphasize that greater stereoscopic efficiency is achieved by adding strongly mismatching components to the two images. These mismatching regions demarcate the parts of the array containing the minimum information required to resolve the depth of the entire array, namely, the surface boundaries. Initial stereoscopic processing appears to be located at the mismatching parts of the images. The implication is that although matching must be necessary for fusion and stereopsis and perhaps sufficient for fusion, it is not sufficient for stereopsis. If it were, our subjects should have had no difficulty with the whole-field condi-
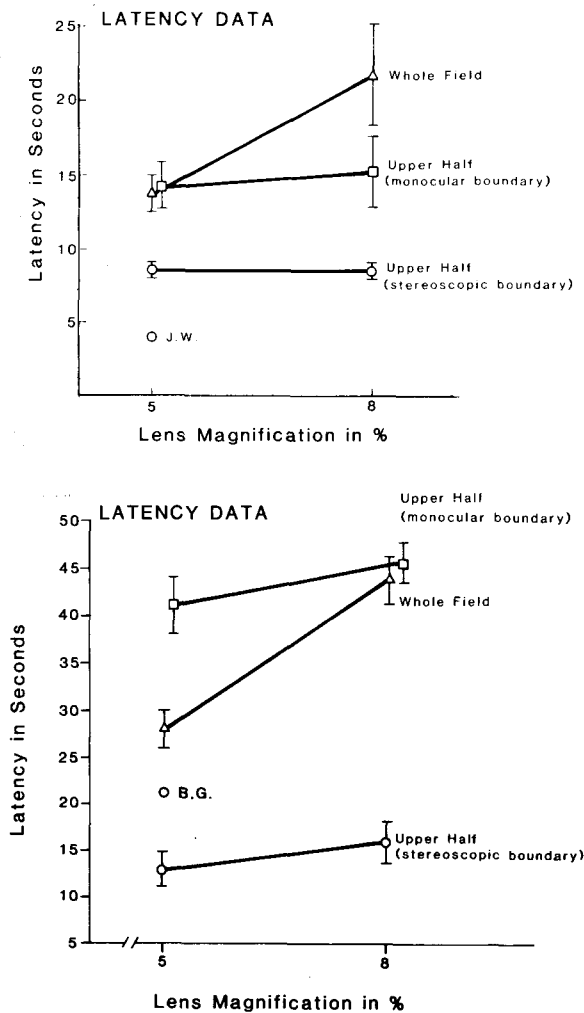
Figure 4. Latency data for two subjects.

tion. As stated earlier, all the patterns appeared fused immediately, regardless of when depth appeared. The important point is that solving the correspondence problem does not seem to solve the stereoscopic problem or render it a trivial one. Matching does not in itself result in the assignment of depth to the matched components.

It is at least logically possible that matching is a low-level process and does not involve reading out coordinate values (disparities) for each matched element into a buffer for individual depth assignment as has generally been assumed in computer models. This may be done only (initially at least) at the discontinuities. All other disparities are, after all, redundant for planar surfaces. It is necessary to keep in mind, however, that slant was ultimately resolved even for our whole-field conditions (even though attenuated), suggesting that a slower processing which does not involve discontinuities comes into play after the fast boundary mode has mapped out the depth field.

Our results suggest that the initial stages of stereoscopic processing are insensitive to changes in disparity occurring along the horizontal dimension (involving only differ-

ences in width of the two images) and to the implicit orientation disparities that always accompany differences in horizontal extent. The critical factor seems to be a change in *horizontal* disparity along the *vertical* dimension. This is a kind of orientation disparity in which vertically aligned image components in one eye correspond to nonvertically aligned components in the other eye. The importance of this factor was first suggested by Wallach and Bacon (1976).[4] They found that the latency for seeing depth was shorter for configurations with vertical misalignment (which they called "transverse disparity") than for configurations with only horizontal differences (which they called "superpositional disparity"). Our stereoscopic boundary condition was the only one that provided transverse disparity. Rogers and Graham's (1983) striking finding that the Craik-O'Brien-Cornsweet effect for stereopsis is considerably weaker in the horizontal dimension than in the vertical dimension also points to the ineffectiveness of a purely horizontal stretching of one image compared with a difference in vertical alignment of the corresponding elements.

Further research is needed on the properties of disparity discontinuities and configurational differences between the images that make these stimulus features particularly effective in producing a fast depth map of the visual field. The properties of the slower process, which is able to operate without such discontinuities, also require further investigation.

REFERENCES

AMES, A., JR. (1946). Binocular vision as affected by relations between uniocular stimulus-patterns in commonplace environments. *American Journal of Psychology*, **59**, 333-357.
ANSTIS, S. M., & ROGERS, B. (1978). A Craik-O'Brien-Cornsweet illusion for visual depth. *Vision Research*, **18**, 213-217.
BARLOW, H. B., BLAKEMORE, C., & PETTIGREW, J. D. (1967). The neural mechanism of binocular depth discrimination. *Journal of Physiology* (London), **193**, 327-342.
GILLAM, B. (1968). Perception of slant when stereopsis and perspective conflict: Experiments with aniseikonic lenses. *Journal of Experimental Psychology*, **72**, 299-305.
GILLAM, B. (1982). A boundary made of stereoscopic processing. *Investigative Ophthalmology* (Suppl.), **22**, 272.
GOGEL, W. C. (1956). Relative visual direction as a factor in relative distance perceptions. *Psychological Monographs*, **418**, 1-19.
GOGEL, W. C. (1972). Depth adjacency and cue effectiveness. *Journal of Experimental Psychology*, **92**, 176-181.
GRIMSON, W. E. L. (1980). *A computer implementation of a theory of human stereo vision*. (A.I. Memo No. 565). Cambridge, MA: M.I.T., Artificial Intelligence Laboratory.
GULICK, W. L., & LAWSON, R. B. (1976). *Human stereopsis: A psychophysical analysis*. New York: Oxford University Press.
MARR, D., & POGGIO, T. (1979). A computational theory of human stereo vision. *Proceedings of the Royal Society* (London), Series B, **204**, 301-328.
MAYHEW, J. E. W., & FRISBY, J. P. (1980). The computation of binocular edges. *Perception*, **9**, 69-86.
NELSON, J. I. (1975). Globality and stereoscopic fusion in binocular vision. *Journal of Theoretical Biology*, **49**, 1-80.
OGLE, K. N. (1950). *Researches in binocular vision*. New York: Hafner.
ROGERS, B., & GRAHAM, M. (1983). Anisotropies in the perception of three-dimensional surfaces. *Science*, **221**, 1409-1411.
SEAGRIM, G. N. (1967). Stereoscopic vision and aniseikonic lenses. I and II. *British Journal of Psychology*, **58**, 337-350 and 337-356.

564   GILLAM, FLAGG, AND FINLAY

SHIPLEY, T., & HYSON, M. (1972). The stereoscopic sense of order—
A classification of stereograms. *American Journal of Optometry and
Archives of the American Academy of Optometry*, **49**, 83-95.

WALLACH, H., & BACON, J. (1976). Two forms of retinal disparity.
Perception & Psychophysics, **19**, 375-382.

WALLACH, H., & LINDAUER, J. (1976). On the definition of retinal dis-
parity. In H. Wallach (Ed.), *On perception* (pp. 57-67). New York:
Times Books. (Reprinted from *Psychologische Beitrage*, 1962, **6**,
521-530.)

NOTES

1. After completing this work, we discovered a reference in an arti-
cle by Shipley and Hyson (1972) to the greater ease of seeing stereoscopic
slant when there is another different stereoscopic slant in the field. This
observation is attributed to Arthur Linksz. The authors also draw at-
tention to the distinction that can be made between continuous and dis-
crete disparity in constructing stereograms.

2. Contrast in the sense of repulsion cannot explain our data unless
it is assumed that a frontal plane has much more contrast effect on a
slanted plane than vice versa. Contrast in the sense of increased salience
of a visual resolution because of the presence of a reference cannot ex-
plain the data because a frontal plane surface, which also appeared to
be in the frontal plane, was also present in the monocular control
(monocular boundary) condition.

3. This definition can be traced to Hering and follows from his con-
cept of cover points. It has recently become popular because of the dis-
covery of "disparity detectors" which appear to record the binocular
coordinates of matched elements (Barlow, Blakemore, & Pettigrew,
1967).

4. Wallach and Lindauer suggested as early as 1962 that stereopsis
was based on differences in configurations in the two eyes rather than
on the disparity of individual elements.