Vision Res. Vol. 18, pp. 117 to 120. Pergamon Press 1978. Printed in Great Britain.

LETTER TO THE EDITORS

SINGLENESS OF VISION AND THE INITIAL APPEARANCE OF BINOCULAR DISPARITY[1]

(*Received 28 January 1976; in revised form 3 December 1976*)

Human perceptual response to static binocular disparity may be operationally defined in two ways. One can measure the resolution of the disparity system, typically in the form of a depth *discrimination* task. Such measurements have determined stereoacuity to be on the order of seconds of retinal arc (e.g. Berry, 1948). Alternatively, one can measure the failure of the same system to resolve disparity, in the form of singleness *judgements*. This failure has commonly been termed "fusion", and the disparity limits within which it occurs, "Panum's fusional area". Most theorists attribute both stereopsis and "fusion" to a central mechanism(s) which simultaneously translates horizontal disparities into the perceived third dimension, and alters the visual directions associated with disparate images in the interest of eliminating diplopia. In the horizontal dimension, these two aspects of binocular vision cannot be separated. In this case, any measurement of disparity resolution by a discrimination procedure is confounded by the presence of depth.

There are non-depth producing stimuli, however, which do allow such measurements to be made for the vertical dimensions of Panum's area, such as the cyclotorsional stimulus shown in Fig. 1.[2] Kaufman and Arditi (1976) recently showed that when a sensitive procedure is employed, torsionally disparate stimuli may be discriminated from non-disparate stimuli with such precision as to call into question the necessity of postulating a central fusional mechanism altogether.

Kertesz and Sullivan (1976) have criticized this experiment on two grounds. First, our 700 msec step presentation of the stimulus may not have provided sufficient time for "fusion" to occur. The authors cite the earlier experiment of Kertesz and Jones (1970) who found that when a large disparity stimulus ($\theta = 4$–5 deg, $d = 22'$–$28'$ in Fig. 1) was presented in a single step from an initial state of zero disparity ("noise" condition in Fig. 1), observers reported diplopia for 10–20 sec before "fusion" occurred. Thus they conclude that "fusion" may occur over an extended period of time. What is not clear, however, is whether "fusion" can be expected to occur less sluggishly with smaller disparities, and if so, how fast? Certainly if a central "fusion" mechanism is to be biologically useful, it ought to be more rapid than compensatory eye movement systems, since the latter can perform similar functions over time. Since the disparities used in the Kaufman and Arditi experiment were minute, it can hardly be expected that the "fusional" process would extend beyond the 700 msec presentation.

The second criticism which Kertesz and Sullivan direct against the Kaufman and Arditi experiment is that even if fusion were to occur within 700 msec, a discrimination could be made between disparate and non-disparate stimuli on the basis of the pre-"fused" initial appearance of the stimulus, regardless of whether or not fusion occurred later on in the stimulus interval. Referring again to the example cited above by Kertesz and Sullivan, an important difference should be noted between the stimulus used by Kertesz and Jones (1970) and that of Kaufman and Arditi (1976). The former authors presented an initially horizontal binocular line to which a disparity was suddenly introduced. This is a strong stimulus for apparent movement which may have enhanced disparity detection in the initial portion of the presentation. The latter authors presented only a fixation marker prior to the introduction of the disparity.

For those readers who have followed this controversy in *Vision Research*, we would like to present another experiment which supports the findings of Kaufman and Arditi, in its examination of the effects of an initial appearance cue on the discrimination of torsionally disparate from non-disparate stimuli. In order to avoid the complexities involved in the use of moving *or* apparently moving stimuli, luminance as a function of time will be the independent variable.

While little is presently known about the response of the visual system to contours in the threshold region, we will assume that if there is a transient response to disparity, it will be reduced or eliminated by the gradual introduction of the stimulus from a sub- to a supra-threshold level of brightness. Obviously, if the thresholds for contours are abrupt transitions from not seeing to seeing, this assumption would be incompatible. Our assumption, however, is not too far fetched in view of the current treatment of the concept of threshold.

METHOD

Apparatus

The experiment was run with the aid of a Digital Equipment Corporation PDP-12 computer. The stimuli, as

[1] This research was supported by Grant GB-36976 from the National Science Foundation. We thank Jean-Claude Falmagne for his suggestions about this research.

[2] In principle, stimuli with a constant vertical disparity could also be used instead of the cyclotorsional stimulus which contains a continuum of vertical disparities. However, such stimuli are powerful for producing compensatory disjunctive eye movements.

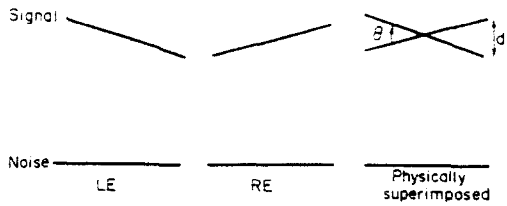

Fig. 1.

shown in Fig. 1, were generated as Lissajou patterns by applying identical in-phase triangular waveforms to the X- and Y-axes of two perpendicularly placed oscilloscopes (Tektronix 5103) whose images were superimposed on a half-silvered mirror. The half-images were segregated by crossed polarizing filters mounted on the oscilloscope faces and at the observer's eyes. The angle of tilt of each of the lines ($\theta/2$ in Fig. 1) was produced by means of a voltage level output device (under program control) and a multiplier, which affectively altered the gain of the signal applied to the Y-axes of the oscilloscopes. The gain of the signal applied to the X-axes was kept constant. The same output voltage was required to produce the opposed tilt in the lines of both half-images because the half-silvered mirror caused the reflected image to be reversed with respect to the non-reflected image. An opaque circular mask of 2° dia over each oscilloscope face delimited the stimulus lines.

Intensity of the stimuli was controlled with the aid of a D/A converter which applied a time varying d.c. voltage with a range of 1.1 V and a resolution of 96 steps under clock control to the Z-axes of both oscilloscopes.

A neutral density filter with a hole subtending 8′ was placed behind the polarizing filters at each oscilloscope face. This hole provided a bright central spot in each half-image which served as a fixation marker, and whose intensity was always 2 log units greater than the rest of each stimulus line. Between trials, then, the stimulus was physically present, but of subthreshold intensity; the brighter and visible fixation marker served the ancillary purpose of increasing the visibility threshold of the lines. All observers reported only the presence of the fixation spot between trials.

Luminance was measured with a MacBeth illuminometer over whose calibrated source was placed a combination of Wratten filters (Nos. 65 and 52). This provided a very close hue match with the light emitted from the oscilloscopes (P31 phosphor) in combination with the polarizing material. The transmission spectra of these filters was taken into account in all measurements.

Procedure

The trials called "signal" were those in which a disparity was present. For these stimuli, disparity may be defined either as the angle θ by which the eyes must counter-rotate to binocularly superimpose the lines, or as the maximum vertical disparity d which is present in the stimulus. The parameter θ is independent of the stimulus line lengths and viewing distance, while d varies inversely with viewing distance. Those called "noise" were simply binocularly superimposed horizontal lines. A session consisted of 100 trials, half of which were "signal" while the other half were "noise". The order of presentation was randomized in ten blocks of ten trials each, by means of a programmed pseudo-random number generator. The observer's task was to give a confidence rating as to whether a "signal" was presented or not. A response of "6" indicated maximal "signal" confidence, "1" indicated maximal "noise" confidence, while the intermediate integers indicated intermediate degrees of confidence.

Between trials, the experimental chamber was totally dark except for the fixation marker. With his head firmly placed in a headrest, the observer initiated each trial by means of a telegraph key, and communicated his response through an intercom. Feedback was given, in which a response of "3", "2" or "1" was interpreted as a "noise" response, while a "4", "5" or "6" was interpreted as a "signal" response.

In one condition (UP), the intensity of each half-image was gradually increased from a subthreshold level of 0.001 cd m^{-2} with luminance as a positive accelerated function of time as shown in Fig. 2. This relation exploited the non-linearity of the visual system to present both a large range of luminances and an approximately linear subjective brightness gradient. The objectively measured points shown in Fig. 2 describe luminances for only one half-image[3]. The total duration of this function was 13.968 sec, but the stimulus was visible only for a certain portion of this interval. All observers were asked to report when the stimulus became visible. The range of this threshold region was 4–6 sec after initiation of the stimulus interval. Thus the stimulus was visible for 8–10 sec. The slope within the threshold region was about 0.05 cd. m^{-2} sec^{-1}.

In the other condition (DOWN), the time-intensity relation was reversed, so that the total amount of luminous flux presented on a trial was kept constant over both conditions. The experiment consisted of one session of UP trials, followed by one session of DOWN trials.

All observers were given informal training with a range of disparities. This also allowed the experimenter to determine their approximate disparity detection thresholds. During the experiment, all observers were allowed unlimited rest breaks at all times.

Three observers were used, one of whom is an author of this paper. All had normal acuity and stereopsis.

Analysis

Discrimination was defined as a significant ($P < 0.05$) χ^2 statistic in which the data P("X"|S) was assumed to be distributed identically as P("X"|N) under the null hypothesis. In some cases, response categories were pooled

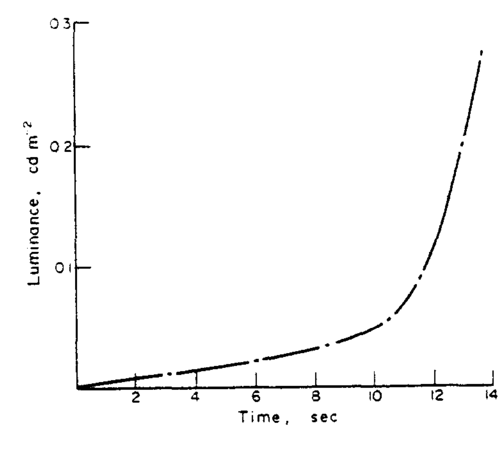

Fig. 2. Time–luminance relation. The points represent objectively measured luminances as described in the text.

[3] Before the experiment was run, the apparent brightness of the other half-image was adjusted to match the brightness of this standard for each observer. In this way, any difference in the sensitivities of the two eyes was compensated for.

in order to provide a minimum cell frequency of five. This test assumes nothing about the underlying shape of the hypothetical Signal and Signal + Noise distributions of Signal Detection Theory. Therefore, while the basic procedures of SDT are employed, the statistical conclusions drawn should be considered distribution- and parameter-free.

RESULTS AND DISCUSSION

ROC curves for the three observers are shown in Fig. 3. Observer SRM discriminated significantly with $P < 0.05$ in both the UP condition ($\chi^2 = 11.700$, d.f. = 3) and the DOWN condition ($\chi^2 = 27.095$, d.f. = 3). Observers JDH and ARA discriminated significantly in the UP condition ($\chi^2 = 13.580$, d.f. = 3, and $\chi^2 = 10.580$, d.f. = 4, respectively) but not in the DOWN condition.

Immediately worth noting about these curves is the wide variability in the observers' sensitivity to disparity. While observers SRM and JDH had sensitivity in accord with previous data (Kaufman and Arditi, 1976), observer ARA was able to discriminate much smaller disparities than has been previously reported. ARA, however, has had a great deal of experience with these stimuli, and has noted gradual improvement in his performance on this task over a period of more than a year. It is not inconceivable that one can learn to discriminate vertical or torsional disparities as finely as horizontal depth-producing disparities, given enough experience.

If the assumption stated in the introduction to this paper is correct, then the data of the UP condition alone are sufficient to show that the previous results of Kaufman and Arditi (1976) are not due to the initial appearance cue. This confirms the view that observers can discriminate torsionally disparate from non-disparate lines with sufficient sensitivity as to raise questions concerning the need for postulating a central fusional mechanism. For all observers the detected maximum vertical disparities at the end of the tilted lines (*d* in Fig. 1) were less than the "fusion" thresholds measured with less sensitive psychophysical procedures (see Kertesz and Jones, 1970; Kertesz, 1972, 1973). One observer's (JDH) threshold approached the limit of monocular resolution, and another observer's (ARA), was perhaps even smaller than this limit.

The DOWN condition was included to test the hypothesis that sensitivity would be greater than in the UP condition since the initial appearance cue is assumed to be lacking in the latter case. It should be noted that for two observers sensitivity was greater in the UP condition than in the DOWN condition. One observer (SRM) exhibited the reverse effect.

The reasons for the differences between the UP and DOWN conditions in observers ARA and JDH are obscure. A "dazzle" effect caused by the fact that the fixation spot in the DOWN condition was 2 log units brighter than the stimulus lines at onset might account for this result. However, if such an effect occurred, it did so only for two out of our three observers.

The data of observer SRM are supportive of the initial appearance hypothesis in that his performance was better in the DOWN condition than in the UP condition. However, it should be emphasized that the disparity which he successfully discriminated in both conditions was still about one half the smallest threshold reported by Kertesz (1973) using a stimulus of the same angular size as that used in the present experiment. Another explanation for the data of SRM which would account both for his requiring a relatively greater disparity and his better performance in the DOWN condition is that a certain amount of actual cyclorotation of the eyes occurred. In the UP condition these compensatory movements would occur as the large disparity stimulus was becoming visible and would cause the disparate lines to assume binocular registry over a similar time course. In the DOWN condition, the time course of counter-rotation would not coincide with the visibility of the lines in this way.

Kertesz and his associates (Kertesz and Jones, 1970; Kertesz, 1972, 1973) have proposed that cyclorotational eye movements do not occur with this kind of stimulus. In three studies they measured these movements objectively. In two of these (Kertesz, 1972, 1973), the stimulus was presented on a light box of 150 ft-L luminance, a very intense stimulus, and in the third study (Kertesz and Jones, 1970), where eye movements were measured with a less sensitive method, the 5 cd m^{-2} stimulus lines were presented on an oscilloscope screen. All of the eye movement techniques used involved apparatus which by necessity had to be located somewhere between the

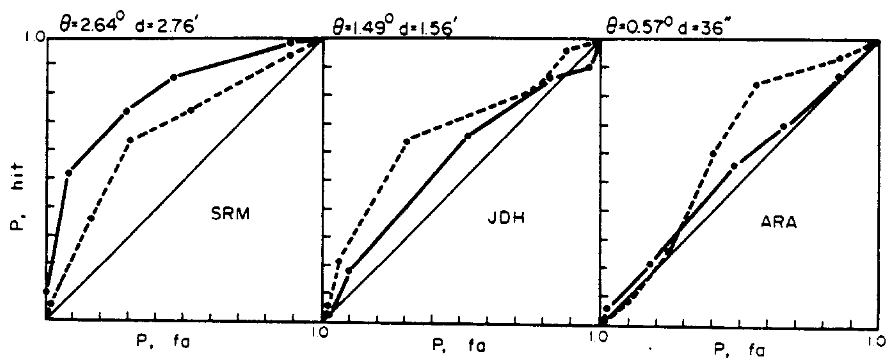

Fig. 3. ROC curves for the three observers. The dashed lines represent performance in the UP condition. The solid lines represent performance in the DOWN condition. Also shown are the torsional disparity θ and the maximum vertical disparity d, for each observer.

observer's eyes and the stimulus. While the stimulus display was unoccluded in both half-images, the rectangular outline of a Dove prism and possibly other objects were visible during these measurements (pers. comm., Kertesz, 1975). In the present experiment, on the other hand, the fixation point and the stimulus were the only objects visible in the field of view. It is possible that in the absence of all competing spatial cues, cyclorotation does occur, as was concluded by earlier investigators (Ogle and Ellerbrock, 1946; Ogle, 1950; Ellerbrock, 1954) who used more subjective eye movement measurements. Recently Kertesz (1976, personal communication) found that under certain conditions cyclorotational eye movements do occur.

Also it should be noted that since the DOWN condition followed the UP condition a certain amount of learning might have occurred for SRM.

In conclusion, the initial appearance of a flashed binocular disparity cannot account for the earlier results of Kaufman and Arditi. It is clear from this experiment that observers can discriminate very small disparities from the absence of disparity whether or not this cue is available.

New York University
Department of Psychology
New York, NY 10003, U.S.A.

ARIES ARDITI
LLOYD KAUFMAN

REFERENCES

Berry R. N. (1948) Quantitative relations among vernier, real depth, and stereoscopic depth acuities. *J. Exp. Psychol.* **38**, 708–721.

Ellerbrock V. J. (1954) Inducement of cyclofusional movements. *Am. J. Optom.* **31**, 553–566.

Kaufman L. and Arditi A. (1976) The fusion illusion. *Vision Res.* **16**, 535–543.

Kertesz A. E. (1972) The effect of stimulus complexity on human cyclofusional response. *Vision Res.* **12**, 699–704.

Kertesz A. E. (1973) Disparity detection within Panum's fusional areas. *Vision Res.* **13**, 1537–1543.

Kertesz A. E. and Jones R. W. (1970) Human cyclofusional response. *Vision Res.* **10**, 891–896.

Kertesz A. E. and Sullivan M. J. (1976) Fusion prevails: a reply to Kaufman and Arditi. *Vision Res.* **16**, 545–549.

Ogle K. N. (1950) In *Researches in Binocular Vision*. Saunders, Philadelphia.

Ogle K. N. and Ellerbrock V. J. (1946) Cyclofusional movements. *Arch. Ophthal.* **36**, 700–735.