Vision Res. Vol. 24, No. 9, pp. 1063–1073, 1984
Printed in Great Britain.

0042-6989/84 $3.00 + 0.00

THE PERCEPTION OF DEPTH IN SIMPLE FIGURES

G. J. MITCHISON* and G. WESTHEIMER
Department of Physiology-Anatomy, University of California, Berkeley, CA 94720, U.S.A.

(*Received* 29 *March* 1983; *in revised form* 8 *December* 1983)

Abstract—When subjects with good stereoscopic acuity are given the task of judging which of two vertical lines lies nearer, the presence of other features nearby alters the perceived depth within the test pair. In the presence of a single flanking line shown with disparity, the test line pair is seen as fronto-parallel when it has disparity in the direction which tends to align it in depth with the flanking line. The notion of "salience" is introduced. This is the summed disparity—weighted approximately inversely with distance—between test objects and their neighbours. We make the hypothesis that objects appear at equal depths when they have equal salience. The salience hypothesis accounts for a variety of depth interaction effects between test lines and adjoining features, such as one or more other lines and a lattice of dots with a disparity gradient. Whether features other than nearest neighbours influence depth judgements depends on the individual. For five good stereo subjects, in two a single line completely masked all effects beyond the nearest neighbour, two others had partial masking, and one had none. If the visual system is interested in corners between planes in depth and in objects protruding from such planes, then salience constitutes a useful indicator for this purpose.

Stereopsis Salience

INTRODUCTION

When an observer is shown a simple three-dimensional figure made of points and line segments and is given no other depth clues, he will use binocular disparity to assign relative depth to the figure's components. This is the basis of stereogram design, discovered by Wheatstone in 1838. As a simple example, consider the case of two lines. If they do not lie at the same distance from the observer, they will subtend different angles at the two eyes (Fig. 1). The visual system is able to use small differences in the relative positions of the images on the retinae to estimate relative depth.

On being presented with a pair of lines in isolation, a subject will generally judge them to be equally far when their physical distances are indeed roughly equal; in other words, when they both lie in the same fronto-parallel plane. Suppose, however, that there are other lines or points present in the visual field. We show here that a subject's perception of equal depth is then affected, to an extent which may be very large, especially when the other objects are near to the test pair.

We can explain this effect by saying that the visual system does not directly use the disparity between two test objects to judge depth. Instead, the judgement is based on what we call the "salience" of these objects, a number derived from disparity differences, which can be calculated for every object in the visual field. Loosely speaking, salience indicates the extent to which an object lies in front of or behind its neighbours. We shall show that salience can account with reasonable accuracy for the perceived depths of test lines in a variety of simple figures.

*Kenneth Craik Laboratory, Department of Physiology, Cambridge CB2 3EG, England, and M.R.C. Laboratory of Molecular Biology, Hills Road, Cambridge. Correspondence should be sent to the former address.

METHODS

A convenient way of measuring stereoscopic performance is to show a subject a test pair of lines repeatedly, each time with a disparity selected at random from an array of seven possible values centered around zero disparity. The subject is asked to make a binary decision as to whether the right or the left line appears nearer. A psychometric curve fitted to the data [Fig. 1(b)] is then analyzed by the method of probits (Finney, 1952) and yields two relevant parameters. One, the slope of the curve, measures the subject's stereoscopic sensitivity by giving a threshold, i.e., half the distance between the disparities at which the subject responds correctly on 75% of occasions when either the right or the left line is nearer. The other is the 50% point on the psychometric curve, the mean, which identifies the relative target positions for which the two lines appear to the subject to be in the fronto-parallel plane. Ordinarily, the mean of the curve will be quite close to zero disparity. This kind of veridicality of perception can be helped by providing error feedback, i.e. by alerting the subject whenever the response is "behind" for a nearer target or vice versa. In the experiments here described we employed normal subjects viewing patterns in a perceptually bland milieu; even when we did not furnish error feedback, the mean of the psychometric curve in the basic test situation was almost always quite close to zero disparity.

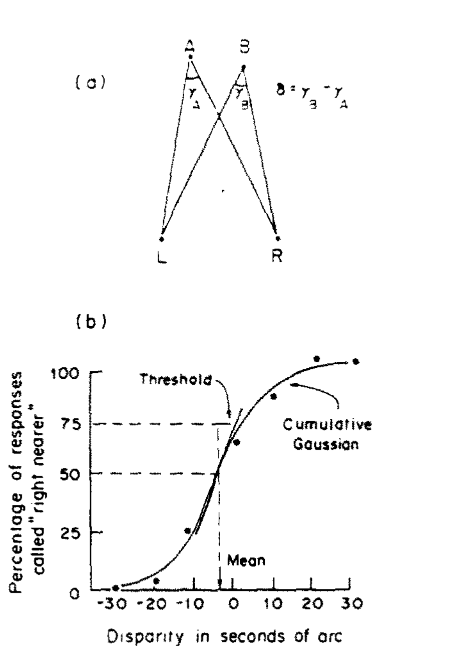


Fig. 1. (a) A geometrical schema of the simplest stereoscopic depth discrimination experiment. Horizontal section containing both eyes and two vertical line targets A and B. The angle δ, the difference between the angles subtended by the eyes at the two targets, is the binocular disparity. Note that, when δ is positive, then B is nearer than A in distance. Thus disparity varies negatively with distance. In fact, a simple calculation shows that, near to the fixation plane, $\text{dist}_B - \text{dist}_A \cong -(d^2/w)\delta$, where d = distance to the fixation plane and w is the interocular separation. (b) Probit method of analyzing data. The percentage of "yes" responses to the question, "Does the right line appear nearer than the left?" is plotted for each of the seven disparities which make up the test ensemble and are used in random order for a total of usually 300 responses. A cumulative Gaussian curve is fitted to the data; the subject's sensitivity is indicated by the slope of the curve at its inflection point and the 50% "yes" response signifies the disparity for fronto-parallel appearance of the line pair. Good subjects, such as those used in this study, have sensitivity of 10 sec of arc or less.

In many of the experiments described here, the determination of thresholds and means was far more difficult than in a conventional measurement for a pair of lines. Subjects sometimes showed considerable variability, both over periods of hours, and sometimes over periods of months, and the only way to surmount this was to collect many responses systematically.

All patterns were generated on two Tektronix 602 cathode-ray display units with short-acting white phosphor (P4), placed 3 m in front of the subject. By means of a beam-splitting pellicle and suitably oriented polaroid filters, it was arranged that the two display units were in accurate register, but one was visible to only the right eye and the other only to the left. Patterns were made up of dots about 0.5 min of arc of visual angle in diameter. Luminance was equivalent to about 40 mL, seen against a uniform background of about 1 mL. When it was intended for the subject to see a continuous line, the dots were aligned in a row 1 min of arc apart. All patterns were restricted to a square area of 1.5 deg side-length centered on the fixation point. No significant features were visible within several degrees of the central area, and the rest of the room was almost dark. The subjects, all of whom had considerable experience and facility in this kind of experiment, observed the screens with natural pupils and wore the required spectacle correction, if any.

Stability of oculo-motor adjustment was assured by a fixation pattern consisting of a small (5 min of arc) central cross and four brackets of similar size outlining a square 45 min of arc side length. It was shown continuously except during the presentation of the stimulus pattern, which lasted for 500 msec and occurred every 2.5 sec.

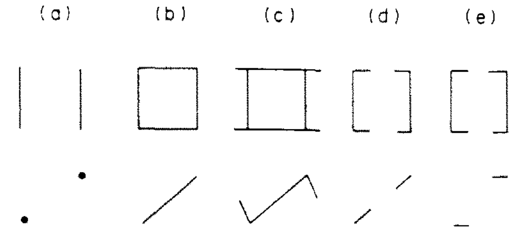
Each stimulus contained some particular aspect to which the subject was trained to attend and respond. Most often it was a vertical line pair, 20 or 12 min of arc apart and 10 or 20 min of arc high, whose members had a different binocular disparity during each presentation. The disparities used formed an equally spaced set of seven, used in random order. The subject had to make a binary judgement after each presentation, setting a switch, for example, according to whether the right or the left member of the test line pair appeared nearer. Results were accumulated in sessions of 200–400 responses, often several conditions being run in a randomly interdigitated fashion without error signals. Thus the subject was always unaware of the "correct" switch setting for any presentation.

The mean and slope of all psychometric curves used for analysis were based on at least 300 responses for each situation and sometimes even as many as 1200. The only exception to this is the data for the experiments in Fig. 5(d) and (e) for S.M., which were based on 200 responses. S.M. is a very reliable subject who shows little variability. The values of the thresholds then had a standard error of at most 10%. The means of the psychometric curves, i.e. of settings for fronto-parallelism, usually had a standard error of 1–2 sec of arc, except on those occasions when our experiments deliberately made such settings difficult or impossible.

We have chosen to retain the units of visual angle throughout, because these can be most readily related to ocular structures, and to other studies of the visual pathways. Conversion to equivalent geometrical relationships in the observer's object space requires knowledge of the interocular distance (about 65 mm in an average subject) and observation distance. For example, we may find that two vertical lines 20 min apart are just perceptibly out of the fronto-parallel plane when they have a disparity of 10 sec of arc. To simulate this situation, two real vertical lines at our observation distance of 3 m would have to be separated by 17.5 mm horizontally and 6.5 mm in the antero-posterior direction. The plane they would define would be inclined 21 deg to the fronto-parallel plane through the fixation point. We rarely used disparities exceeding 1 min of arc. It follows that the corresponding real objects, were they all of the same physical dimensions, would differ by less than 1% from each other in the visual angle they subtended, an amount that is actually well below threshold for size difference detection. Because the two targets would then be at unequal distances from the two eyes, there would also be some disparity in the vertical direction, but it would be less than 0.22 sec of arc in all experiments described here. These calculations show that, in the situations considered here, horizontal binocular disparity is so much more readily detected than other spatial properties of objects, that it is quite feasible to study its properties in isolation from competing clues to space perception.

RESULTS

We take as our point of departure the observation reported by one of us (Westheimer, 1979) that there is a considerable difference in disparity discrimination between two patterns containing identical stereo cues: a pair of vertical lines and a square made of the same vertical lines but now connected by two horizontal lines. Every subject examined in this laboratory exhibits a higher disparity threshold when the pattern is a square rather than just two lines [Fig. 2(a) and (b)]. The finding was, in fact, reported as early as 1937 by Werner. One subject, S.M., who has exceedingly good stereoscopic acuity, has such a high stereo threshold for a square target that it can hardly be measured. She has recently investigated this phenomenon in more detail (McKee, 1983), and has suggested that the deterioration in threshold occurs when features of different disparities are connected by parts of a figure.



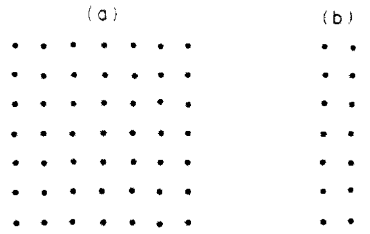
	(a)	(b)	(c)	(d)	(e)
SM	6.0 ±0 7	∞	16.3±16	99±13	6.0±0 7
GW	9.4 ±1.1	14.3±1 5	148±1 0	158±14	12 5±1 2
MG	8.9 ±0 9	196±1.9	11.3±1 2	14.1 ±0.9	152±16
RY	5.9 ±1 3	170 ±1 9	65±04	188 ±1.9	9.9±1 0

Fig. 2. A variety of patterns all sharing the basic component of two vertical lines 20 min of arc apart and 20 min high. The top row shows how they appear on the oscilloscope screen. Below is a view from above, representing them in depth. The four subjects' disparity thresholds (in sec of arc) for the various figures are listed underneath. (a) A pair of lines, showing low thresholds. (b) The square, showing elevated thresholds, especially for S.M. (c) The square, with its horizontals extended, their endpoints being given the disparity of the fixation plane, (which is the mean disparity of the uprights). Thresholds are lower than for the square (except perhaps for G.W.). S.M.'s threshold is dramatically lower. (d) The square with two pieces removed from the horizontals to make two brackets. The endpoints of the horizontals are given the disparity which makes the whole figure planar. Thresholds are elevated in comparison to the two lines. S.M. has an especially high threshold. (e) The brackets, but now the endpoints have the same disparity as the uprights they are joined to. Thresholds are generally lower (except perhaps for M.G.), and S.M. has a much lower threshold than for the coplanar brackets (d). In fact, her threshold is the same as for two lines (a).

We have some experimental evidence which argues against this connectivity hypothesis, at least as the sole explanation. First, thresholds can be greatly improved by extending the horizontals of the square a small distance [Fig. 2(c)], giving their endpoints a disparity which is the mean of those of the uprights. Of course, the resulting figure is still connected. We can also find thresholds comparable to those for the square in a nonconnected figure. If we take a square, and remove the central parts of the horizontal lines, so producing two brackets, then subjects show higher thresholds than for the two vertical lines [Fig. 2(d)]. This is only true when we give the endpoints of the broken horizontal lines a disparity which would make the whole figure planar. If we assign to these endpoints the same disparity as the two uprights, then thresholds are reduced for most subjects, especially for S.M., [Fig. 2(e)].

This suggested to us that the poor performance with the square figure is related to its being a planar figure. When the horizontals are extended, this introduces line segments which run out of the plane of the square. And the thresholds with brackets are only high when they are coplanar. We therefore argued that measurements are made relative to a plane somehow defined by the visual system. When two lines have to be compared and nothing else is visible, then the visual system uses a hypothetical fronto-parallel plane. When a square is presented, we should argue that the plane is partly re-defined to be parallel to this square and the re-definition is so complete in S.M.'s case as to prevent her reading any depth from the components of the square. To put it succinctly, for her the square becomes its own reference plane.

Unlike S.M., most subjects perceive depth reasonably well in the square configuration [Fig. 2(b)]. We therefore sought a figure which would serve for them as an analogue of the square. A good choice for this purpose turned out to be a fairly large 7 × 7 lattice of points, arranged with a square spacing, separated by 8 arc min vertically and horizontally. A fixed amount of disparity was introduced between adjoining columns of this lattice, so that it ought to appear rotated around a vertical axis through the fixation point either to the subject's right or his left. By varying the amount and direction of the disparity from trial to trial in the manner described in the



	(a)	(b)
G.W.	83 ± 12	14.4 ± 1.5
M.G.	140 ± 24	13.5 ± 1.7
R.Y.	160 ± 30	6.9 ± 0.6

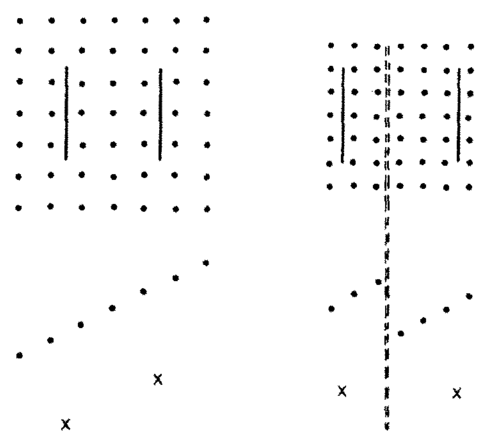
Fig. 3. Disparity thresholds (sec of arc) for three observers for targets consisting of (a) a 7 × 7 lattice of dots with 8 min of arc vertical and horizontal separation, and (b) a pair of columns from this lattice. In each case the threshold indicates the disparity (in sec of arc) between a pair of adjacent columns needed to detect a rotation of the configuration out of the fronto-parallel plane towards the observer's right or left.

methods section, a depth threshold for this configuration can be obtained. For convenience it will be identified as the disparity between adjoining columns (8 arc min apart) at which the subject can judge the tilt of the whole plane correctly on 75% of occasions.

It is almost impossible to obtain a threshold for this configuration [see Fig. 3(a)]. This is a particularly striking phenomenon, because the disparity between the outermost pair of columns is six times that between any adjoining columns. To ensure that dotted vertical lines, on which disparity judgments have to be performed in this pattern, are indeed capable of supporting good stereoacuity, a simple disparity threshold was obtained for a single pair of columns of dots, 8 min of arc apart, seen by themselves. Figure 3(b) shows that this target did indeed give good thresholds. Thus we have succeeded in designing a configuration which does for other observers what a square does for S.M.: its tilt out of the fronto-parallel plane is almost impossible to detect, though its openly displayed components, neither connected nor part of some deliberately misleading form, exhibit perfectly good stereoacuity in isolation.

Two new experiments bring us closer to an understanding of what is involved. The first addresses the question of whether the tilt of the plane containing the lattice of dots is in fact not detected, or whether it is detected but not identified because it lacks appropriate reference. To this end, a pair of vertical lines, 24 min of arc long and 24 min of arc apart, was shown in addition to the lattice of dots [Fig. 4(a)]. The subject had to judge whether the right or the left member of the line pair appeared nearer to him. The dots formed a backdrop and had a tilt with respect to the fronto-parallel plane, with the right edge or the left edge nearer in random order. For better separation, the plane which the dots made up was placed 1 min of arc behind the fixation plane, but that is not a necessary ingredient of the experiment. The subject's responses were accumulated in two separate lots, one for each direction of the orientation of the background sheet of dots. As described in the Methods section, we obtained a threshold for disparity detection of the line pair and also a mean value of the disparity, that is to say, a setting of the two lines in which they both appeared to be lined up in the fronto-parallel plane. No error feedback was provided for the subject and the random nature of the target and background disparity gave them no clue of any data trend.

The results, [Fig. 4(a)], give a clear answer to the original question: the mean setting for fronto-parallelism of the line pair depends on the tilt of the backdrop sheet of dots. Moreover, for most of the



	(a)	(b)
S.M.	44.0	−4.0
G.W.	25.4	−2.0
M.G.	24.3	0.9
R.Y.	32.0	−3.0

Fig. 4. (a) Two lines set with a disparity of 1 min of arc in front of the lattice of dots shown in Fig. 3. Subjects performed the usual threshold procedure for the two lines, while the background was presented in two opposite tilts, randomly mixed. (Below the frontal view we show the view in depth of the configuration in one of its two directions of tilt, towards the subject's left.) From the psychometric curve [Fig. 1(b)], the difference in the means for the two opposite tilts was calculated. If the two lines had been aligned parallel to the lattice, a difference of 24 arc sec would have been found. The values (in arc sec) for four subjects are shown below. (b) The same task, but now the background lattice has been split into two halves, and one of the halves brought forward in depth so that the points located behind the reference lines have equal disparities. In this lattice, the spacing is 6 min of arc between dots, vertically and horizontally. The dotted line in the figure indicates the split in the plane. If slope of the lattice is used to align the test lines, they should show a difference of the means similar to that found in (a); if salience is used, then there should be little difference in the means. The subjects' responses, shown below, indicate that the latter is correct.

subjects there is a startling numerical concordance. If the two targets were set for fronto-parallelism according to the tilt of the background lattice of dots seen at the same time rather than the fixation plane, then the difference in mean settings should be 24 sec of arc. This is the value obtained for two of the subjects, the other two giving a value which was a litttle higher. This means that the tilt of the sheet of dots, to which these subjects cannot respond with any reasonable accuracy when it alone is the stimulus, is nevertheless available to their visual system, for it can provide a quantitatively accurate reference for the setting of the two target lines..

Now, we could imagine two ways in which the tilt of the plane might be detected. One is to use the local tilt of the plane, and the other is to compare the disparity of the background near to the two target lines. To distinguish between these two possibilities, we split the plane of dots into two halves, then separated them in depth in such a way that the regions near to the target lines had the same average disparity, [Fig. 4(b)]. Thus the tilt of the planes was preserved, but the local disparity difference near the target lines was removed. A depth discontinuity could be perceived in this background, but was not conspicuous during the task of comparing the test lines. The data, [Fig. 4(b)], show that there is now very little difference between the mean setting under the two conditions of opposite background slant direction. We can, therefore, conclude that the difference in mean disparity setting of the two target lines has its origin not in the disparity slope of the background of dots, but in the actual disparity value of the background dots in the vicinity of the target lines.

LOCAL INTERACTIONS BETWEEN LINES

Our experiments so far have all been organized around the concept of a reference plane. We asked next whether the effects we have seen could be explained by local interactions which might in certain cases combine to generate something like a reference plane. Interaction concepts are not unknown in the extensive literature on stereopsis (Anstis *et al.*, 1978; Gulick and Lawson, 1976; Jaensch, 1911).

To test this, we sought a figure yet simpler than those we have used so far, which could reveal interactions amongst very few elements. We therefore displayed two test lines, separated by 12 min of arc and 10 min of arc in height. A further line, also 10 min in height, was set either to the left or right of the test lines, 12 min of arc from the nearer member of the pair. The disparity of this additional line, which we shall refer to as the inducing line, was varied, and in randomly interleaved trials it was set either in the fixation plane, or 50 or 100 sec of arc in front of, or behind, the fixation plane. The subject's task was to judge which member of the test pair appeared nearer to him, and to set a switch accordingly. No error signal was provided.

Figure 5(a) shows that the subjects' mean for the test pair is consistently displaced by the inducing line in such a direction as would tend to align the test pair and the inducing line. The shift in the mean is roughly proportional to the disparity of the inducing line.

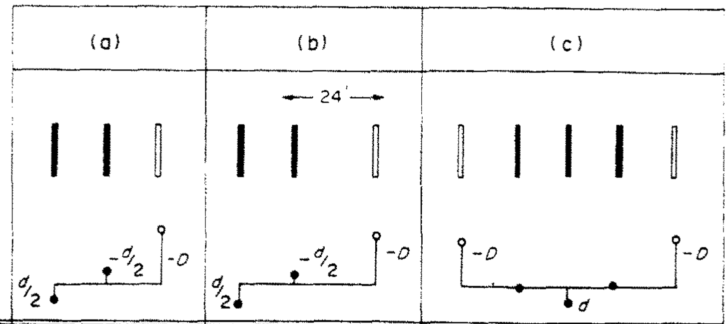
It is not necessary for the inducing line and the test lines to be equally spaced. If the inducing line is 24 min of arc away from the nearest test line, then there is a change in the mean which is in the same direction as in the previous experiment, and increases with disparity of the inducing line, but is of a smaller size, [Fig. 5(b)].

Now we can try to interpret this displacement of the mean by saying that the depth of a given line is, at least partially, measured relative to its neighbours. If the inducing line serves more effectively in this role as a depth standard for the nearer of the two test lines, then we may expect such a shift of the mean. We shall analyze these arguments more thoroughly in the Discussion. For the present, notice that the three line experiment which we are invoking as evidence for local interactions could also be interpreted in the language of the reference plane. One could say that the inducing line twists the test lines somewhat towards their common plane, but is too feeble a stimulus to bring about a complete re-orientation. The next experiment opposes such a view.

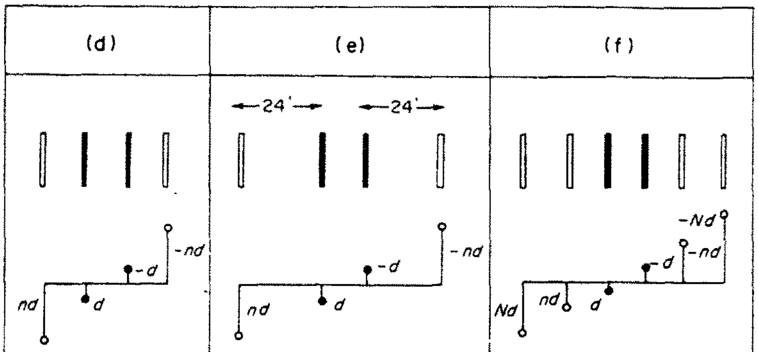
Instead of judging whether one test line is in front of another, a subject can be asked to say whether a central line lies in front of, or behind, two flanking lines which lie in the fixation plane. Thresholds for this task are very low, and even without error signals the mean of the psychometric curve (i.e. that disparity of the center line for which forced responses are equally distributed between "seen in front" and "seen behind") is usually within a few seconds of arc of the fixation plane. To this arrangement, we now added two further flanking lines, which were both either in the fixation plane, or given the same disparity of 50 or 100 sec of arc behind the fixation plane or in front of it, [Fig. 5(c)]. Again, we call these lines of variable disparity "inducing" lines. All lines in the display were 10 arc min high, and separated from their neighbours by 12 min of arc.

For this task we also find a strong effect of the disparity of the inducing lines upon the mean, (Fig. 6). If the inducing lines lie in front of the fixation plane, for instance, the test line is judged to be in the same plane as its two flanking lines when it actually lies behind the fixation plane. This configuration therefore resembles a "V", and it is hard to see how one could account for it by the concept of coplanarity. On the other hand, we shall see in the Discussion that the local interaction hypothesis explains the findings very well.

Having established this, the next experiment was intended to see whether a simple configuration of 4 lines could reproduce the extraordinarily high thresholds found in the earlier experiment with planes of dots [Fig. 3(a)]. We presented the subject with a set



	(a)	(b)	(c)
R.Y.	16.4 (D = 50)	15.2 (D = 100)	22.5 (D = 100)
G.W.	13.4	10.8	20.0
M.G.	——	——	7.5



	(d)	(e)	(f)	
R.Y.	n = 3.5	n = 6.5	N = 0 n = 5	N = 10 n = 3.5
G.W.	4	10	—	—
M.G.	9	11	8	4
G.M.	3	5.5	3	3
S.M.	3	4.5	3	3

Fig. 5. Experiments using rows of lines, separated by 12 min of arc, except where shown. The test lines are filled in, the inducing lines shown as outlines. All lines in experiments (a) and (b) were 10 min of arc high, all other lines were 20 min of arc high. The disparities of lines relative to the fixation plane are shown schematically underneath, with circles corresponding to those lines which lie above them in the figure. Below this is shown the data from five subjects. Saliences can be calculated by inspecting these figures. (a). Using equation (1), the salience of the left test line is $L_1 = -w_1 d - w_2(D + d/2)$ and that of the right test line is $L_2 = w_1 d - w_1(D - d/2)$. A term in this sum is positive when that line lies behind the test line in question, and negative when it lies in front. The saliences of the two test lines are equal when $d(5w_1 + w_2) = 2D(w_1 - w_2)$, or, on putting $r = w_2/w_1$, when $d = 2D(1 - r)/(5 + r)$. This is the calculation done in text, reproduced here for convenience. Data give the effect of 50 sec of arc disparity in the inducing line on the mean setting for apparent fronto-parallelism of the test pair, expressed in the shift of the mean in seconds of arc. The shift is in the direction tending to align the test pair with the inducing line. (b) Letting w_2 be the weighting for lines separated by 24 min of arc, and w_3 that for 36 min of arc apart, we find $L_1 = -w_1 d - w_3(D + d/2)$, and $L_2 = w_1 d - w_2(D - d/2)$. $L_1 = L_2$ when $d(4w_1 + w_2 + w_3) = 2D(w_2 - w_3)$, or, with r as before and $s = w_3/w_1$, when $d = 2D(r - s)/(4 + r + s)$. Data as under (a), except that the inducing line has 100 sec of arc disparity. (c) For the centre test line we obtain $L_1 = -2w_1 d - 2w_2(d + D_d)$, and for the flanking lines $L_2 = w_1 d - w_1 D - w_3 D$. $L_1 = L_2$ when $d = D(1 - 2r + s)/(3 + 2r)$. Data, from Fig. 6, show the disparity shift of the center line necessary for it to appear coplanar with its nearest neighbours when the outer pair is presented with 100 sec disparity. (d) For the left test line we obtain $L_1 = -2w_1 d + w_1(n - 1)d - w_2(n - 1)d$. By symmetry, $L_2 = -L_1$, so that the saliences are equal when $L_1 = 0$, or $n = (3 + r)/(1 - r)$. (e) Proceeding as for (d), we find $L_1 = -2w_1 d + w_2(n - 1)d - w_3(n + 1)d$. Then $L_1 = 0$ when $n = (2 + r + s)/(r - s)$. (f) Here $L_1 = -2w_1 d + w_1(n - 1)d - w_2(n + 1)d + w_2(N - 1)d - w_3(N + 1)d$, and $L_1 = 0$ when $(1 - r)n = 3 + 2r + s - N(r - s)$.

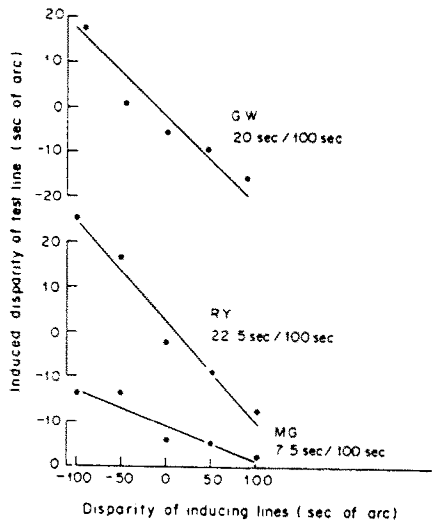


Fig. 6. The data from three subjects from the experiment shown in Fig. 5(c). The mean of the psychometric curve for the disparity of the central line relative to the flanking lines, which are set in the fixation plane, is plotted against the disparity of the outermost lines. Appealing to the diagram in Fig. 5(c), we are plotting against D that value of d which makes the central line appear at equal depth to its two nearest neighbours. Five values of D were used, $D = 100$, 50, 0, -50 and -100 sec of arc. A line fitted to these points (by the method of least squares) is used to estimate the average shift in d per 100 sec of arc shift in D.

of four vertical lines, 20 min of arc high. The centre pair were 12 arc min apart, and the flanking pair were either 12 min [Fig. 5(d)], or 24 min [Fig. 5(e)], away from the nearer of the center pair. The subject's task was to judge which member of the center pair appeared nearer to him and set a switch accordingly. No error signal was provided. The disparity between the outer pair of lines was always a constant multiple, n, of the disparity of the inner pair, [Fig. 5(d)]. The latter was set at random to one of a set of seven equally spaced disparities, as described in the Methods section. Several such series were run in a randomly interdigitated form, for example with $n = 0$, 2, 4, 6, and 10.

The results of the experiment in the case where the four lines are equally spaced are given for three typical subjects in Fig. 7. These show that, for a particular value of n, thresholds become infinite; that is, the subject is unable to identify any depth difference of the two test lines no matter what their disparity. For values of n larger than this critical number, the subject again makes accurate stereo judgements about the test pair, but consistently judges the nearer of the lines to be farther. This yields a negative threshold. The critical value of n is given for all subjects in Fig. 5(d), and is given in Fig. 5(e) for the experiment where the flanking lines are 24 min from the nearer of the test lines.

The final experiment was designed to investigate the range of interactions by seeing whether two flanking lines added to the previous arrangement would have any effect upon the perceived depth of the test pair. We therefore added an outer pair of lines, 12 arc min beyond those lines on either side of the test pair. Their depths took multiples Nd and $-Nd$ of the depths d and $-d$ of the innermost pair, [Fig. 5(f)]. We then measured that value of n which gave an infinite threshold for the test pair, and asked how it changed for two widely different values of N, $N = 0$ and $N = 10$. Figure 5(f) shows that for two subjects, S.M. and G.M., there was essentially no change in n with N. Both R.Y. and M.G. showed appreciable changes, those for M.G. being larger. This suggests that there is a considerable variation between subjects in the range of local interactions.

DISCUSSION

Our findings show that the perceived relative depth of two test objects, using stereoscopic cues alone, is influenced by the presence of other objects nearby in the visual field. This effect is dramatically shown by the grid of dots [Fig. 3(a)], where essentially no depth can be read from any columns of dots, even though two adjacent columns in isolation provide an excellent stereo stimulus. The effect is also shown quite strongly even in such a simple figure as a row of three lines, two of which have to be compared in depth, while the third exerts an "inducing" effect upon them.

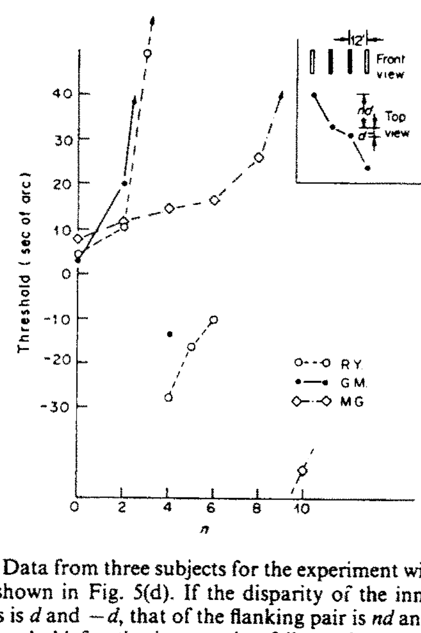


Fig. 7. Data from three subjects for the experiment with four lines, shown in Fig. 5(d). If the disparity of the inner pair of lines is d and $-d$, that of the flanking pair is nd and $-nd$. The threshold for the inner pair of lines, the test pair, is plotted against n. For all subjects, threshold increases with n, becoming infinite for some n, then turning negative. For one subject shown here, G.M., the threshold becomes infinite when $n = 3$. For this value of n, the four lines form a plane in depth. For the other subjects shown here, the critical value of n is larger, corresponding to a configuration where the flanking lines are tilted more steeply in depth than the test pair.

Our earlier experiments were motivated by the concept of a reference plane: we assumed that the visual system creates such a plane, relative to which depth comparisons of a test pair are made, and we assumed that its tilt could be influenced by visual objects near to a test pair. However, the experiment of Fig. 5(c) is not easily explained in this way, as there is no single plane involved. It is, therefore, tempting to look for more primitive interactions in the stereo domain, and to ask that these lead to the notion of a reference plane in a natural way. Our experiments using small numbers of lines were designed to explore such interactions in a simple setting.

Now according to the "Adjacency Principle" (Gogel, 1963; Gogel and Mershon, 1977), the perception of many parameters of visual figures shows a strong effect of objects upon their neighbours. Our findings support this principle in the case of depth, but prompt us to go further and suggest that the perceived depth of an object is entirely determined by the disparity difference between itself and its neighbours. Then we can explain why no depth can be read from two lines in an extended plane of objects, for the two lines have the same disparity difference between themselves and their neighbours and so appear to be at equal depths.

PERCEIVED DEPTH AND SALIENCE

To make this hypothesis more precise, let us define the *salience* of a visual object to be the linear sum, suitably weighted for distance or proximity, of the disparity differences between itself and other objects nearby. If we denote by d the disparity of the test object relative to some fixed reference (e.g. the fixation plane) and the disparity of its neighbours by d_i, for $i = 1,n$, then the salience L is given by

$$L = \Sigma w_i \cdot (d_i - d) \qquad (1)$$

where the w_i are weighting factors. Our basic hypothesis is that two objects have equal perceived depths when they have equal salience.

Note first that, for two isolated test objects, salience is proportional to their negative disparity difference, as would be expected from geometry (see Fig. 1(a)). Consider next the case of the three lines in a row [Fig. 5(a)]. Suppose the left-most pair are to be compared, and suppose their disparities are $d/2$ and $-d/2$ relative to the fixation plane for the left- and right-hand member, respectively. Let the third line have a disparity $-D$ in front of the fixation plane. Then the salience for the left test line is $L_1 = -w_1 \cdot d - w_2 \cdot (d/2 + D)$, while that for the right test line is $L_2 = w_1 \cdot d - w_1 \cdot (D - d/2)$. Here w_1 is the weighting factor for adjacent lines, 12 minutes apart, and w_2 is that for lines 24 min apart, separated by one intervening line.

Now, according to our hypothesis, the test lines appear to have equal depth when $L_1 = L_2$, which implies $d(5w_1 + w_2) = 2D\,(w_1 - w_2)$. If we set $r = w_2/w_1$, this becomes

$$d = 2D\,(1 - r)/(5 + r).$$

This shows that d depends linearly upon D, as was to be expected from our definition [equation (1)]. The values of r needed to fit our subjects' data are given in Table 1(a).

We can now apply this rule to other experiments. In the five line configuration [Fig. 5(c)] subjects must decide whether the central line lies in front of, or behind, its two neighbours, these lying in the fixation plane. If d denotes the disparity between the central line and its neighbours, and $-D$ that of the outer lines relative to the fixation plane, we find by similar calculations [Fig. 5(c)] that $d = D(1 - 2r + s)/(3 + 2r)$ when the central line appears at equal depth to its neighbours. Here r is defined as before, and $s = w_3/w_1$, where w_3 is the weighting for lines 36 min of arc apart, separated by two others.

Consider next the experiment with four lines in a row [Fig. 5(d)], where the central pair are the test pair, with disparity d and $-d$ relative to the fixation plane, and the outer pair take disparities nd and $-nd$. Here we use equation (1) to calculate that value of n for which the saliences of the test pair are equal, which should give infinite thresholds according to our hypothesis. This yields $n = (3 + r)/(1 - r)$. A similar calculation can be done for the six line experiments [Fig. 5(f)], where the outermost lines have disparities Nd and $-Nd$. We expect infinite thresholds when $n = [3 + 2r + s - N(r - s)]/(1 - r)$.

Table 1 shows that it is possible to choose values of r and s which give excellent fits for the data for all of our subjects. Interestingly, there are considerable differences between subjects. For S.M. and G.M., the only interactions are between nearest neighbours ($r = s = 0$). For the other subjects, there are considerable influences of other lines, particularly in the case of M.G.

DISTANCE OR NEIGHBOUR RELATIONS?

So far we have only considered configurations in which the elements are equally spaced. Let us now apply the salience argument to the case of two test lines, 12 min apart, with an inducing line 24 min from the nearer of the test pair [Fig 5(b)]. Let us make the assumption that the w_i depend only on distance. Then the weighting factor for the inducing line to the nearer test line will be what we previously called w_2; namely, the weighting for two lines with one intervening line in an equally spaced array. Similarly, the weighting factor for the inducing line to the further test line will be what we previously called w_3. We then find [Fig. 5(b)] that $d = 2D(r - s)/(4 + r + s)$, where $r = w_2/w_1$ and $s = w_3/w_1$ as before. If we apply the same assumptions to the four line configuration with the flanking lines 24 min from the test pair, we find

Table 1. Experiments using configurations of equally spaced lines

	RY		GW		MG		GM		SM	
	Data	$r = .2$ $s = .05$	Data	$r = 0.25$ $s = .1$	Data	$r = .5$ $s = .3$	Data	$r = 0$ $s = 0$	Data	$r = 0$ $s = 0$
(a)	16.4	15.4	13.4	14.3	—	—	—	—	—	—
(b)	22.5	19.1	20	17	7.5	7.5	—	—	—	—
(c)	3.5	4.0	4.0	4.3	9	7	3	3.0	3	3
(d) N = 0	5	4.3	—	—	8	8.6	3	3	3	3
N = 10	3.5	2.8	—	—	4	4.6	3	3	3	3

Data for each subject (left column) is given with calculated values of salience (right column). There are two free parameters, r and s, in these calculations. For each subject they are chosen to give a good fit over all the experiments. The left column shows schematically the experimental configurations. The calculations are given in the legend to Fig. 5.

(a) A test pair and one inducing line [Fig. 5(a)]. The number given here is the mean, in seconds of arc, of the psychometric curve, when the inducing line has disparity 50 sec of arc. To eliminate intrinsic bias, an average $(M_1 - M_2)/2$ is taken, where M_1 is the mean when the inducing line has disparity $+50$ sec of arc and M_2 is the mean when the disparity is -50 sec of arc.

(b) The experiment depicted in Fig. 5(c). The subject judges whether the central line lies in front of, or behind, its two neighbours. The number given here is the shift in the mean of the psychometric curve, in seconds, per 100 sec change in disparity of the flanking lines.

(c) The experiment shown in Fig. 5(d). If the test lines (the inner pair) have disparities d and $-d$, the flanking lines have disparities nd and $-nd$ (reading from left to right). The number given is that value of n which gives an infinite threshold.

(d) The experiment shown in Fig. 5(f). The configuration is as in (c) above, but two further flanking lines have been added, with disparities Nd and $-Nd$. The value of n which gives infinite thresholds is shown for $N = 0$ and $N = 10$.

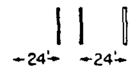
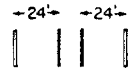
[Fig. 5(e)] that infinite thresholds are predicted for $n = (2 + r + s)/(r - s)$.

Table 2, (second column), shows what happens when we try to fit the data from these experiments using the values of r and s previously obtained. We find a reasonable fit for M.G., poor fits for R.Y. and G.W., while S.M. and G.M. give flagrant disagreements with the calculations. The results for the last two subjects provide a clue as to what is taking place. We know that they showed no effect of lines other than their nearest neighbours in those experiments where all the lines were 12 min apart (Table 1). If distance were all that mattered, lines 24 min away should have no effect. But the flanking lines 24 min from the test pair in Table 2(b) clearly do have a large effect. However, these lines are again the nearest neighbours of the test pair, though more distant.

We are therefore led to the conclusion that, for the two subjects S.M. and G.M., only nearest neighbours contribute to the local depth. This is revealed more clearly by comparing the four line experiment where the flanking lines are 24 min from the test pair [Fig. 5(e)] with the six line experiment [Fig. 5(f)]. In the first case, the flanking lines have a large effect for these two subjects, giving infinite thresholds when $n = 4.5$ (S.M.) and $n = 5.5$ (G.M.). But, in the six line configuration, even when the equivalent flanking lines take widely different disparity values, $N = 0$ and $N = 10$, they have no effect at all upon that value of n which gives infinite thresholds. They are at the same distance from the test pair in both configurations, but in the four line experiments they are nearest neighbours; in the six line experiment they are not.

This predominance of nearest neighbours does not appear to exist for M.G., whose data are all fitted well by weighting functions which depend only upon distance (Tables 1 and 2). The other two subjects show something intermediate. Let us assume that there is a "masking factor", m, which represents the amount by which a weighting function is reduced when an intervening line is introduced. Then we can put $w_2 = mv_2$, and $w_3 = mv_3$, where w_2, w_3 are as before, and v_2, v_3 are the corresponding weighting factors with one fewer intervening lines. If we use v_2 and v_3 in place of w_2 and w_3 for the experiments with unequally spaced lines, we can fit all the subjects' data

Table 2. Two configurations with unequal distances between lines

	RY			GW			MG			GM			SM		
	Data	$r = 0.2$, $s = .05$	$m = .4$	Data	$r = .25$, $s = .1$	$m = .57$	Data	$r = .5$, $s = .3$	$m = 1$	Data	$r = 0$, $s = 0$	$m = 0$, $R = .44$	Data	$r = 0$, $s = 0$	$m = 0$, $R = .57$
	15.2	7.0	16.4	10.8	6.9	11.2	—	—	—	—	—	—	—	—	—
	7	15	7.1	10	15.7	10	~15	14	14	5.5	∞	5.5	4.5	∞	4.5

Data from subjects (first column) are compared with calculations of salience for (1) pure distance weighting (second column), using the values of r and s previously estimated from the experiments listed in Table 1; (2) weighting which combines distance and "masking factor," m, for intervening lines (third column). For S.M. and G.M., where $m = 0$, R cannot be determined from $R = r/m$, so we give that value which fits $n = (2 + r + s)/(r - s)$, [Fig. 5(e)], with $r = R$ and $s = 0$. The distance between the test lines (solid) is 12 min of arc. In (a), which is the configuration shown in Fig. 5(b), the numbers given are the shift in the mean in sec of arc when the inducing line (outline) is given a disparity of 100 sec of arc relative to the fixation plane. As in Table 1(a), an average is taken for disparities of + 100 sec of arc and − 100 sec of arc of the inducing line. In (b), which is the configuration of Fig. 5(e), the number given is that value of n (the multiple of the disparity of the flanking lines relative to the particular disparity of the test pair), which gives infinite threshold.

by taking $m = 0$ for G.W. and S.M., $m = 0.4$ for R.Y., $m = 0.57$ for G.W., and $m = 1$ for M.G. These correspond to 100% masking for G.M. and S.M., 60% for R.Y., 43% for G.W., and no masking for M.G.

Interestingly, if we define $R = v_2/w_1 = r/m$, this being the ratio of weighting factors with no intervening lines, then we get very similar values for all our subjects. For R.Y., G.W., M.G., G.M., and S.M., we find $R = 0.5$, 0.44, 0.5, 0.44, 0.57 respectively (see Table 2 for the values for S.M. and G.M.). This suggests the intriguing possibility that the weighting for distance may be similar for all subjects, the principal differences lying in the degree of masking by intervening lines. The fact that R is about 1/2, on comparing disparity influences of lines at distance of 24 min with those at 12 min, suggests that weighting varies approximately inversely with distance. Very tentatively, we may infer the weighting factor for a larger distance. Assuming that the effect of two intervening lines is to introduce a masking factor m^2, we may define u_3 by $w_3 = mv_3 = m^2u_3$, so that u_3 is the weighting factor for 36 min without intervening lines. If we then compare the weightings for 36 against 12 min, using $S = u_3/w_1$, we find $S = 0.31, 0.31, 0.30$ for R.Y., G.W., and M.G., which is again in good agreement with an inverse distance law, which would predict $S = 1/3$.

We may combine masking factors and the inverse dependence upon distance by expressing the weighting factor w as $w = m^k/|x|$ where k is the number of intervening objects, m the masking factor and $|x|$ the absolute distance. This gives a precise mathematical form to the more general assertion (Gogel and Mershon, 1977) that adjacency effects vary "inversely" with distance. It is interesting to note that the $1/x$ law leads, in the case of strict nearest neighbour dependence, to exact coplanar alignment of test lines with their neighbours [e.g. $n = 3$ for Fig. 5(d), and $n = 5$ for Fig. 5(e)].

OTHER FIGURES

Let us now return to the first experiments which we described, involving the square and its variants. Consider the extended square [Fig. 2(c)]. If we imagine the end points and the uprights of the extended square changed into 4 lines, then our experience with the rows of lines above tells us that it should be possible to read depth from the central pair with good thresholds. Similarly, if we translate the uprights and end-points of the brackets into four lines in row, then the planar configuration, which corresponds to $n = 3$, should have a very poor threshold. In the configuration where each bracket is fronto-parallel, thresholds should be good, as indeed they are.

This analogy is not good enough, however, for it supposes that the central pair of lines is the test pair. In the bracket configuration, the subject is free to use the uprights, corresponding to the outer pair of lines. According to salience calculations, this should give a good threshold, and, of course, this is not what is found.

One quite natural explanation is that the visual system may be poor at comparing depths of objects when they are separated by intervening objects at different depths. Suppose that the comparison is made by summing saliences in an "area of attention" around each of the chosen test objects. If these test objects are close together (as in all of our experiments with rows of lines), then these areas can be small. But when the test objects are far apart, the areas may be large and send a combination of signals from the test objects and others nearby. For instance, the signals from the uprights of the planar brackets will be averaged with the smaller saliences of the other points of the figure, and so give a higher threshold. For the extended square [Fig. 2(c)], the uprights will have larger saliences than the same lines in the square [Fig. 2(b)]. As long as the ends of the horizontals contribute less than the uprights when the latter are being attended to, we can expect a larger salience than for the square, so accounting for the lower thresholds.

Turning next to the grid of points [Fig. 3(a)], we might try to explain the higher thresholds there by saying that two lines within the grid have identical saliences. For the sum in (1) probably does not extend significantly beyond the third neighbour, this being the width of the grid on either side of the central pair of lines. However, the subjects were not restricted to using the central pair of lines for extracting depth from the figure. Again, we have to appeal to the notion that comparisons are difficult between those columns which are separated by intervening columns at different depths. This might give a two- or threefold increase in thresholds. The much higher thresholds found with all our subjects for this figure, and the infinite threshold found with S.M. for a square, show that we have not yet a complete explanation for all the phenomena of depth perception in our figures.

Finally, the aligning of two test bars in front of the plane of dots [Fig. 4(a)] can be readily explained by the salience hypothesis. In the display we used, the environments of the vertical bars were identical, up to two vertical columns of dots on either side, and the contributions to saliences from other columns would be small for all our subjects. The test lines will therefore have approximately equal saliences when they have the same disparity relative to their neighbours in the grid of dots, which means they would be aligned parallel to the grid.

CONCLUSIONS

The perception of depth can be accounted for, quite economically and with considerable numerical accuracy, by the concept of salience, which is the summed disparity—suitably weighted—between a test object and its neighbours. The weighting appears to vary in a roughly inverse manner with distance for all of our subjects. There is also a "masking" effect, which reduces the weighting when lines intervene between the given lines and others nearby. For some subjects, this masking is complete so that salience depends entirely upon nearest neighbours. For others, it is partial, and for one subject there was no detectable masking of this kind.

Why does the visual system use salience? One possibility is that it represents a stage in the generation of planes. Salience is nonzero when objects protrude from planes, or lie at corners between planes. Two objects have equal saliences if they are roughly coplanar with other objects nearby. If the visual system is interested in corners between planes, and objects which protrude from planes, then salience constitutes a useful indicator for this purpose.

Various caveats need to be uttered. First, all our experiments involve comparisons between two or three test lines, and we do not know that the system generates the same signal for all objects, even when they are not being specially attended to. However, this does seem a reasonable assumption. Second, our subjects were chosen from many candidates for their excellent stereoacuity. It is quite rare to find subjects with stereo thresholds for a two line comparison as low as those shown in Fig. 2(a). It is conceivable, therefore, that we have selected for some special stereo mechanism, although we think this unlikely.

There are clearly many questions which remain to be answered. We have only explored depth interactions amongst objects which lie along the horizontal meridian. What happens when perturbing objects lie at some arbitrary point in the visual field, and what then is the definition of neighbourhood relationships? There are also questions which might be posed at the level of neurophysiology. Do the interactions we see occur amongst disparity tuned cells in areas 17 or 18? In particular, are there cells which handle relative disparities, as would be expected from the definition of salience? A masking mechanism could be based upon inhibitory gating of inputs from cells in other regions of the visual field. Can any such connections be seen?

Acknowledgements—Research support from the National Eye Institute, U.S. Public Health Service under grant EY-00220 is gratefully acknowledged. The work was done while G.M. was on leave of absence from the M.R.C. Laboratory of Molecular Biology, Hills Road, Cambridge, England. He thanks the Smith-Kettlewell Institute for hospitality. We are most grateful to Dr S. P. McKee for serving as a subject in our experiments and for a critical reading of the manuscript. We have benefited from discussions with many of our colleagues, especially Drs F. H. C. Crick and S. P. McKee.

REFERENCES

Anstis S. M., Howard J. P. and Rogers B. (1978) A Craik-Cornsweet illusion for visual depth. *Vision Res.* **18**, 213–217.

Finney D. J. (1952) *Probit Analysis.* Cambridge Univ. Press.

Gogel W. C. (1963) The visual perception of size and distance. *Vision Res.* **3**, 101–120.

Gogel W. C. and Mershon D. H. (1977) Local autonomy in visual space. *Scand. J. Psychol.* **18**, 237–250.

Gulick W. L. and Lawson R. B. (1976) *Human Stereopsis. A Psychophysical Analysis.* Oxford Univ. Press.

Jaensch E. R. (1911) Ueber die Wahrnehmung des Raumes. *Z. Psychol.* Ergaenzungsband 6.

McKee S. P. (1983) The spatial requirements for fine stereoacuity. *Vision Res.* **23**, 191–198.

Werner H. (1937) Dynamical theory of depth perception. *Psychol. Monogr* **49**, 1–127.

Westheimer G. (1979) The spatial sense of the eye. *Invest. Ophthal. visual Sci.* **18**, 893–912.