Sensitivity to depth relief on slanted surfaces

Andrew Glennerster
University Laboratory of Physiology, Oxford, UK

Suzanne McKee
Smith-Kettlewell Eye Research Institute, San Francisco,CA, USA

The finest stereoacuity is known to depend on the disparity of a target relative to other visible points. Here we show that a more important factor in determining sensitivity to displacement can be the disparity of a target relative to an invisible interpolation plane through other neighboring points. We tested the sensitivity of observers to displacements of the central column of a regular grid of dots that was either fronto-parallel or slanted about a vertical axis. We found that subjects' sensitivity to displacement was better predicted by a model based on the disparity of a target with respect to the grid plane than it was by a model based on disparity with respect to other reference points. In control conditions carried out on one subject, we found that this result did not depend on adaptation to the grid slant because it also occurred when the direction of grid slant varied from trial to trial. Nor did it depend on the perception of slant, because the data were similar for trials on which the grid was perceived as approximately fronto-parallel or markedly slanted. Our results indicate that sensitivity to the depth component of the target displacement is based on disparity relative to a local reference plane.

Keywords: binocular stereopsis, relative disparity, reference frame

Introduction

There is now compelling physiological evidence that the initial processing of binocular disparity in the visual system is carried out in a retinal coordinate frame, using the absolute disparities between features in the left and right retinal images (Cumming & Parker, 1999, 2000) These physiological findings stand in stark contrast to psychophysical evidence that the visual system is sensitive to the relative disparity between points, a quantity that is independent of eye position (e.g., Andrews, Glennerster, & Parker, 2001; Erkelens & Collewijn, 1985; McKee, Welch, Taylor, & Bowne, 1990; Westheimer, 1979). There are data suggesting that cells in extra-striate areas may respond selectively to the relative disparity between surfaces, over a range of absolute disparities (Eifuku & Wurtz, 1999; Thomas, Cumming, & Parker, 2002). This may be an important step toward generating a representation of depth that is independent of eye movements.

However, animals move their heads as well as their eyes. Relative disparities are head-based, because they are measured relative to the Vieth-Müller circle, which describes the locus of zero retinal disparity in the plane containing the eyes and the fixated point. As a result, if you look at a bumpy surface and rotate your head to the right and then the left all the relative disparities generated by the surface will change. Somehow, these changing relative disparities contribute to a stable perception of surface shape.

Within the region of space most commonly studied in stereo experiments (i.e., for points approximately straight ahead of the observer), relative disparity describes the disparity of a point with respect to a fronto-parallel plane through another point. A series of psychophysical findings has challenged the idea that relative disparity described in this way is the quantity that is important to the visual system (Glennerster & McKee, 1999; Glennerster, McKee, & Birch, 2002; Mitchison & McKee, 1987; Mitchison & Westheimer, 1984). Collectively, these studies investigate various aspects of stereoscopic processing when the target is presented close to a slanted surface. They provide strong evidence that the important variable to the visual system is the disparity of a point with respect to the slanted reference plane rather than disparity with respect to the fixation plane or, indeed, any fronto-parallel plane.

This conclusion is relevant to a debate about whether shifter circuits are used in binocular processing (Anderson & van Essen, 1987; Nishihara, 1987; Quam, 1987). The proposal is that an intended vergence eye movement or an attentional shift to a different depth plane could alter local circuitry such that features at the attended depth plane would be processed at a fine scale resolution that is normally applied only to features in the fixation plane. Shifter circuits were proposed as a mechanism of mimicking the effect of vergence eye movements, which Marr and Poggio (1979) had originally proposed as the method of bringing fine scale analysis to bear at a particular depth plane. In the shifter circuit models, fine scale analysis could be switched to a new depth plane without an eye movement. The new depth plane was always assumed to be fronto-parallel. If shifter circuits are to be invoked to explain the psychophysical data on slanted surfaces, there would need to be, first, a signal about the surface slant and, second, a wider range of circuits between which to shift.

Mitchison and McKee (1985,1987) were the first to propose that stereo processing might use as its input disparity with respect to a plane defined by neighboring points.

doi:10.1167/4.5.3 Received June 26, 2003; published May 6, 2004 ISSN 1534-7362 © 2004 ARVO

They investigated the correspondence rules used by the visual system when presented with ambiguous stereograms and found that the disparity of matched points is minimized with respect to an interpolation plane through the surface. The salience model that Mitchison and Westheimer (1984) proposed to account for the perceived depth of points is closely related. In addition, Glennerster and McKee (1999) found that thresholds for comparing the depths of two features were determined largely by their disparities with respect to a local reference plane. Finally, Glennerster et al. (2002) showed that detectability of a target displacement depended on how much its disparity changed with respect to the local reference plane. Thus, there is evidence that three central aspects of stereoscopic depth processing – correspondence, the magnitude of perceived depth and sensitivity to the relative depths of points – are determined by disparity with respect to a local interpolated plane.

In this work, we explore further the effect of a slanted reference plane on thresholds for detecting the displacement of a target and discuss how the visual system could compute and store disparities with respect to a local interpolated plane.

Methods

Apparatus

The initial experiments were carried out in San Francisco (Figure 2), where stimuli were composed of dots drawn by computer-generated signals on two Hewlett-Packard 1332A monitors, each equipped with a P4 phosphor. The images on the monitors were superimposed by a beam-splitting pellicle. Orthogonally oriented polarizes placed in front of the monitors and the subject's eyes ensured that only one screen was visible to each eye. Stimuli were viewed in a dimly lit room. The background luminance was low (0.005 cd/m^2 measured with a Pritchard photometer), and the dots were bright (space-averaged luminance of 6 cd/m^2 for a 1.6 by 1.6 arcmin lattice). Viewing distance was 1.5 m.

For the data collected in Oxford (Figures 2-4), stimuli were presented on two CRT monitors viewed through front-silvered mirrors in a Wheatstone configuration and at a viewing distance of 2.65 m (for details, see Andrews et al., 2001). Dots were 55 cd/m^2, 2-arcmin width presented on a dark background (0.4 cd/m^2). Screen luminances were linearized and dot edges anti-aliased to allow accurate subpixel shifts.

Stimuli

The stimulus was a regular, 7 by 7 square grid of bright dots, either fronto-parallel or slanted about a vertical axis. For the experiments in San Francisco, the grid size was 2°, whereas for those in Oxford it was 4°. Exposure duration was 600 ms, and the interstimulus interval was 500 ms. This was sufficiently long to weaken any apparent motion signal generated by the displacement between the target and reference stimuli. A fixation marker was presented between trials.

Psychometric procedure

Subjects judged in which of two intervals the central grid column was shifted laterally, in depth or a combination of both. In the other interval, it was not displaced (i.e., it appeared in the center of the grid and with zero disparity, as shown by the box in Figures 1 and 2). Incorrect responses were signaled by a tone. The shifted location of the target column was constant for one run of 50 trials (see individual experiments for details). Data points are based on a minimum of 200 trails. Error bars show the *SD* of the binomial distribution.

Results

Experiment 1: Lateral and stereo acuity with a fronto-parallel reference plane

In Experiment 1, the grid of dots was fronto-parallel. Subjects judged in which of two intervals the central grid column was shifted (in the other interval it was always presented in the center of the grid with zero disparity). The different locations tested were at 0 and ±0.22 arcmin disparity at a range of lateral displacements either side of the center of the grid, as shown in Figure 1. The shifted location of the central, target column in the signal interval was constant for one run of trials. For targets at zero disparity, performance improved with lateral displacement. As one would expect, for targets at ±0.22 arcmin, performance was always better than for targets at zero disparity at the same lateral position, and there was no asymmetry in performance for displacements to the left or right (a second subject showed a similar pattern; not shown).

The solid curve shows the predictions of a model in which information about the lateral displacement and the disparity of the target are combined independently, calculated as follows. We collected data for target locations at a range of disparities but with zero lateral displacement (only two of these data points are shown). When d' is plotted against disparity the slope of the best fitting straight line, constrained to pass through the origin, gives a measure of d' per arcmin of disparity (k_1). For this subject, k_1 = 7.2. Similarly, the zero disparity data shown in Figure 1 was used, by the same method, to calculate k_2, the expected d' per arcmin of lateral displacement. Detectability, d', was defined as $d' = \sqrt{2}F^{-1}(P)$, where P is the proportion of correct responses and F^{-1} is the inverse of the cumulative Gaussian function. For this subject, $k_2 = 1.3$.

Then, for each target position, we computed the expected d' contribution from disparity ($d'_d = k_1 d$, where d is target disparity) and from lateral displacement ($d'_l = k_2 l$, where l is target lateral displacement). According to the signal detection integration model, or 'd' summation' (Green & Swets, 1966), the expected detectability of the target, d'_t , is

$$d'^2_t = d'^2_d + d'^2_l, \tag{1}$$

if the disparity and lateral position signals are combined independently. We adjusted these d' estimates to account for cue-independent errors (as if the subject made a random response on a small proportion of trials (Wichmann & Hill, 2001)). The best fit for this error rate, λ, was computed once for the entire data set (λ = 0 for the data in Figure 1). λ is the only free parameter in the model and is constrained to lie between 0 and 0.06. In Figure 1, the d' predictions shown by the solid line have been converted to proportion correct, P, using the formula

$$P = (F(d'_t / \sqrt{2}) - 0.5)(1-\lambda) + 0.5$$

where F is the cumulative Gaussian function.

Experiment 2: Slanting the reference plane

Figure 2 shows results for the same judgments and same paradigm as Experiment 1, but now with the grid slanted about a vertical axis. Data are shown for three subjects. Two (SPM and CQ) were tested in San Francisco and one in Oxford (MDB). As described in "Methods," there were important differences between the stimuli in the two laboratories. In particular, the grid size for subject SPM and CQ was 2 x 2º, and the target disparity was ±0.22 arcmin in addition to the lateral displacements shown on the abscissa. For subject MDB, the grid size was 4 x 4º, and the target disparity was ±0.4 arcmin. The data for MDB are re-plotted from Glennerster et al. (2002).

The data in Figure 2 show a clear asymmetry that depends on the direction in which the grid is slanted. Broadly, when the target is close to the surface performance is poor (e.g., open triangles on the left for the slant shown in Figure 2a), and when it is far from the surface, performance is better (e.g., closed triangles for the same condition). This pattern holds for crossed and uncrossed disparities, for both slants and for all three subjects.

The smooth curves show the predictions of the d' summation model (Equation 1). This is very similar to the model shown in Figure 1, except that instead of disparity and lateral displacement, the cues to be combined are now displacement along and disparity with respect to the reference plane. Of course, in the case of a fronto-parallel reference plane, there is no difference between these. As in Experiment 1, we calculated (separately for each grid slant) (i) k_1, detectability per arcmin of disparity when the target had no lateral displacement and (ii) k_2, detectability per arcmin of target displacement along the plane of the grid. The values of k_1 and k_2 are both lower than in Experiment 1, compatible with the known increase in stereoacuity thresholds in the presence of a slanted reference plane (Kumar & Glaser, 1992). k_1 and k_2 for the three subjects were SPM 6.2 and 1.1; CQ 4.4 and 2.0; and MDB 3.1 and 1.3. Then, for each target position, we computed the expected d' contribution from disparity ($d'_d = k_1 d_r$), where d_r is target

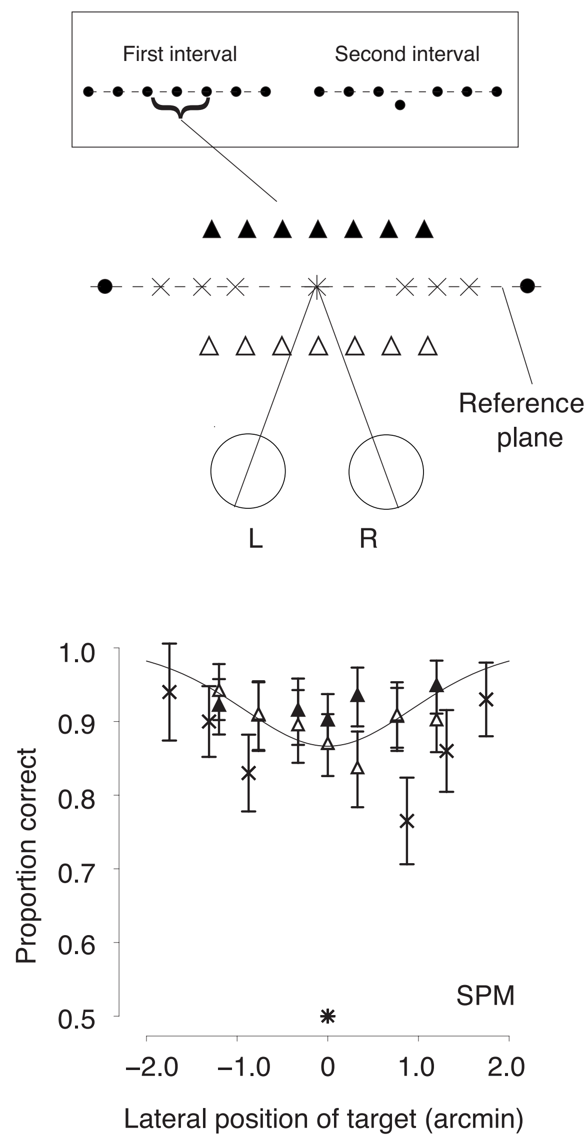

Figure 1. Results from Experiment 1 for a fronto-parallel grid. The task was to detect in which of two intervals the target had been shifted to a location away from the center (which is marked by a star). An example of one trial is shown in the box. The diagram beneath the box shows the target locations we tested. It is a schematic plan view of the stimulus with the eyes drawn artificially close to the surface. Targets with a crossed or uncrossed disparity of 0.22 arcmin (open and closed triangles, respectively) were easier to detect than those at zero disparity (shown by the crosses) for any given lateral displacement. For the data shown by the triangles, the solid curve shows the predicted performance if information about the lateral and depth shift of the target are combined, as described in the text.

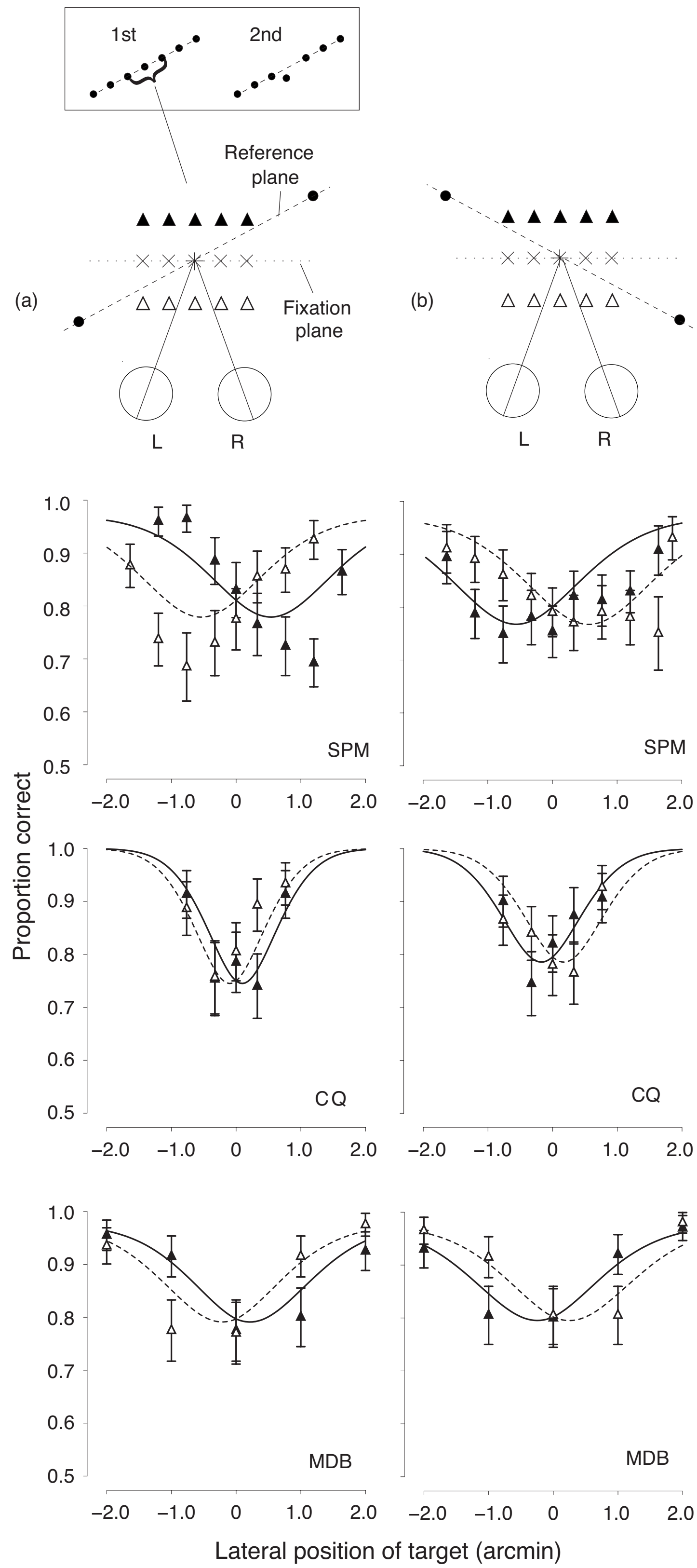

Figure 2. Results for a slanted grid (Experiment 2). Symbols are as for Figure 1. The disparity gradient of the grid was 0.1 for the plots in the left column (a) and -0.1 in (b). Data for three subjects are shown. (There were some differences in the stimuli for SPM and CQ, measured in San Francisco, and for MDB, measured in Oxford: see "Methods".) Data for crossed and uncrossed target locations are shown by open and closed symbols, as in Figure 1. The smooth solid and dashed curves show predictions for these two conditions respectively (see text). This model is the same as that shown in Figure 1 provided that in both cases the disparity of the target is measured with respect to the plane of the grid.

disparity with respect to the reference plane; i.e., the plane of the grid) and from the component of lateral displacement $d'_l = k_2 l_r$, where l_r is target displacement along the reference plane). Note that the target disparities were larger (±0.4) for subject MDB, but the lateral displacements we tested were the same for all subjects. As before, the expected detectability of the target, d'_t, is given by Equation 1. As in Figure 2, d'_t was converted to percentage correct to plot the curves in Figure 1. The solid curves show predicted performance for uncrossed target disparities and the dashed curves predictions for crossed disparities.

The crucial difference between this surface model and a fronto-parallel model is that disparities, d_r, are computed with respect to the reference plane not with respect to the fixation plane (or any other fronto-parallel plane). It is this element of the model that gives rise to the asymmetry in the predictions and the dependence on grid slant. Any model that assumes that the disparity and lateral displacement of the target provide independent information will predict a symmetrical pattern of data, as in Figure 1. To evaluate the two models, we compared the fit of each model to the data using a X^2 statistic. For all three subjects, the fit of the surface model is better than the fronto-parallel model. It should be pointed out that in no case do the data fall within the 95% confidence interval of the model, although for subject MDB $X^2 = 32$ for the surface model, just outside the confidence interval of 30. The X^2 values are as follows: for SPM, fronto-parallel model $X^2 = 388$, surface model $X^2 = 183$, 95% confidence interval $X^2 = 49$ (34 d.f.); for CQ, fronto-parallel model $X^2 = 44.4$, surface model $X^2 = 42.1$, 95% confidence interval $X^2 = 30.1$ (19 d.f.); and for MDB, fronto-parallel model $X^2 = 73.5$, surface model $X^2 = 32.0$, 95% confidence interval $X^2 = 30.1$ (19 d.f.). Values of the one free parameter (the cue-independent miss-rate, λ) were for SPM, 0.06; for CQ, 0; and for MDB, 0.05. These were calculated using all the data shown in Figure 2 for each subject and a fronto-parallel model fit.

Failures of the surface model are generally that it has underestimated the magnitude of the asymmetry in the data. This can be seen clearly at the points where the model curves differ most (e.g., ±1 arcmin for SPM and MDB and ±0.25 arcmin for CQ). In a related experiment, Petrov and Glennerster (2004) also found that a model like the one proposed here underestimated the magnitude of the asymmetry in the data in two subjects (different subjects from those used here). They pointed out that a nonlinear relationship between d' and cue magnitude, similar to that found in contrast detection experiments, could help explain the deviations from the simple model shown here. Despite its failings, the model we have presented here provides a qualitative prediction, indicating the situations in which performance is likely to be better for crossed or uncrossed disparities. The fronto-parallel model, on the other hand, fails to capture these patterns.

The effect of the slanted grid is smallest for subject CQ. This is predictable from this subject's stereoacuity, k_1, and the lateral acuity, k_2. The ratio of these two (2.4:1) is much smaller than for subject SPM (5.6:1), and hence the degree of predicted asymmetry is less (the ratio k_1:k_2 is similar for CQ and MDB, but the stimulus disparity was different for these two subjects, hence the predictions are different, too). Thus, although the data from subject CQ are less useful in distinguishing between rival models than the data of other subjects, they are, nonetheless, compatible with the predictions of the surface model.

Experiment 3: The perception of slant

The purpose of this experiment was to determine whether the effects on sensitivity that we observed in Experiment 2 were a consequence of the pattern of disparities in the stimulus or whether the perception of slant the observer experiences also plays a role in determining thresholds. Glennerster and McKee (1999) showed that depth increment thresholds were lowest in a plane close to a slanted plane that was perceived to be fronto-parallel, but they did not distinguish hypotheses based on the subject's experience from those based simply on the pattern of disparities in the stimulus. We also wished to test whether the effects were only present when the subject is presented with the same grid slant repeatedly, in which case the mechanism responsible for the effect might reflect medium term adaptation to the mean slant experienced over a period of time. For example, a possible hypothesis is that after prolonged exposure to a slanted surface, the visual system reorganizes itself so that the properties normally associated with the horopter (perception of this plane as fronto-parallel, high stereoacuity close to this plane, correspondence matches chosen on the basis of proximity to this plane, etc.) all become associated with a new, slanted plane.

The first hypothesis concerns the perceived slant of the grid. The disparity gradients of the grid stimuli used in Experiment 2 correspond to extreme slants, and yet for the most part they were barely perceived to be slanted at all. The gradients were ± 0.1 and the viewing distance was either 1.5 m (subject SPM and CQ) or 2.65 m (subject MDB). These gradients correspond to physical slants of 67° and 76° from fronto-parallel, respectively. This underestimation of surface slant in regular grid-like stimuli is well known (e.g., Cagenello & Rogers, 1993; Mitchison & McKee, 1990; Mitchison & Westheimer, 1984; Wallach & Bacon, 1976). One might suppose that this underestimation of slant is crucial for the effect on thresholds that we observed.

In the following two control experiments, we examined whether the pattern of sensitivity to displacement of the target was affected by (a) the length of time a subject is exposed to one direction of slant and (b) the subject's perception of slant. To reduce prolonged exposure to one slant, we randomly interleaved trials with opposite directions of slant. As Figure 3 shows, although sensitivity is slightly lower overall than in the non-interleaved experiment, the pattern of results shows the same type of asymmetry as in the previous experiment (Figure 2). Performance is considerably poorer than in Experiment 2, and the predictions from Figure 2 no longer provide a good description of the data. All the same, it is clear that the data has not reverted to the symmetrical pattern predicted by the fronto-parallel model.

To determine the effects of perception of grid slant, we asked the subject to press a second mouse button after they had responded to the target displacement. In this experiment, the grid disparity gradient was 0.1 throughout. The three choices were (i) the grid appeared approximately fronto-parallel, (ii) the grid appeared to have a very large slant, or (iii) neither of the above. In fact, the subject rarely chose the third option, and the other two were chosen about equally often (718, 728, and 154 trials, respectively). Thus, for this subject and this condition (non-interleaved slant and no perspective) the stimulus slant was bi-stable. Other examples of bi-stable stereoscopic slant perception are discussed by van Ee, van Dam, and Erkelens (2002).

Figure 4 shows data for four key target locations (those in which the greatest asymmetry is expected), analyzed separately according to the subjective appearances of the grid. Filled symbols show data for uncrossed stimuli, open symbols data for crossed stimuli. The size of the symbols indi-

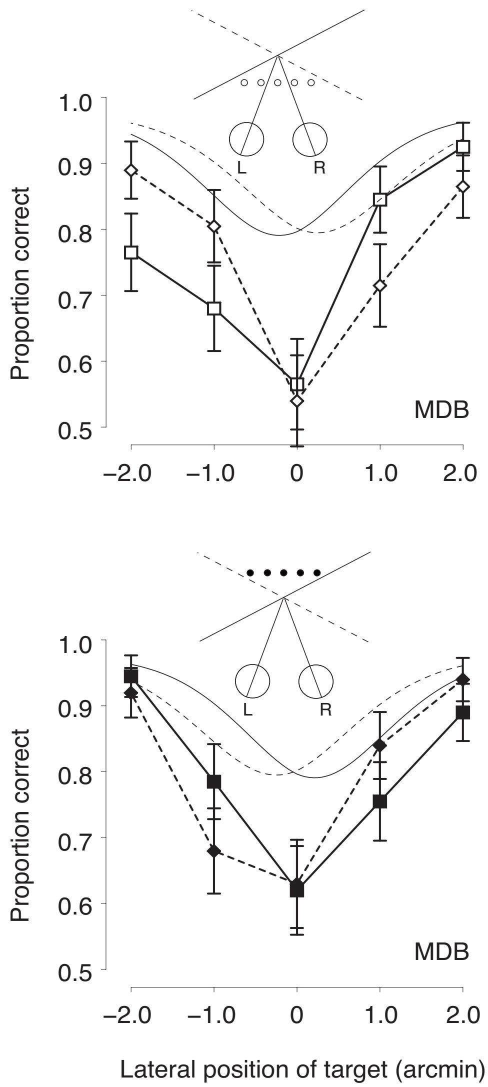

Figure 3. Data from Experiment 3 in which the slant of the grid was randomly varied from trial to trial. Squares show the data for trials in which the grid disparity gradient was +0.1 (solid line in icon) and diamonds show data for -0.1 (dotted line). Data for crossed target disparities are shown on the left (open symbols) and for uncrossed on the right (filled symbols). Note that here the two points plotted for any particular lateral position refer to data gathered within a single run. The format differs from that used to present data in Figure 2, where data for one grid slant was shown on each plot. The curves show predicted performance, replotted from Figure 2 (solid and dashed curves correspond to the solid and dashed lines indicating the grid slant in the icon).

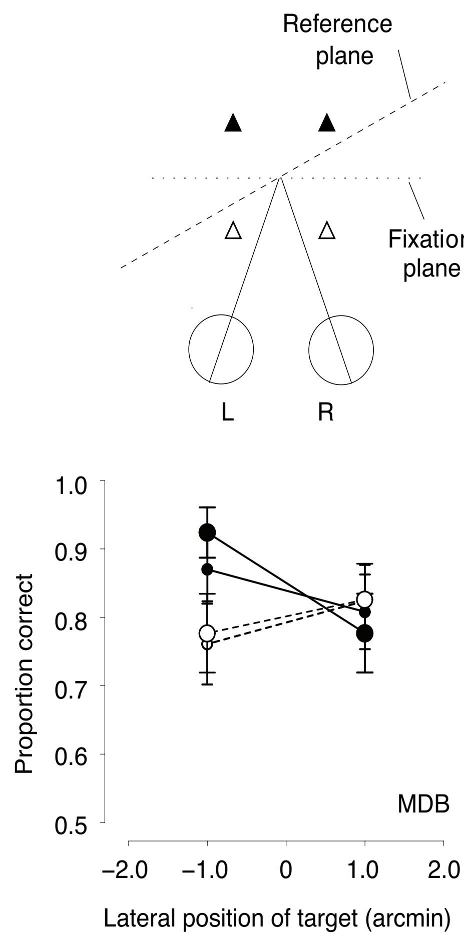

Figure 4. Data from an experiment testing the effect of the subject's perception of grid slant. Grid slant was constant throughout (disparity gradient 0.1) and only four target positions were tested (see icon). After indicating which interval contained the displaced target column, the subject recorded his perception of the grid slant. Large and small symbols show the data when the grid was seen as strongly slanted or approximately fronto-parallel, respectively. The subject could also indicate that the appearance was neither of the above. Data for crossed and uncrossed target positions are shown by unfilled and filled symbols, respectively, as in previous figures. The data are asymmetrical, as in Figure 2, for both types of perception of the grid slant.

cates whether the subject perceived the grid to be flat or strongly slanted. The asymmetry observed in Experiment 2 is evident in both sets of data. Thus, for this subject at least, the effect of the grid on performance does not depend critically on the subject perceiving it as slanted.

The results from these two experiments, albeit sparse and from only one subject, do not support the hypothesis described above (i.e., that a slow reorganization of mechanisms that normally operate close to the fixation plane might occur around a new, slanted plane). However, a more thorough study would be required to make any firm conclusions.

Discussion

The results presented here explore further the demonstrations by Glennerster and McKee (1999) and Glennerster et al. (2002) that the disparity of a target with respect to an invisible interpolation plane can be the critical parameter determining sensitivity to displacement. We have reached this conclusion indirectly, examining displacements that consist of both lateral and depth components. In those subjects who are particularly sensitive to shifts in depth, the results can be used to infer something about the disparity signal that critically affects performance. The argument that binocular processing delivers a disparity signal that is proportional to the depth of a feature with respect to a surface is radically different from the traditional view in which the visual system is sensitive primarily to the disparities of points either with respect to the fixation plane (absolute disparity) or with respect to a fronto-parallel plane through visible points (relative disparity).

Specifically, the data here support the following new conclusions. First, the effect of an interpolation plane is most dramatic for subjects with fine stereoacuity. This fits the predictions of the model, as Figure 2 demonstrates. The predictions (shown by the curves) are based on the sensitivity to pure stereo or pure lateral displacements of the target. For a subject whose stereoacuity is very good, such as subject SPM, the model predicts that the effect of the grid slant will be large, as the data show. On the other hand, for a subject with poorer stereoacuity (relative to lateral acuity) such as CQ, the effect of the grid slant is predicted to be small. Again, the data bear this out. Thus, although the data for CQ are less useful in distinguishing between rival models (such as the surface-model and fronto-parallel model), they are nonetheless compatible with the predictions of the surface model.

The second conclusion must be more tentative, because we have presented data from only one subject. In this case, the subjective perception of slant in the grid had little or no effect on performance. This is evident from the data in Figure 4, where responses are separated out according to the subject's perception of the grid slant. What matters is the grid's disparity gradient and the disparity of points with respect to the interpolation plane. The same conclusion is supported indirectly by the data on interleaved presentation of different grid slants (Figure 3). In this experiment, the subjective grid slant was much stronger than when the grid slant was the same on every trial, yet the data show asymmetries in the same direction in both conditions (cf. subject MDB in Figure 2). The data from this interleaved experiment also demonstrate that the effect of the grid occurs on a single trial rather than being due to longer term adaptation.

Glennerster and McKee (1999) raised the idea that an important factor determining stereoacuity might be the perceived depth difference between points. They showed that the minimum threshold for detecting depth increments tended to occur when the depth difference between target and comparison line was perceived to be small, rather than when the actual disparity difference was small. The data from that study are also compatible with the hypothesis that the disparity of points with respect to the reference surface is the crucial variable. If so, the perceived slant of the reference surface and other lines is not important. The data from the current experiment lend support to this view.

The task in these experiments could potentially be done with one eye closed. However, the results would be different. This is most clearly seen by considering the data for crossed and uncrossed disparities at zero lateral displacement in Figure 1 and Figure 2. These data points would lie close to 0.5 (i.e., chance), if the visual system used only the monocular displacement because stereo thresholds are better than lateral displacement thresholds ($k_1 > k_2$), and the lateral displacement in each eye is only half of the total disparity. For non-zero lateral displacements, the data would lie slightly above the crosses in Figure 1 for one direction of lateral displacements (because the disparity component would increase the lateral displacement in that eye) and slightly below the crosses for the other direction. This is clearly not a good description of the data.

We have shown that in these experiments the disparity signal that best predicts displacement thresholds is the disparity between the target and the reference plane. We have not discussed the way in which the visual system might calculate this quantity (or one that co-varies under the conditions we examined). In the following section, we describe one possible method in which separate metrics for measuring feature position are used in each monocular image. This method is equivalent to calculating disparity with respect to an interpolation plane. Other possible methods exist that are not equivalent but which are sufficiently similar to lead to indistinguishable predictions in our experiment. One is to compute the change in disparity gradient at the target point. Another closely correlated measure is disparity curvature: It could be defined here as the change in disparity gradient at the target point divided by the cyclopean visual angle between the neighboring grid columns. Mitchison (1993) has suggested that disparities are interpolated between features and that subsequent neural operations, such as center-surround mechanisms, could use the interpolated disparity field as its input. Further experiments are required to test between these possibilities (see Petrov & Glennerster, 2004, and discussion by Lappin & Craft, 2000).

A monocular metric for position

This section describes a way to use an image-based metric to define the position of image points and from these compute a measure of disparity that is independent of surface slant. The idea is not new: Koenderink and van Doorn (1991) used this method in their structure-from-motion algorithm. They demonstrated a method of recovering the bas relief structure of a set of points from two images. Bas relief (both mathematically and in art) defines the ratio of depths of features (with respect to some background plane) but not their absolute depths. Like Koenderink and van Doorn (1991), we assume orthographic projection.

Figure 5 shows three plan views of a set of points. The black triangles form a surface in front of which lie a red diamond, a yellow square, and a blue circle. The surface is fronto-parallel in the middle plot and rotated by ±30° in the other plots, with the other three points rigidly attached to the surface. The abscissa and ordinate show the position of the points in the x (lateral) and z depth directions, respectively. Thus, this is an object about 4-cm wide presented at about 150 cm from the observer in different orientations. The left and right eyes' views of the object as seen in the left hand plot (+30° slant) are shown beneath it. The differences between the left and right eyes' views have been exaggerated. The arrows indicate the horizontal width of the surface in the left and right eyes' views, w_l and w_r. It is possible to define the location of all the features using these monocular widths. Taking the bottom left hand triangle as the origin in each image, the horizontal location of the i^{th} feature, P_i, is $x_{l_i} w_l$ in the left eye's image and $x_{r_i} w_r$ in the right eye. The vertical location of features in each eye is equal under orthographic projection and can be ignored here. In the example shown in Figure 5, $x_l = x_r = 0$ for triangles on the left hand edge of the surface, $x_l = x_r = 1$ for triangles on the right, and $x_l = x_r = 0.5$ for triangles in the center. For all points on the surface $x_{l_i} = x_{r_i}$. This follows from the fact that under orthographic projection the left eye's image of a surface slanted about a vertical axis is a uniform horizontal expansion/compression of the right eye's image. $(x_{r_i} - x_{l_i})$ provides a measure of disparity with respect to the plane. Figure 6 illustrates this claim.

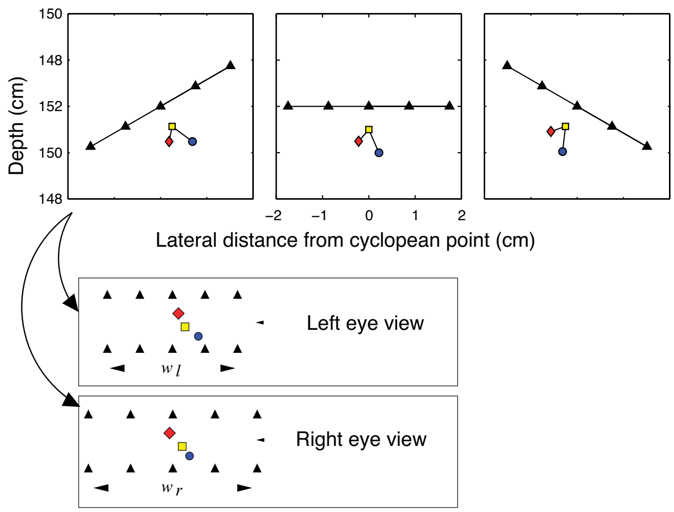

Figure 5. Plan view of an object whose disparities are analyzed in Figure 6 and Figure 7. The object consists of 10 black triangles in a plane and three protruding points shown as a red diamond, yellow square, and blue circle. The surface is fronto-parallel in the center plot and has been rotated (with the other points rigidly attached) by ±30° in the left and right hand plots. The positions of all the points are shown by their lateral (x) and depth (z) coordinates with respect to the cyclopean eye (i.e., a point midway between the eyes where the inter-ocular axis defines the x direction). The boxes below illustrate the left and right eyes' views of the object when it has the orientation shown in the left-hand plot but with exaggerated disparities (the correct disparities are plotted in Figure 6). For this orientation, the width of the stimulus in the right eye, w_r, is greater than the width in the left eye, w_l. Figure 6 and Figure 7 show the effect of "normalizing" image locations using these widths.

Figure 6 shows three plots, each corresponding to the three orientations of the object shown in Figure 5. They show the difference between the normalized positions of features in the left and right eye plotted against their mean normalized positions [i.e., $(x_{r_i} - x_{l_i})$ is plotted against $(x_{l_i} + x_{r_i})/2)$]. This would be a traditional plot of disparity against lateral position were it not for the normalization step, which was applied as follows. The original horizontal locations of each point have been divided by w_l in the left eye and w_r in the right eye. The values of w_l and w_r are (1 –

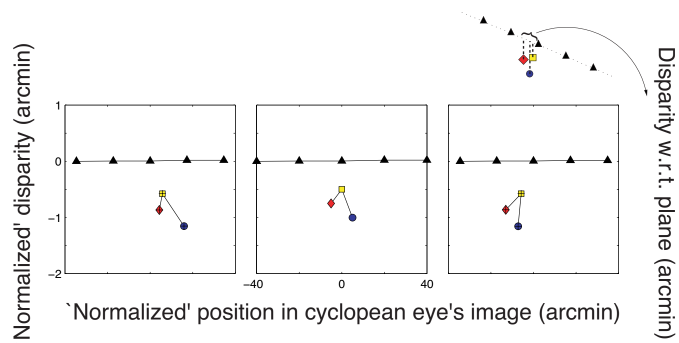

Figure 6. For each plot, the ordinate shows the normalized disparity of each point, which is the difference between the normalized positions of features in the left and right eye, as defined in the text. The abscissa shows the mean of the normalized positions of features in the left and right eye. The three plots correspond to the different orientations of the surface, as in Figure 5. Points on the surface now have zero normalized disparity. For the two slanted surfaces, the crosses show the true disparity difference between the three protruding features and the plane behind them, as illustrated in the icon above.

$g/2$) and ($1 + g/2$) where g is the disparity gradient of the surface. The disparities of the triangles, square, diamond, and circle plotted in Figure 6 are the differences between the normalized horizontal location of these features in the left and right eyes.

A consequence of the normalization is that the surface points have zero disparity in the new metric and the protruding features have a disparity that is relative to the plane of the surface. To confirm that this is indeed what the normalized disparity measures, we have computed precisely what the disparity of each feature is relative to the surface behind it (measured along a cyclopean line of sight, as illustrated in the icon above). Because this computation used a correct perspective projection rather than assuming orthographic projection, the disparities (shown by the crosses for the two slanted surfaces in Figure 6) are very slightly different, but negligibly so for this viewing distance.

Figure 7 shows how the measure of disparity illustrated in Figure 6 can provide an invariant representation of the depth relief of the three protruding points. The normalized disparities of the three points vary with surface slant, as Figure 6 shows (compare disparities of the protruding points in the left and middle plots), but the ratio of these disparity values is almost entirely constant across a wide range of slants. This is, of course, not the case for the disparities of the points with respect to the fixation plane (dashed lines). We have described a simple example of a surface slanted about a vertical axis but the principle is extendible to all slants (indeed any affine image distortion) as Koenderink and van Doorn (1991) describe.

Thresholds for lateral shifts in position are known to be worse than stereoacuity thresholds (e.g., Berry, 1948; Westheimer & McKee, 1979). Indeed, our own data confirm this result. The superiority of stereo thresholds seems to rule out the monocular re-scaling or normalization account of our disparity results. The argument is that if normalized monocular data are sufficiently accurate to act as the input to a disparity mechanism, then they should be accessible for monocular, two-dimensional (2D) judgments and yield sensitivities that are at least as great as for stereo tasks. Instead, we are suggesting here that the normalization of monocular components could be embedded in the calculation of disparity and hence not be accessible for judgments of 2D relative position. An analogous independence of lateral and stereo sensitivities is present in the responses of disparity sensitive neurons, which are generally less sensitive to lateral shifts in position than to shifts in disparity (e.g., Ohzawa, DeAngelis, & Freeman, 1997).

Conclusions

Although the mechanism may remain unclear as yet, there are obvious advantages to computing relief in a surface-based frame of reference. Relative disparities are head-based, as discussed in the "Introduction." At some stage the visual system must factor out these changing disparities if it is to arrive at a stable perception of surface shape independent of head movement. The suggestion here is that by calculating disparities relative to a locally defined reference frame, a significant part of that job could be done at an early stage in binocular processing.

The disparity of points with respect to a local reference plane is useful information for the visual system to extract. The ratio of such disparities is almost entirely invariant to movements of the observer around the surface (see Figure 7). The fact that displacement thresholds are determined by the disparity of points relative to a surface (rather than by disparities relative to each other) suggests that the visual system computes them at an early stage, where the magnitude of these disparities presents a fundamental limit on performance.

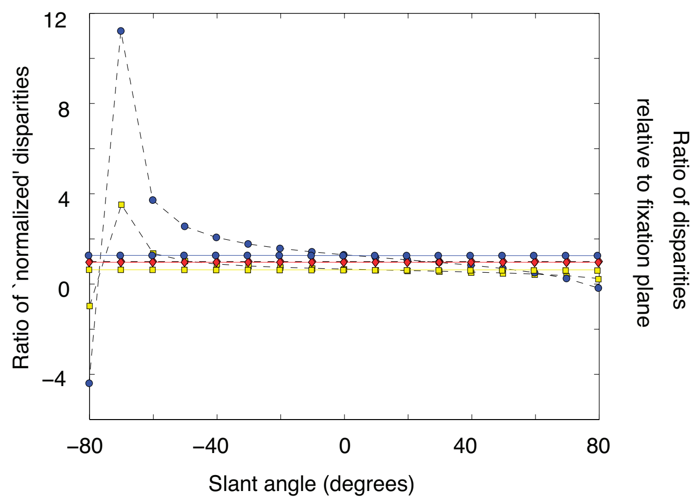

Figure 7. The ratio of normalized disparities of features is invariant to the slant of the surface over a wide range of slants. Figure 6 shows the normalized disparities of the blue circle, red diamond, and yellow square for surface slants -30, 0, and +30 (objects shown in Figure 5). Here, for a range of surface slants, the ratio of normalized disparities is shown relative to the normalized disparity of the red diamond (so the red diamonds are, by definition, at 1). The ratio of disparities shown by the dashed lines are computed from disparities measured with respect to the fixation plane. Again, points for the red diamond are at 1 by definition, because they show the ratio of the disparities of the blue circle and yellow square compared to that of the red diamond. At zero slant, normalized disparities of features are the same as their disparities with respect to the fixation plane, so the dashed and solid curves coincide.

Acknowledgments

We are grateful to Andrew Parker and Martin Birch for their help. This work was supported by the Wellcome Trust and National Eye Institute Grant EY06644. AG is a Royal Society University Research Fellow.

Commercial Relationships: None.
Corresponding author: Andrew Glennerster.
Email: ag@physiol.ox.ac.uk.
Address: University Laboratory of Physiology, Oxford, UK.

References

Anderson, C. H., & Van Essen, D. C. (1987). Shifter circuits: A computational strategy for dynamic aspects of visual processing. *Proceedings of the National Academy of Sciences, 84*, 6297-6301. [PubMed]

Andrews, T. J., Glennerster, A., & Parker, A. J. (2001). Stereoacuity thresholds in the presence of a reference surface. *Vision Research*, *41,* 3051-3061. [PubMed]

Berry, R. N. (1948). Quantitative relations among vernier real depth and stereoscopic depth acuities. *Journal of Experimental Psychology*, *38,* 708-721.

Cagenello, R., & Rogers, B. J. (1993). Anisotropies in the perception of stereoscopic surfaces - the role of orientation disparity. *Vision Research*, *33,* 2189-2201. [PubMed]

Cumming, B. G., & Parker, A. J. (1999). Binocular neurons in V1 of awake monkeys are selective for absolute, not relative disparity. *Journal of Neuroscience*, *19,* 1981-2088. [PubMed]

Cumming, B. G., & Parker, A. J. (2000). Local disparity not perceived depth is signalled by binocular neurons in cortical area V1 of the macaque. *Journal of Neuroscience*, *20*, 4758-4767. [PubMed]

Eifuku, S., & Wurtz, R. H. (1999). Response to motion in extrastriate area MSTl: Disparity sensitivity. *Journal of Neurophysiology*, *82,* 2462-2475. [PubMed]

Erkelens, C. J., & Collewijn, H. (1985). Motion perception during dichoptic viewing of moving random-dot stereograms. *Vision Research*, *25,* 583-588. [PubMed]

Glennerster, A., & McKee, S. P. (1999). Bias and sensitivity of stereo judgments in the presence of a slanted reference plane. *Vision Research*, *39,* 3057-3069. [PubMed]

Glennerster, A., McKee, S. P., & Birch, M. D. (2002). Evidence of surface-based processing of binocular disparity. *Current Biology*, *12,* 825-828. [PubMed]

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: John Wiley & Sons.

Koenderink, J. J., & van Doorn, A. J. (1991). Affine structure from motion. *Journal of the Optical Society of America* A, *8,* 377-385. [PubMed]

Kumar, T., & Glaser, D. A. (1992). Depth discrimination of a line is improved by adding other lines nearby. *Vision Research*, *32,* 1667-1676. [PubMed]

Lappin, J. S., & Craft, W. D. (2000). Foundations of spatial vision: From retinal images to perceived shapes. *Psychological Review*, *107,* 6-38. [PubMed]

Marr, D., & Poggio, T. (1979). A computational theory of human stereo vision. *Proceedings of the Royal Society of London (B)*, *204,* 301-328. [PubMed]

McKee, S. P., Welch, L., Taylor, D. G., & Bowne, S. F. (1990). Finding the common bond: Stereoacuity and the other hyperacuities. *Vision Research*, *30,* 879-891. [PubMed]

Mitchison, G. J. (1993). The neural representation of stereoscopic depth contrast. *Perception*, *22,* 1415-1426. [PubMed]

Mitchison, G. J., & McKee, S. P. (1985). Interpolation in stereoscopic matching. *Nature*, *315,* 402-404. [PubMed]

Mitchison, G. J., & McKee, S. P. (1987). The resolution of ambiguous stereoscopic matches by interpolation. *Vision Research*, *27*, 285-294. [PubMed]

Mitchison, G. J., & McKee, S. P. (1990). Mechanisms underlying the anisotropy of stereoscopic tilt perception. *Vision Research*, *30,* 1781-1791. [PubMed]

Mitchison, G. J., & Westheimer, G. (1984). The perception of depth in simple figures. *Vision Research*, *24*, 1063-1073. [PubMed]

Nishihara, H. K. (1987). Practical real-time imaging stereo matcher. In M. A. Fischler & O. Firschein (Eds.), *Readings in computer vision* (pp. 63-72). Los Altos, CA: Kauffman.

Ohzawa, I., DeAngelis, G. C., & Freeman, R. D. (1997). Encoding of binocular disparity by complex cells in the cat's visual cortex. *Journal of Neurophysiology*, *77,* 2879-2909. [PubMed]

Petrov, Y., & Glennerster, A. (2004). The role of a local reference in stereoscopic detection of depth relief. *Vision Research*, *44,* 367-376. [PubMed]

Quam, L. H. (1987). Hierarchical warp stereo. In M. A. Fischler & O. Firschein (Eds.), *Readings in computer vision* (pp. 80-86). Los Altos, CA: Kauffman.

Thomas, O. M., Cumming, B. G., & Parker, A. J. (2002). A specialization for relative disparity in V2. *Nature Neuroscience*, *5,* 472-478. [PubMed]

van Ee, R., van Dam, L. C. J., & Erkelens, C. J. (2002). Bistability in perceived slant when binocular disparity and monocular perspective specify different slants. *Journal of Vision*, *2*(9), 597-607, http://journalofvision.org/2/9/2/, doi:10.1167/2.9.2. [PubMed][Article]

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

Westheimer, G. (1979). Cooperative neural processes involved in stereoscopic acuity. *Experimental Brain Research*, *36,* 585-597. [PubMed]

Westheimer, G., & McKee, S. P. (1979). What prior uniocular processing is necessary for stereopsis? *Investigative Ophthalmology and Visual Science*, *18,* 614-621. [PubMed]

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function I:fitting, sampling and goodness-of-fit. *Perception and Psychophysics*, *63,* 1293-1313. [PubMed]