Spatial Vision, Vol. 20, No. 1–2, pp. 107–138 (2007)

Also available online - www.brill.nl/sv

Cue combination and the effect of horizontal disparity and perspective on stereoacuity

ANNA M. ZALEVSKI [1,*], G. BRUCE HENNING [2] and N. JEREMY HILL [3]

[1] *Department of Experimental Psychology, South Parks Road, Oxford, OX1 3UD, UK*
[2] *The Sensory Research Unit, Department of Experimental Psychology, South Parks Road, Oxford, OX1 3UD, UK*
[3] *Max-Planck-Institut für biologische Kybernetik, Spemannstr. 38, 72076 Tübingen, Germany*

Received 31 January 2006; accepted 10 May 2006

Abstract—Relative depth judgments of vertical lines based on horizontal disparity deteriorate enormously when the lines form part of closed configurations (Westheimer, 1979). In studies showing this effect, perspective was not manipulated and thus produced inconsistency between horizontal disparity and perspective. We show that stereoacuity improves dramatically when perspective and horizontal disparity are made consistent. Observers appear to use unhelpful perspective cues in judging the relative depth of the vertical sides of rectangles in a way not incompatible with a form of cue weighting. However, 95% confidence intervals for the weights derived for cues usually exceed the *a-priori* [0–1] range.

Keywords: Stereoacuity; cue combination.

1. INTRODUCTION

Most observers are good at estimating the relative depth of visual objects like two vertical lines. One source of information about depth is horizontal disparity — the difference in the lateral position of the images on the retinae. Observers can use horizontal disparity to make fine relative-depth judgments — displacements of the image in one eye by as little as 2–6 seconds of arc result in noticeable changes in depth (Andersen and Weymouth, 1923; Berry, 1948; Bourdon, 1900; Howard, H. J., 1919; Howard, I. P., 2002; Ogle, 1953; Westheimer and McKee, 1978). Stereoacuity, the precision of judgments in stereoscopic disparity, is one commonly used measure of relative-depth discrimination.

However, detecting the relative depth produced by horizontal disparities in vertical lines becomes impossible for some observers when the lines are linked; that is,

*To whom correspondence should be addressed. E-mail: anna.zalevski@worcester-oxford.com

when they form part of a real or imaginary closed configuration such as a square or rectangle. As a result of the 'connection' stereoacuity deteriorates significantly, and becomes unmeasureably large in some observers.

McKee (1983) reported that the loss in stereoacuity in closed figures was first noted by Werner (1937), and results (Fig. 1) from Westheimer (1979) show the dramatic deterioration in performance. The first column of Fig. 1 indicates three types of stimuli: separate vertical lines, the same lines connected horizontally (rectangular configuration), or an inverted 'square bracket'. Stereoacuity for two observers is shown in the remaining columns as thresholds (seconds of arc — bold font) and standard deviations. Both observers had good stereoacuity for separate lines but with the rectangular configuration, the threshold for one observer deteriorated almost 7-fold, and could not even be measured for the other. The loss of stereoacuity forms the principal interest of this study.

The deterioration in stereo-thresholds for closed configurations was confirmed by McKee (1983). When combined, the results of McKee, and those of Mitchison and Westheimer (Mitchison and Westheimer, 1984) note individual differences, with some (including experienced observers) unable to see relative depth at all with simple closed configurations and others showing only moderate deterioration.

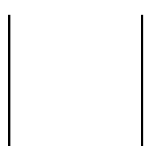
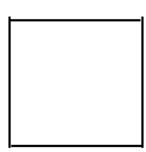
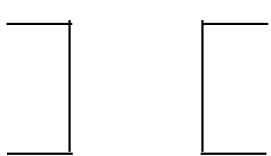
Stimulus	**Stereoacuity thresholds (arcsec) for two observers**	
	Observer 1	**Observer 2**
	**11.7** ± 1.2	**4.5** ± 0.4
	**76.6** ± 11.2	**>100**
	**15.6** ± 1.9	**6.7** ± 0.6

Figure 1. The column labeled 'Stimulus' depicts three stimulus configurations: pairs of lines, rectangles, and inverted square brackets. The second and third columns show stereoacuity thresholds in bold font (and corresponding standard deviations) for two observers obtained with each stimulus configuration (adapted from Westheimer (1979)).

Introducing small modifications to the rectangular configuration has significant effects on stereo thresholds. For example, breaking the horizontal lines, such as in a 'bracket' configuration ([ ]), slightly improved stereoacuity thresholds (Mitchison and Westheimer, 1984). But when the endpoints faced away from each other (] [) — the inverted bracket configuration of Fig. 1 — stereo thresholds were very nearly as good as for a pair of separate vertical lines (Westheimer, 1979).

Monocular cues, like relative size and linear perspective, also provide depth information (Epstein, 1966; Epstein and Landauer, 1969; Howard, 2002; Gogel, 1969; Schlosberg, 1950). In the real world, the linear dimensions of retinal images scale inversely with distance, so that an object close to the observer, subtends a greater visual angle than when it is farther away. This is called 'vertical scaling' and we shall refer to this and related cues as 'foreshortening' or 'perspective' cues. Perspective cues were fixed and thus inconsistent with changing horizontal disparities in experiments where closed forms produced big losses in stereoacuity.

Training with feedback helps some observers (Kumar and Glaser, 1992). Nonetheless, with some configurations, different for each observer, thresholds remain elevated despite extensive training.

It has been suggested that the poor performance resulted from the negative impact of 'connectivity' when features at different disparities formed a figure (McKee, 1983). It has also been proposed that, in the absence of any other depth cues, the visual system uses a hypothetical fronto-parallel plane as a reference frame in order to assess the relative depth of two vertical lines (Mitcheson and Westheimer, 1984). Arguments of this form might now be considered as attempts to characterize Bayesian priors (Knill and Saunders, 2003). Judging relative depth in a rectangular configuration is more difficult, it is sometimes argued, because the reference frame becomes partially re-defined to be parallel to the target rectangle. In the extreme case, the rectangle becomes its own reference plane, and observers lose the ability to see relative depth in the constituents of the rectangle. In other cases, slant reversal occurs (Gillam, 1968; Young *et al.*, 1993). In partial agreement with McKee's hypothesis, it was suggested that it is the monocular figure induced by connectivity — the rectangle — that is important (Stevens and Brookes, 1988); while disparity differences suggest relative depth, the rectangle's shape somehow induces the perception of a figure in a fronto-parallel plane — again an explanation that might be cast in terms of Bayesian priors (Knill and Saunders, 2003).

In a natural environment both binocular and monocular cues are available for depth judgements. When two vertical lines of the same length are viewed (Fig. 2(a)), the line closer to the observer subtends a greater visual angle — this is called foreshortening. If the lengths of the vertical lines are not manipulated with changing disparity, then observers relying exclusively on foreshortening as a cue to depth would not see the lines at different depths even though disparity information indicates that one line or edge of the target is nearer than the other. Horizontal disparity usually provides good cues to depth so that foreshortening may play a negligible role in threshold depth judgments of separate lines.

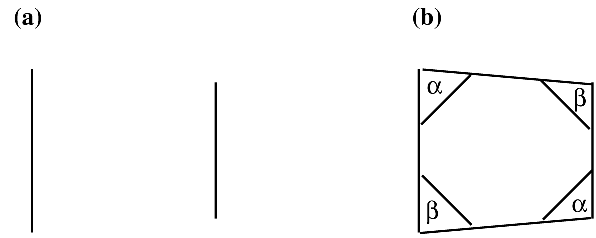


Figure 2. (a) shows a stimulus composed of separate vertical lines: here the only monocular information about depth comes from the difference in the projected length of the lines (not drawn to scale). (b) illustrates a rectangular stimulus slanted about the vertical axis. Here monocular information may again be derived from the difference in the projected length of the vertical lines, but there is additional, linear perspective information from the connecting lines. The lines converge towards the right, and angles β are larger than angles α (not drawn to scale).

However, foreshortening, and associated cues, may be much more effective with closed forms because additional information about relative depth comes from 'perspective' cues: the non-horizontal orientation of the projections of the connecting lines, and the projected acute and obtuse angles they form with the verticals; α and β, respectively, in Fig. 2(b). When disparity- and perspective-based cues provide different depth information, the cues are said to be in conflict. The effect of cue conflict may depend on the relative weights observers give to perspective and horizontal disparity in judging depth and the weights may well change from one stimulus configuration to another. Closed configurations, for example, normally provide many more perspective-based cues than pairs of vertical lines so it seems likely that observers would give more weight to perspective-based cues when viewing closed figures than when viewing pairs of vertical lines. When perspective cues are unhelpful in revealing relative depth — and they are unhelpful when they do not change with changing horizontal disparity — their inclusion in forming relative depth judgments would hurt stereoacuity particulary if they were used almost exclusively, i.e. when heavily weighted relative to disparity.

The importance of horizontal disparity and perspective cue conflict has been shown in supra-threshold studies examining slant perception (Gillam, 1968; Gillam and Ryan, 1992, 1993; Hillis *et al.*, 2006; Knill and Saunders, 2003; Ryan and Gillam, 1994; Stevens and Brookes, 1988; Youngs, 1976). Linear perspective can be a very strong cue to relative depth with some observers appearing to use binocular disparity while others appear to use perspective when the cues are in conflict. Superiority of linear perspective over binocular disparity in slant judgments has been reported (Stevens and Brookes, 1988; Youngs, 1976). Other research shows the importance of perspective in slant judgments but not a clear-cut advantage of this cue over horizontal disparity. For instance Gillam and Ryan, using grids with horizontal and/or vertical lines with a perspective that specified a fronto-parallel surface and a disparity that specified slant, found that surfaces covered with horizontal lines, or horizontal and vertical lines, gave attenuated perceived slant (Gillam and Ryan, 1992). Later they reported that perspective and disparity contributed about equally to perceived slant with the slant estimation reduced by 50% in the presence of cue conflict (Ryan and Gillam, 1994).

In the experiments showing large losses in stereoacuity with closed configurations, foreshortening and the associated perspective cues were fixed: neither the length of the vertically orientated lines, nor the vertical sides and horizontal orientation of the connecting lines of the closed configurations were manipulated in accordance with the changing horizontal disparity. We propose to test an alternative explanation of the loss of stereoacuity with closed forms based on conflict between, or differential weighting of, horizontal disparity and perspective cues. When cue conflict is removed by making perspective and horizontal-disparity cues consistent, stereoacuity with closed forms is almost as good as with pairs of lines.

We measured relative depth sensitivity for closed configurations (rectangules) and open configurations (two separate vertical lines) in three main conditions: a zero-vertical-scaling condition (typifying stereoacuity experiments that show deterioration in performance with closed figures); a slanted-plane condition (similar to viewing a real 3-D scene where horizontal disparity and perspective cues provided consistent information); and a zero-horizontal-disparity condition with only perspective cues providing non-zero relative depth information. (Of course in the latter two conditions the images of the 'rectangles' became trapezoidal but we keep the label, 'rectangle', for the sake of simplicity.) A supplementary monocular condition was also used.

2. METHODS

The two classes of stimuli used in all experiments are illustrated in Fig. 3(a): either two vertical lines, or the same lines joined to form a closed (rectilinear) figure. Vertical lines of both stimuli subtended approximately 25 minutes of visual angle at the observers' eyes and were about 20 minutes of arc apart, a lateral separation that produces good stereo thresholds (McKee, 1983).

The stimuli were computer-generated and displayed on two 21-inch flat-screen Sony FD Trinitron monitors (GDM F500R, one for each eye) at a refresh rate of 100 Hz. The stimuli were presented in a modified Wheatstone stereoscope at an optical distance of 3 m. A chin-rest was used.

The luminance of the uniform background of each monitor was about 0.5 cd/m^2, and the luminance of the line stimuli, about 7.2 cd/m^2. The test room was darkened and features, except for the target stimuli and the monitors, covered with black cloth. An 'anti-aliasing' technique based on a Gaussian kernel with 1-pixel standard deviation was used to allow the production of sub-pixel disparities between the two retinal images (Vollmerhausen and Driggers, 2000).

A two-alternative forced-choice procedure was used. There were two temporal intervals on each trial and a stereoscopically presented stimulus was shown in each observation interval. The observers reported, by pressing keys, the interval in which the left-hand line (PL, in Fig. 3(a)) appeared to lie nearer than the right-hand line

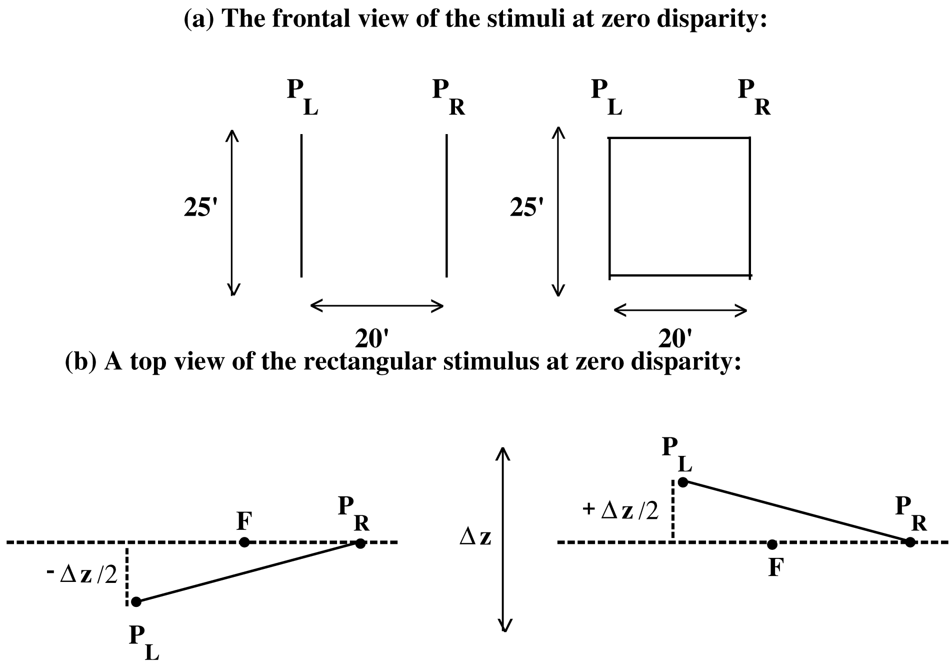


Figure 3. (a) shows a front view of the stimuli: two vertical lines and a rectangle. (b) shows a top view of the rectangular stimulus with PL at a virtual depth distance corresponding to either $+\Delta z/2$ or $-\Delta z/2$ depending on the presentation interval. PR and PL are connected in the rectangular configuration as shown, but were not connected in the line configuration. The fixation point, F, is not shown during the presentation of the stimuli.

(PR). The right-hand line was present in both intervals in the plane of the virtual screen. The left-hand line was displaced by a virtual depth towards the observer in one randomly chosen interval and by the same virtual distance away from the observer in the other (Fig. 3(b)).

The two, one-second long presentation intervals on each trial were separated by a 0.2 second pause. Auditory markers indicated the beginning and end of the observation intervals and provided feedback at the end of the second interval. A binocular fixation cross, in the plane of the virtual screen with arms subtending 30 seconds of arc, was displayed for 2 seconds between trials.

Performance was measured in blocks of 30 trials, with the first five being practice. Disparity was held constant in each block. A number of different disparity values were used to generate psychometric functions of at least five points, usually ranging from above 90% to below 60% correct responses. Each block was usually repeated at least three times. The experimental sessions lasted between an hour and two hours, using extensively trained observers.

There were four experimental conditions, and performance with lines and ‘rectangles’ was compared in each of them:

1. *The Zero-Vertical-Scaling Condition.* Orthogonal projection was used in the vertical dimension so the lines had the same vertical length despite changing disparity. Thus the cues conveyed inconsistent information about relative depth.
2. *The Slanted-Plane Condition.* Separate perspective projections were computed for each eye, and the lengths of vertical lines changed according to the changes in depth that corresponded to the disparity. The horizontal disparity and perspective cues were consistent with the real rotation of a plane containing the stimuli about a fixed vertical axis in space thus providing consistent information about relative depth.
3. *The Zero-Horizontal-Disparity Condition.* Horizontal disparity was not manipulated, but the foreshortening normally associated with particular disparities was introduced: the appropriate perspective projection for each eye was computed from the cyclopean virtual object and presented to the eyes. Horizontal disparity was consistent with objects lying in the fronto-parallel plane through the fixation point and did not change. The stimuli, viewed monocularly, were trapezoidal.
4. *The Monocular Condition.* This condition was exactly the same as in the binocular Zero-Horizontal-Disparity Condition except that observers wore an eye patch over their left eye so that only the right monitor could be seen.

Four observers with normal, or corrected-to-normal, acuity took part in this experiment. Two were authors (AMZ and GBH). Two were naïve to the aims of the study and and one (PJG) had never participated in a psychophysical experiment.

3. RESULTS AND DISCUSSION

The results are all plotted in the same way: they show on semi-logarithmic coordinates and in separate panels for each observer, the proportion of correct responses as a function of disparity difference (seconds of arc). Solid symbols represent results from the 'rectangle' configurations, while open symbols represent results from the vertical-lines configurations. Where possible, Weibull functions were fitted to each psychometric function using the maximum-likehood techniques of Wichmann and Hill (2001a, b). Discrimination 'thresholds' at 60%, 75%, and 90% correct are shown with their 68% (boxes) and 95% (fins) confidence intervals. The 'disparity' corresponding to the 75% correct level is included in each panel. When the observers could not perform the task and psychometric functions could not be fitted, data points are shown with vertical error bars representing ± 1 standard deviation in the proportion correct on the assumption of binomial variability.

Consider first the zero-vertical-scaling condition typifying the conditions used in previous studies (Westheimer, 1979; McKee, 1983).

3.1. Zero vertical scaling

Figure 4 presents results of the zero-vertical-scaling condition; i.e. only horizontal disparity was manipulated. All observers achieved 75% correct discrimination with lines. There is, however, considerable variability among the observers as has often reported in stereoacuity studies.

Performance was very much worse with the closed configuration; three out of four observers did not reach 75% correct discrimination at any of the measured disparities. (LIB's threshold was twice her corresponding stereo threshold with lines.) Consistent with other findings (McKee, 1983; Westheimer, 1979), these results show significant deterioration in depth discrimination performance with vertical lines when the lines are horizontally connected.

The results of the next condition show what happens when perspective cues appropriate for the changes in horizontal disparity are provided.

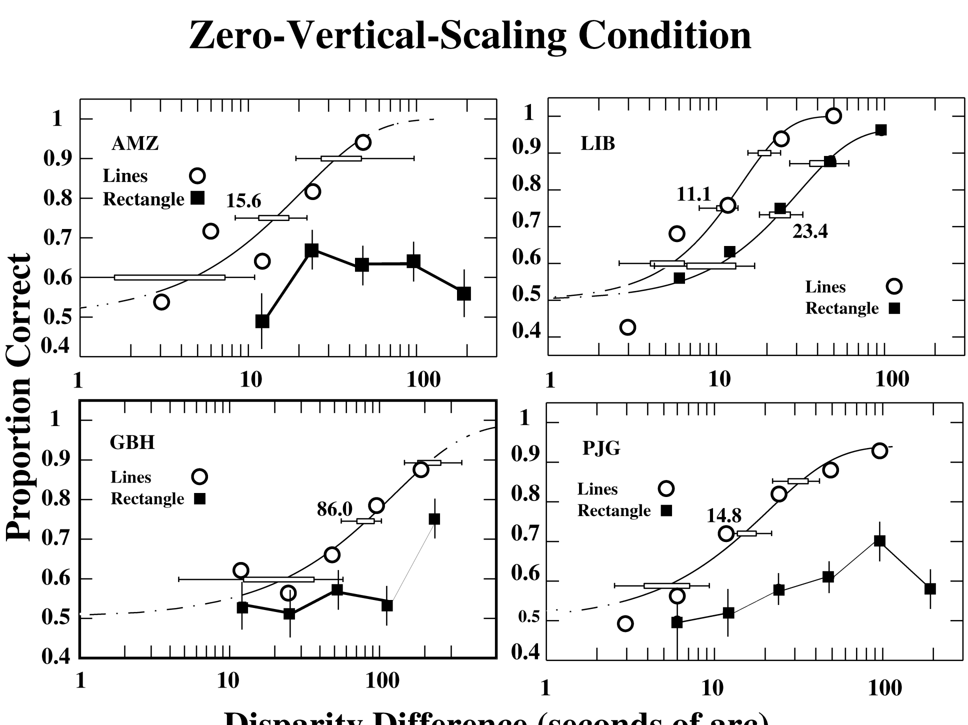


Figure 4. Zero-vertical-scaling condition. Each panel shows separately for four observers the proportion of correct responses as a function of disparity difference on semi-log co-ordinates. Thresholds at 60%, 75% and 90% correct responses are shown with the corresponding 68% (boxes) and 95% (fins) confidence intervals. The smooth curves are the best fitting Weibull functions. When discrimination did not reach 75% correct, vertical error bars, represent ± 1 standard deviation in the proportion correct (assuming binomial variability). Solid symbols represent the rectangular (closed) configuration, and open symbols represent the line configuration.

3.2. Slanted plane

Figure 5 presents results from the slanted-plane condition where horizontal disparity and perspective cues were consistent. With lines, the 75% thresholds for 3 observers show only a slight improvement over their performance in the zero-vertical-scaling condition. The exception was the worst observer GBH, whose 75% threshold improved by a factor of almost 3 when the appropriate foreshortening was present.

The striking result, however, was found when observers viewed the closed configuration. With consistent information from disparity and perspective, the observers who were previously unable to perform the task now had stereo thresholds at 75% correct ranging from 29.07 to 53.84 seconds of arc. LIB almost halved her corresponding stereo threshold when the cues were consistent.

There was a dramatic improvement in performance with the closed figures when consistent perspective information was added. Discrimination in the cue-consistent, 'slanted-plane' condition was still poorer for rectangles than for lines, but the difference was much smaller than without vertical scaling. Thus the principal

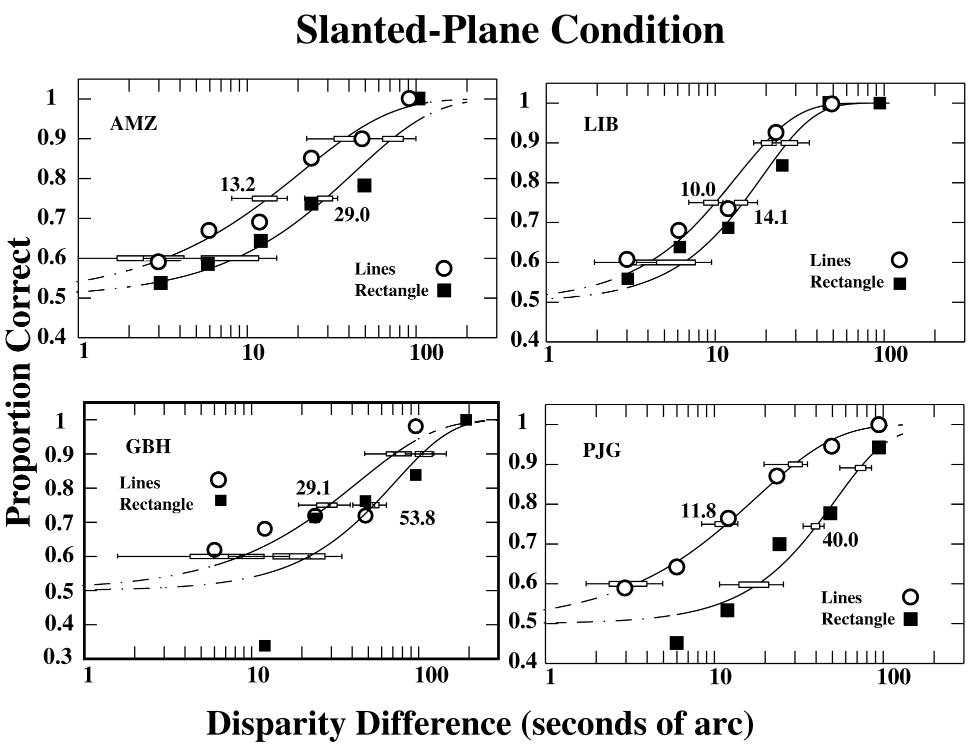


Figure 5. Slanted-plane condition. Each panel shows the proportion of correct responses as a function of disparity difference on semi-log co-ordinates separately for four observers. Thresholds at 60%, 75% and 90% correct responses are shown with the corresponding 68% (boxes) and 95% (fins) confidence intervals. The smooth curves are the best fitting Weibull functions. Solid symbols represent the rectangular (closed) configuration, and open symbols represent the line configuration.

determinant of the large deterioration in stereoacuity with closed configurations appears to be cue conflict in the experimental conditions where changing horizontal disparity suggests changes in depth, but with unchanging vertical scaling, the unchanging foreshortening and perspective cues suggest no change in depth. The effect is much stronger with closed figures than with lines because there are more effective perspective cues that depend on vertical scaling when the stimulus configuration is closed.

Even where the cues are consistent and cue conflict has been removed, the closed figures produced worse performance than pairs of lines. The difference may reflect some of the factors postulated to account for the large effect (McKee, 1983; Mitchison and Westheimer, 1984; Stevens and Brookes, 1988), but a simpler explanation may be that perspective information for the closed figure in the slanted-plane condition is much richer than for the pairs of lines and that, with the closed forms, the observers rely at least partly on perspective cues. With no useful cue from perspective, their performance is hurt by the use of perpective cues. Such behaviour could be modelled by non-optimal cue combination as outlined in Section 5. For all the observers except GBH, any non-optimal use of cues based on perspective is likely to result in poorer performance. With the exception of GBH, performance with pairs of lines is relatively unaffected by the introduction of perspective cues, probably because perspective cue with pairs of lines is limited to a change in the length of the lines — foreshortening — that is very small at all but GBH's stereo thresholds and therefore unlikely ever to have been useful with a stimulus consisting only of a pair of vertical lines.

We also introduced a zero-horizontal-disparity condition, in which horizontal disparity was not manipulated.

3.3. Zero horizontal disparity

The zero-horizontal-disparity condition was designed to measure discrimination thresholds for relative depth where the useful information was provided by cues dependent on vertical scaling, and thus on linear perspective. The introduction of vertical scaling normally associated with particular horizontal disparities in 'rectangles' produces trapezoidal shapes in which the projection of one vertical line (the nearer) is longer than the other, thus suggesting a slanted plane. But horizontal disparity was fixed (at zero) so that observers using horizontal disparity by itself would see no changes in depth.

To simplify comparison with the other conditions, performance is plotted as a function of 'disparity difference'. Disparity difference represents the change in disparity that would have corresponded to the change in the projected length of the target lines in the slanted-plane condition even through, in the current condition, the actual horizontal disparity was zero in both observation intervals.

Figure 6 shows on average a 10-fold deterioration in threshold with lines compared to the previous two conditions. Estimation of relative depth was markedly easier when rectangles were viewed — by a factor of 2–3 on average. Thus with

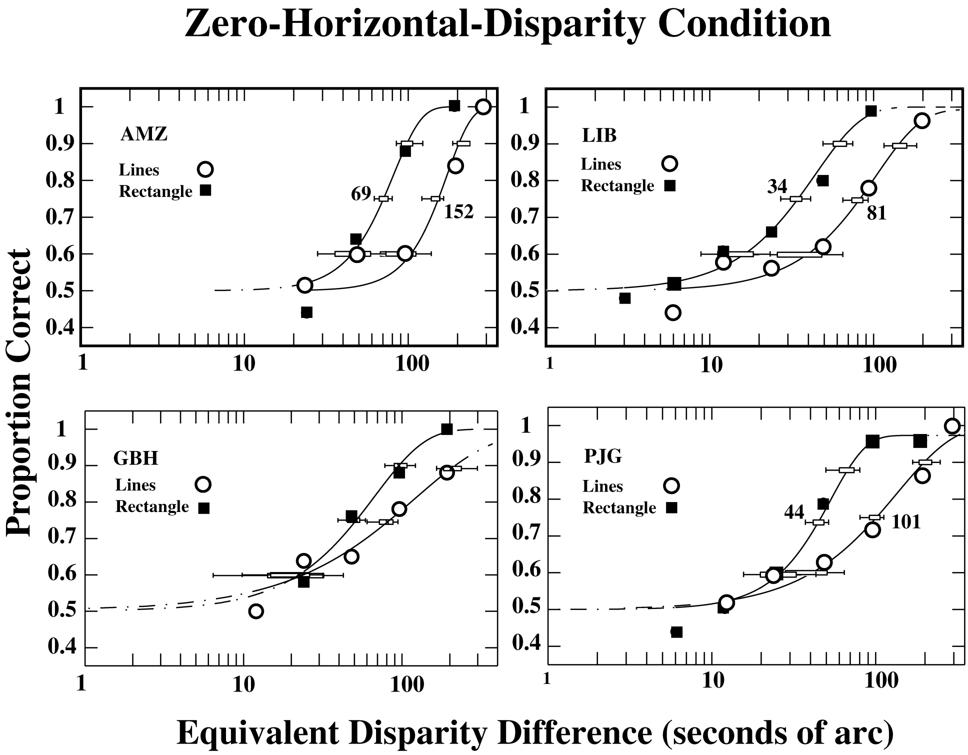


Figure 6. Zero-horizontal-disparity condition. Each panel shows the proportion of correct responses as a function of equivalent disparity difference on semi-log co-ordinates separately for four observers. Thresholds at 60%, 75% and 90% correct responses are shown with the corresponding 68% (boxes) and 95% (fins) confidence intervals. The smooth curves are the best fitting Weibull functions. Solid symbols represent rectangular (closed) configuration, and open symbols represent the line configuration.

zero horizontal disparity the pattern of results found in the other two conditions was reversed. This is hardly surprising since in this condition the only useful depth cues are those based on perspective or foreshortening and the 'rectangle' has many more such cues than pairs of vertical lines.

3.4. Monocular condition

Figure 7 compares, for three observers separately, performance with pairs of vertical lines viewed binocularly (solid circles) or monocularly (open circles) in the zero-horizontal-disparity condition. Observers PJG and AMZ performed in a roughly similar way irrespective of whether the stimuli were viewed with one or with both eyes, while LIB's binocular performance was superior to her monocular performance.

Figure 8 compares binocular (solid squares) and monocular (open squares) performance with rectangles in the zero-horizontal-disparity condition.

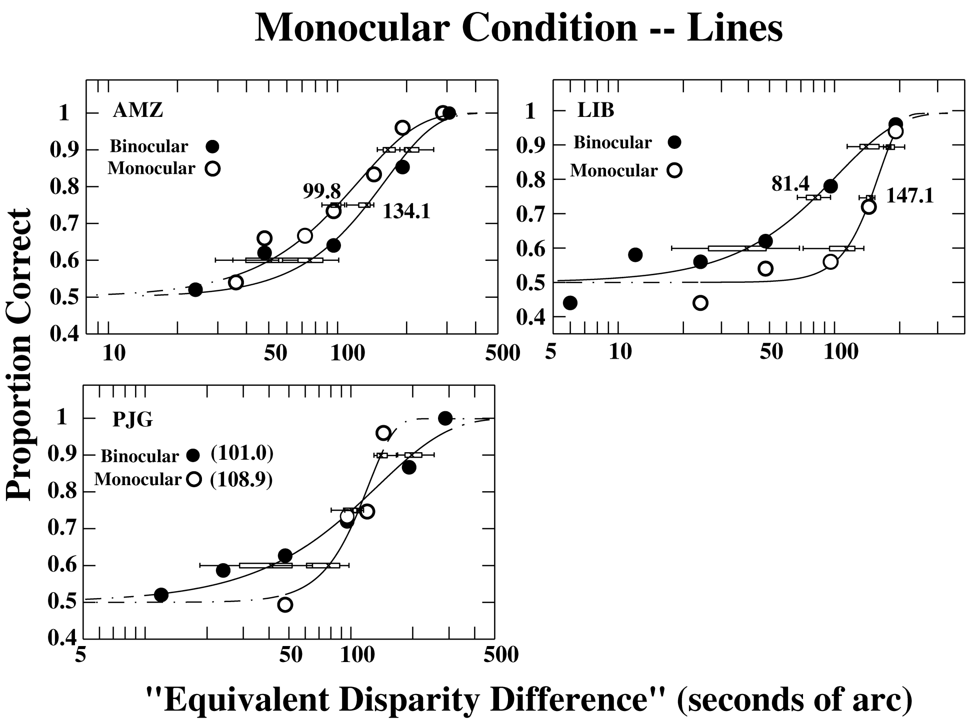


Figure 7. Monocular condition. Each panel shows the proportion of correct responses as a function of 'equivalent disparity difference' on semi-log co-ordinates separately for three observers. Closed circles represent binocular results, while open circles represent monocular results — both sets with pairs of vertical lines and in the zero-horizontal-disparity condition. Thresholds at 60%, 75% and 90% correct responses are shown with the corresponding 68% (boxes) and 95% (fins) confidence intervals. The smooth curves are the best fitting Weibull functions. (Note the change of scale for observer PJG.)

We had hoped that the monocular condition would allow us to measure the observers' sensitivity to perspective cues directly and without interference from information based on horizontal disparity. Discussion is deferred until after we have considered cue combination.

4. GENERAL DISCUSSION

We confirmed the earlier findings showing a significant deterioration in performance with closed configurations (McKee, 1983; Westheimer, 1979). Three out of four observers were unable to reach 75% correct responses with rectangles at any of the disparities we used, but they failed only when vertical scaling was not manipulated. When perspective cues and horizontal disparity provided consistent information about the relative depth of the closed forms, the large losses in stereoacuity almost disappeared. Thus, the effect reported by Westheimer and by McKee appears to arise to a large extent from the inconsistency between horizontal-disparity and perspective cues in their stimuli.

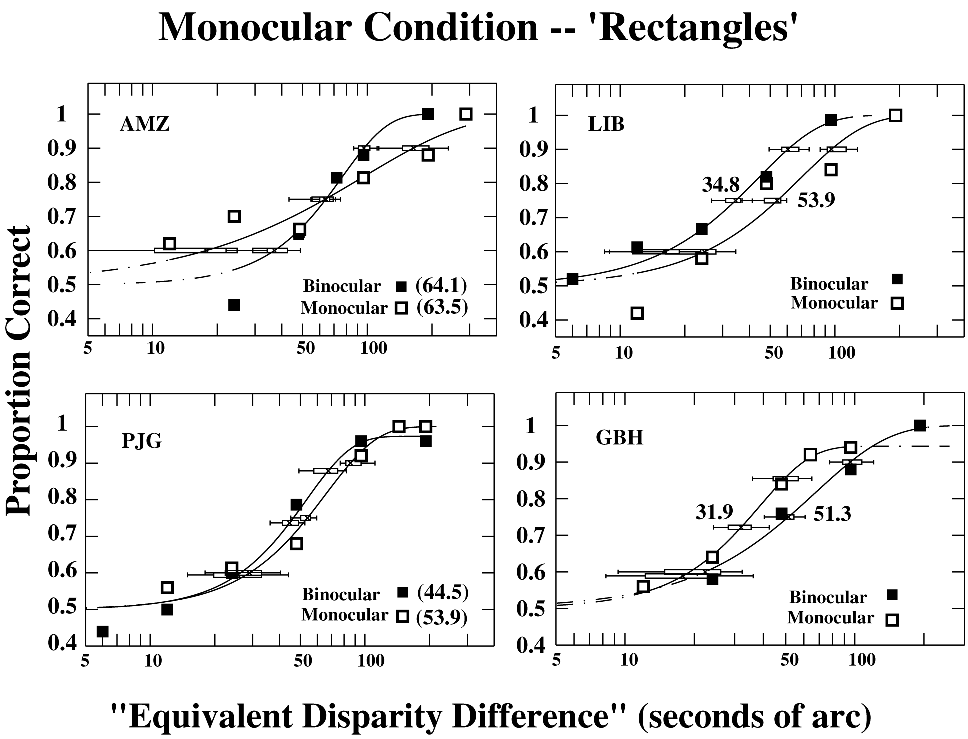


Figure 8. Monocular condition. Each panel shows the proportion of correct responses as a function of 'equivalent disparity difference' on semi-log co-ordinates separately for four observers. Closed squares represent binocular results, while open squares represent monocular results — both sets with rectangles in the zero-horizontal-disparity condition. Thresholds at 60%, 75% and 90% correct responses are shown with corresponding 68% (boxes) and 95% (fins) confidence intervals. The smooth curves are the best fitting Weibull functions.

We found generally good stereo sensitivity for lines irrespective of whether cues were in conflict or not. One interpretation of these results is that the additional perspective information contained in closed forms hurts depth estimation because the less precise perspective cues are included in the formation of that judgment, particularly in experimental conditions where perspective provide no information about experimentally produced changes in depth. Their inclusion in a condition where they are uninformative is clearly not optimal cue combination. It is conceivable, however, that the combination of depth cues is determined in a normal environment and is not flexible in the way that cue-combination experiments involving several modalities suggest.

When perspective alone was manipulated, depth judgments are better with closed than with open configurations. Consistent with other stereoacuity research (Kumar and Glaser, 1992; Mitchison and Westheimer, 1984), considerable variability among as well as within the observers' performance was found.

4.1. Relative cue-weighting and cue-switching explored

It is likely that the normal visual system can use either horizontal disparity, or perspective, or both, to aid relative depth judgements. The information provided by each cue may be thought of as entering into a weighted estimate, where a sensible weighting takes into account both the estimate provided by a given cue and its variability, as well as any other information that might be helpful in assessing the likely accuracy of the cue in any given circumstances (Ernst and Banks, 2002; Ernst *et al.*, 2000; Hillis *et al.*, 2002; Sivia, 1996; Strang, 1988). For example, as we find at long stimulus durations, perspective cues may be less reliable than those based on horizontal disparity but the relative reliability might be reversed at shorter durations or with stimuli that have richer texture. It might even be the case that the weights for different cues alternate from time to time — a possibility beyond the scope of a paper dealing with stereoacuity but one which has been addressed in slant perception (van Ee *et al.*, 2002, 2003). If the observers in fact change the weights assigned to different classes of cue, then they switch sufficiently infrequently in our experiments that the weights assigned to a particular class of cues are not completely hidden in our results.

Leaving the possibility of rapid weight-switching aside, let us consider what happens when the cues are in 'conflict' in the slightly unusual sense that one cue changes from observation interval to observation interval but the other cue does not. In the zero-vertical-scaling condition, horizontal disparity suggests slant, while perspective cues are consistent with a rectangular figure in the fronto-parallel plane. The situation is reversed in the zero-horizontal-disparity condition. Do observers use the information that suggests changes in depth between the observation intervals, or the information that indicates no change? Do they switch between the cues depending on the condition, or do they only give different weighting to the cues? (Cue switching, of course, can be viewed as an extreme form of cue combination in which a cue is given weight one in some condition and weight zero in another.)

Previous stereoacuity experiments report finding much variability among observers (Hillis *et al.*, 2006; Knill and Saunders, 2003; Kumar and Glaser, 1992; McKee, 1983; Mitchison and Westheimer, 1984). The large individual differences, possibly resulting from differences in cue-weighting, have also been reported in slant research (Allison and Howard, 2000; Gillam, 1968; Rosas *et al.*, 2004; Sato and Howard, 2001). Indeed, the weight assigned to cues of potential use in relative depth perception might depend on the observer, the experimental condition, and the configuration. Furthermore, it seems that different weights may be used by the same individual depending on the stimuli viewed and on other, yet unexplored factors — duration and richness of texture, for instance.

Let us now consider in detail cue combination in the context of two-alternative forced-choice experiments.

5. CUE COMBINATION

Let us assume that the decisions about relative-depth are derived from a decision axis (or decision statistic), F, called 'Forwardness', say. The two classes of cue — those based on horizontal disparity and those based on vertical scaling (perspective) — map to this axis. On the 'Forwardness' axis, values less than zero are taken to indicate that the left line is closer than the right line. When the left line is displaced by a certain virtual depth $+\Delta z/2$ away from the observer, we should expect the mean value of the decision statistic to be greater than zero. Similarly when the left line is in fact displaced by a certain virtual depth $-\Delta z/2$ towards the observer, we should expect the mean value of the decision statistic to be less than zero.

In order to proceed, some assumptions must be made about the mapping from stimulus characteristics to the decision axis. It is not possible, of course, to observe the transformation from the stimulus to the decision axis directly. By measuring the observers' performance, we can only observe the results of their decisions with a given stimulus and attempt to infer how these consequences were constrained by some underlying decision statistic, often with the assumption of characteristics of an unknown noise. Further, the decision axis itself is elusive because any monotonic transformation of it will allow the same behaviour to be derived (Green and Swets, 1966).

Let us assume first that the mapping to the decision axis is a monotonic function both from horizontal disparity and from vertical scaling at least over the ranges used in our experiments, and second that the decision axis can be treated as continuous. (The assumption of continuity is not really required but as long as the decision axis is reasonably approximated as continuous, a transformation can be found that makes the decision statistic in at least one of the observation intervals Gaussian in form.)

Consider first the mapping of depth in the display as coded by foreshortening or perspective cues, P, to the decision axis. We need to determine the distribution of the 'Forwardness' statistic based solely on perspective, F_{P}, in both observation intervals of the 2-AFC task that was used.

The means and variances of each probability density for the intervals in which the virtual images are actually forward (f) and backward (b), if the observers used perspective alone, are given by: μ_{Pf}, σ^2_{Pf}, μ_{Pb} and σ^2_{Pb}, respectively. The variability in 'Forwardness' comes from an unknown source of noise. Since by assumption we treat the decision axis as a monotonic transformation of depth-given-by-perspective, we can define a equivalent decision axis such that at least one of the distributions, for example the 'backward' distribution, is Gaussian (Papoulis and Pillai, 2002). Further, it seems reasonable to use a Gaussian approximation to the 'forward' distribution because the stimulus differences are very small.

In the 2-AFC task employed, the observer gets one sample from the backward distribution and one from the forward distribution on each trial. In order to maximize the percentage of corrrect responses, the observer should choose the interval that produces the larger F value as the interval in which the left line was forward. Thus the probability of the observer's being correct is the probability that the difference between a sample of the decision statistic drawn from the distribution that arises from the virtual stimulus' being forward exceeds a sample drawn from the distribution that arises when the virtual stimulus is backward. This probability is readily estimated by combining the density functions from each observation interval to produce the difference distribution, ΔF_{P}. The observation intervals are sufficiently far apart in time that the samples of the decision statistic from each will be treated as independent. Thus, because the parent distributions are Gaussian, the difference distribution will also be Gaussian, with mean $\mu_{\mathrm{Pb}} - \mu_{\mathrm{Pf}}$ and, because the two samples on each trial are assumed to be independent, the difference distribution will have variance $\sigma_{\mathrm{Pf}}^2 + \sigma_{\mathrm{Pb}}^2$. For convenience, and to make the equations simpler and readily recognized, let us define $2\sigma_{\mathrm{P}}^2 = \sigma_{\mathrm{Pf}}^2 + \sigma_{\mathrm{Pb}}^2$.

The probability of an observer's being correct is the probability that the difference distribution is greater than zero: that probability is given by the area under the difference distribution above zero, and if the difference distribution is Gaussian, then that area is related to the statistic, z. The z-value for perspective-based cues, z_{P}, is given by:

$$z_{\mathrm{P}} = \frac{\mu_{\mathrm{Pb}} - \mu_{\mathrm{Pf}}}{\sqrt{2}\sigma_{\mathrm{P}}}, \tag{1}$$

and the probability of being correct using perspective is given by:

$$P(0 \leqslant z_{\mathrm{P}}) = 0.5 + (2\pi)^{-1/2} \int_0^{z_{\mathrm{P}}} \mathrm{e}^{-\frac{1}{2}x^2}\,\mathrm{d}x. \tag{2}$$

Thus the z-values are simply related to proportion correct. (Note that $\sqrt{2}\sigma_{\mathrm{P}}$ in the denominator in equation (1) is merely a convenient way to write $\sqrt{\sigma_{\mathrm{Pf}}^2 + \sigma_{\mathrm{Pb}}^2}$.)

The same line of reasoning gives the z-value for horizontal disparity, z_{D}, as:

$$z_{\mathrm{D}} = \frac{\mu_{\mathrm{Db}} - \mu_{\mathrm{Df}}}{\sqrt{2}\sigma_{\mathrm{D}}}, \tag{3}$$

and the probability of being correct using horizontal disparity alone is given by:

$$P(0 \leqslant z_{\mathrm{D}}) = 0.5 + (2\pi)^{-1/2} \int_0^{z_{\mathrm{D}}} \mathrm{e}^{-\frac{1}{2}x^2}\,\mathrm{d}x. \tag{4}$$

A cue to depth providing no information on which feedback depends, does not help in relative depth estimation but this is not the same as the absence of depth information from that cue. While a cue that does not change from observation interval to observation interval gives the observer no useful information about depth, the cue is still present and may influence decisions even though, ideally, it should be ignored by the observers. Further, in the stereoscopic depth-discrimination experiments under consideration, we cannot, with the closed-form stimuli, measure the proportion correct (or the associated z_{P} and z_{D}) with either cue alone because there are no easy binocular conditions in which either perspective-based or disparity-based cues form the sole possible source of information about relative depth. As a result, it becomes difficult to measure performance with each cue separately and thus difficult to get estimates of their individual variability (on which optimally weighted contributions depend and on the basis of which tests of optimal cue combination rest).

The next section indicates how the cues to depth might be combined, how the weights for each class of cue might be determined, and how the weights might behave under different experimental conditions.

5.1. Cue weighting

One way to view the contribution of different cues is by imagining each cue to contribute to a weighted estimate, where a sensible weighting takes into account both the estimate provided by the given cue and its variability (Backus and Banks, 1999; Ernst and Banks, 2002; Ernst *et al.*, 2000; Hillis *et al.*, 2002; Jacobs, 1999; Landy *et al.*, 1995; Rosas *et al.*, 2005; Sivia, 1996; Strang, 1988; van Ee and Erkelens, 1998; van Ee *et al.*, 1999). Such a method of cue combination emphasizes cues with smaller variance by giving their estimates more weight. Thus, it is sometimes argued, our visual system may combine individual estimates in a linear fashion to produce a weighted estimate, S, such that:

$$S = \sum_i \omega_i S_i, \tag{5}$$

where ω_i is the weighting for the ith cue. The sum of weights over all cues is usually normalized to unity:

$$\sum_i \omega_i = 1. \tag{6}$$

When the cues are uncorrelated (Oruç *et al.*, 2003), the optimal weight for the ith cue is equal to:

$$\omega_i = \frac{1/\sigma_i^2}{\sum_k 1/\sigma_k^2}, \tag{7}$$

where k in the normalizing constant of the denominator is taken over all cues.

It is interesting to investigate whether the two classes of depth cue — horizontal disparity and perspective — are optimally combined. In a sense, of course, they cannot be optimally combined because an optimal observer would give the unhelpful perspective cues zero weight in the zero-vertical-scaling condition where their use can only hurt depth discrimination. Thus, if cue-weighting is the appropriate metaphor, the weights appear not to be readily adjusted for depth discrimination tasks. One possibility is that the weights for the sort of cues manipulated in depth discrimination experiments are fixed — possibly adjustable in terms of early experience in childhood but then no longer flexible.

The rules that govern relative-depth estimation in an impoverished experimental environment, however, may be complex. For instance, when judging depth based on inconsistent cues, the observers could use the information on which the feedback in the experiments depends or, possibly and inappropriately in such a case, the cues that provide no information about depth in the given task but which may have been useful in normal visual experience. They could switch between the cues depending on the condition, give different weighting to the cues depending on the stimulus and the feedback, or weight the cues on the basis of their long experience with the way in which their individual visual systems can process them: some observers may be unable to ignore (uninformative) perspective information, despite extensive training and feedback. Indeed, the results point to high weighting for the perspective cues when viewing closed forms — weightings based, perhaps, on early experience and not easily altered, i.e. fixed, and perhaps optimal for natural viewing conditions.

5.2. Cue weighting by linear combination with fixed and optimal cues

Consider the technique of optimal weighting applied to the results of depth discrimination experiments. Let us assume that the cues to depth are uncorrelated and that, in combining cues, human observers behave optimally in normal circumstances as they are reported to do in some experiments (Bakus and Banks, 1999; Ernst *et al.*, 2000; Ernst and Banks, 2002; Hillis *et al.*, 2002; Jacobs, 1999; Knill and Saunders, 2003; Landy *et al.*, 1995; van Ee *et al.*, 1999). When the cues are consistent, optimal behaviour involves using weights for perspective determined by the reciprocals of the appropriate variances: σ^2_{Pf}, σ^2_{Pb}, σ^2_{Df} and σ^2_{Db} (Oruç *et al.*, 2003; Sivia 1996). Consider the slanted-plane condition first.

5.2.1. Linear cue combination in the slanted-plane condition. Figure 9 shows the process involved in linear cue weighting in the 2-AFC experiments. (For convenience, we assume a particular sequence in the combination but the actual sequence of the linear combination of the independent Gaussian variables is irrelevant.) Weights for cues based on horizontal disparity and perspective multiply the forward (f) and backward (b) statistics that come from each class of cue. The weighted cues from the stimulus-forward interval are added (Figs. 9a and 9b), giving a new Gaussian distribution with a mean equal to the sum of the weighted means, and, assuming the cues to be uncorrelated, a new variance equal to the sums of the weighted variances. Similarly, the weighted cues from the stimulus-backward interval are added to give a new distribution, both in Fig. 9c. The difference distribution between the new forward and the new backward distributions, is shown in Fig. 9d. It is Gaussian in shape (because the forward and backward distributions are assumed to be Gaussian) with the peak at the value of its mean, $\omega_{\mathrm{D}}(\mu_{\mathrm{Db}} - \mu_{\mathrm{Df}}) + \omega_{\mathrm{P}}(\mu_{\mathrm{Pb}} - \mu_{\mathrm{Pf}})$, and a standard deviation equal to $\sqrt{2}\sqrt{\omega^2_{\mathrm{D}}\sigma^2_{\mathrm{D}} + \omega^2_{\mathrm{P}}\sigma^2_{\mathrm{P}}}$. The difference distribution is calculated under the assumption that the information from the two observation intervals is independent and leads to a z-value for the slanted-

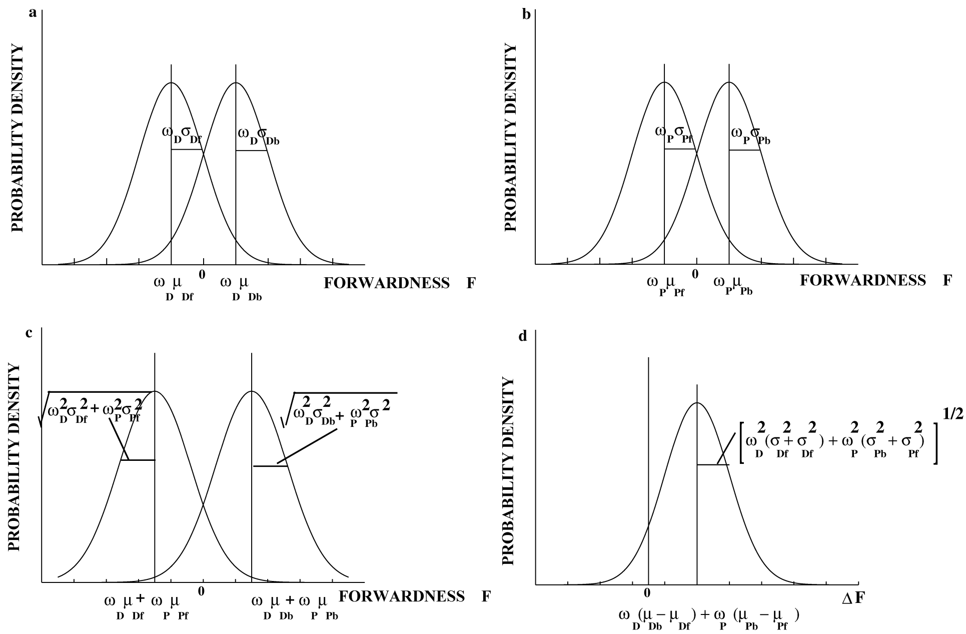


Figure 9. Figure shows the model for the observers' behaviour under the assumption of monotonic and continuous mapping of depth in the display as mapped from vertical scaling (perspective) cues and from horizontal disparity to the decision axis. Two density distributions on the Forwardness axis are shown for each type of cue, one for each of the two temporal intervals in each trial of the 2-AFC experiment. The means and variances for the intervals in which the stimuli were forward (f) and backward (b) are given by: μ_{Df}, σ^2_{Df}, μ_{Pf} and σ^2_{Pf} and by μ_{Db}, σ^2_{Db}, μ_{Pb} and σ^2_{Pb}, in Figs. 9a and 9b, respectively. Figure 9c shows the combined weighted forward and backward density functions produced on the assumption that the observation intervals are independent. Figure 9d is the final difference distribution on Δ_{F} with the mean and variance shown.

plane condition, in which the cues are consistent, z_{C}:

$$z_{\mathrm{C}} = \frac{\omega_{\mathrm{D}}(\mu_{\mathrm{Db}} - \mu_{\mathrm{Df}}) + \omega_{\mathrm{P}}(\mu_{\mathrm{Pb}} - \mu_{\mathrm{Pf}})}{\sqrt{2}\sqrt{\omega_{\mathrm{D}}^2\sigma_{\mathrm{D}}^2 + \omega_{\mathrm{P}}^2\sigma_{\mathrm{P}}^2}}. \tag{8}$$

The z_{C}-value is related to proportion correct determined from the performance measured in the slanted-plane condition. Note, however, that neither σ_{D}^2 nor σ_{P}^2 of equation (8) can be observed directly and that both are merely convenient representations of sums of variances.

At this point it is convenient to break equation (8) into separate terms:

$$z_{\mathrm{C}} = \frac{\omega_{\mathrm{D}}(\mu_{\mathrm{Db}} - \mu_{\mathrm{Df}})}{\sqrt{2}\sqrt{\omega_{\mathrm{D}}^2\sigma_{\mathrm{D}}^2 + \omega_{\mathrm{P}}^2\sigma_{\mathrm{P}}^2}} + \frac{\omega_{\mathrm{P}}(\mu_{\mathrm{Pb}} - \mu_{\mathrm{Pf}})}{\sqrt{2}\sqrt{\omega_{\mathrm{D}}^2\sigma_{\mathrm{D}}^2 + \omega_{\mathrm{P}}^2\sigma_{\mathrm{P}}^2}}. \tag{9}$$

The two terms of equation (9) have numerators that depend on only one source of information, either disparity or perspective, although their common denominator depends on weights and variances of both classes of cue.

Equation (9) describes the situation in the consistent (slanted plane) condition. But, if the weights are fixed and not dependent on the usefulness of a cue in the experiments, then equation (9) reveals what happens once the weights ω_{D} and ω_{P} are known for lines and for closed forms, in the zero-vertical-scaling and zero-horizontal-disparity conditions where cues to depth are not consistent: Let us first consider the zero-vertical-scaling condition.

5.2.2. Zero vertical scaling. If the horizontal disparity and perspective weights do not change with the viewing conditions, then the weights for disparity are the same in the slanted-plane and zero-vertical-scaling conditions. The underlying weights for disparity and perspective cues remain as they were in the slanted-plane condition because that is the (real-life) condition in which the weights are assumed to have been fixed.

In the zero-vertical-scaling condition, the mean difference for perspective is zero: $\mu_{\mathrm{Pb}} - \mu_{\mathrm{Pf}} = 0$ (because the (zero) vertical scaling is identical in both observation intervals). Thus the second term of equation (9) is eliminated and the z-value for the zero-vertical-scaling condition, z_{H}, is given by:

$$z_{\mathrm{H}} = \frac{\omega_{\mathrm{D}}(\mu_{\mathrm{Db}} - \mu_{\mathrm{Df}})}{\sqrt{2}\sqrt{\omega_{\mathrm{D}}^2\sigma_{\mathrm{D}}^2 + \omega_{\mathrm{P}}^2\sigma_{\mathrm{P}}^2}}. \tag{10}$$

Although the second term of equation (9) is zero, neither the variance of the perspective cues, σ_{P}^2, nor the weight given to perspective, ω_{P}, disappears; together they help determine the value of z_{H}. Moreover, if the weight for perspective is close to one, as it might be for some observers viewing ‘rectangles’, the weight for disparity, ω_{D}, would be close to zero and very poor performance would result — just as we find in the zero-vertical scaling condition.

5.2.3. Zero horizontal disparity. An argument similar to that for the zero-vertical-scaling condition suggests that the z-value for the zero-horizontal-disparity condition, z_{V}, is given by:

$$z_{\mathrm{V}} = \frac{\omega_{\mathrm{P}}(\mu_{\mathrm{Pb}} - \mu_{\mathrm{Pf}})}{\sqrt{2}\sqrt{\omega_{\mathrm{D}}^2\sigma_{\mathrm{D}}^2 + \omega_{\mathrm{P}}^2\sigma_{\mathrm{P}}^2}}. \tag{11}$$

Inspection of equations (9)–(11) leads to the prediction that when the weights are fixed, the z-value for performance in the slanted-plane condition should be the sum of z_{H} and z_{V} and this prediction can be tested with experimental data.

Figure 10 presents the data as a function of ‘disparity difference’ for each observer. Data for pairs of vertical lines are shown as open circles (Fig. 10a–d), data for rectangles, as filled squares (Fig. 10e–h). Each figure shows the z-values for the zero-vertical-scaling condition, z_{H} (dotted lines), the z-values for the zero-horizontal-disparity condition, z_{V} (dashed lines), and the sum $z_{\mathrm{H}} + z_{\mathrm{V}}$ (dashed-dotted lines). The data points represent the z-values actually obtained in the slanted-

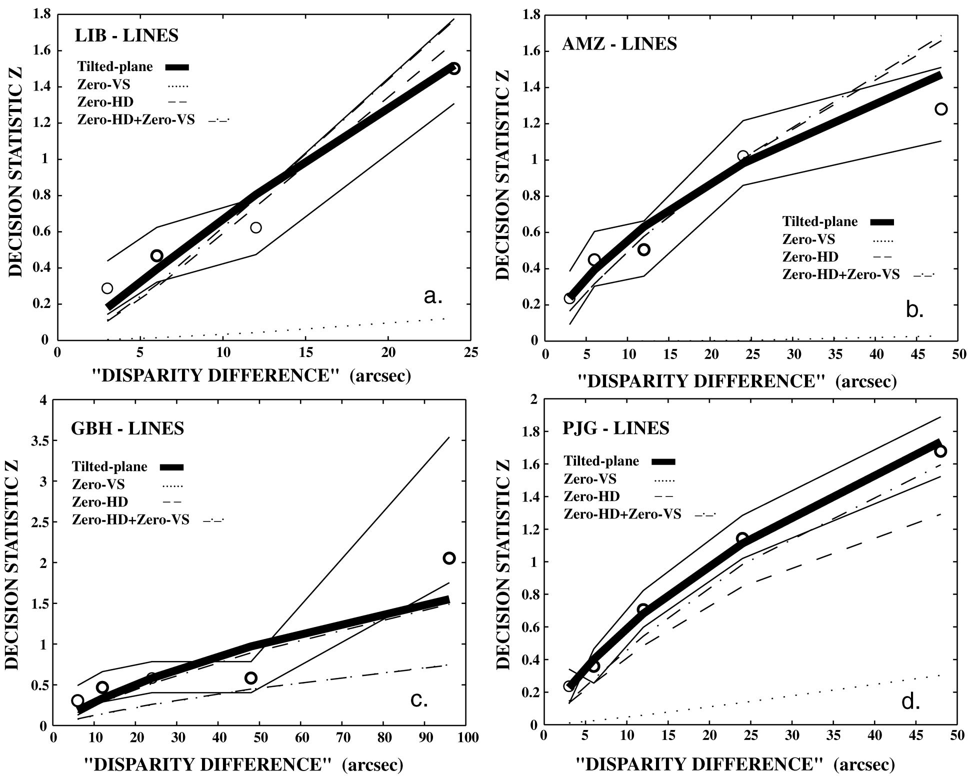


Figure 10. Figure shows for each observer separately the z-values for pairs of vertical lines (a–d) and rectangles (e–h) as a function of disparity difference. Each panel shows the z-values for the zero-vertical-scaling condition, z_{H} (dotted lines), the z-values for the zero-horizontal-disparity condition, z_{V} (dashed lines), and the sum $z_{\mathrm{H}} + z_{\mathrm{V}}$ predicted from equation (9) for the slanted-plane condition (dashed-dotted lines). The data points represent the z-values actually obtained in the slanted-plane condition, z_{C} (calculated from the data of Fig. 5), together with 95% confidence intervals shown as thin solid lines. The thick solid lines show the maximum likelihood fit also from Fig. 5.

plane condition, z_{C}, derived from the data in Fig. 5 together with 95% confidence intervals for that data shown as thin solid lines. The thick solid lines are the maximum likelihood fits also from Fig. 5.

Figure 10 shows that the sum of z_{H} and z_{V} falls approximately within 95% confidence intervals of z_{C} except for AMZ with rectangular stimuli (Fig. 10f). This gives some support to the prediction of equation (9). Let us now return to the cue weighting analysis.

So far, the only observable variables are z_{C}, z_{V} and z_{H}. To proceed further, let us assume that optimal weights are used. That is, $\omega_{\mathrm{D}} = \frac{k}{2\sigma_{\mathrm{D}}^2}$ and $\omega_{\mathrm{P}} = \frac{k}{2\sigma_{\mathrm{P}}^2}$. Given that $\omega_{\mathrm{D}} + \omega_{\mathrm{P}} = 1$, the weights become:

$$\omega_{\mathrm{D}} = \frac{\sigma_{\mathrm{P}}^2}{\sigma_{\mathrm{D}}^2 + \sigma_{\mathrm{P}}^2} \qquad (12)$$

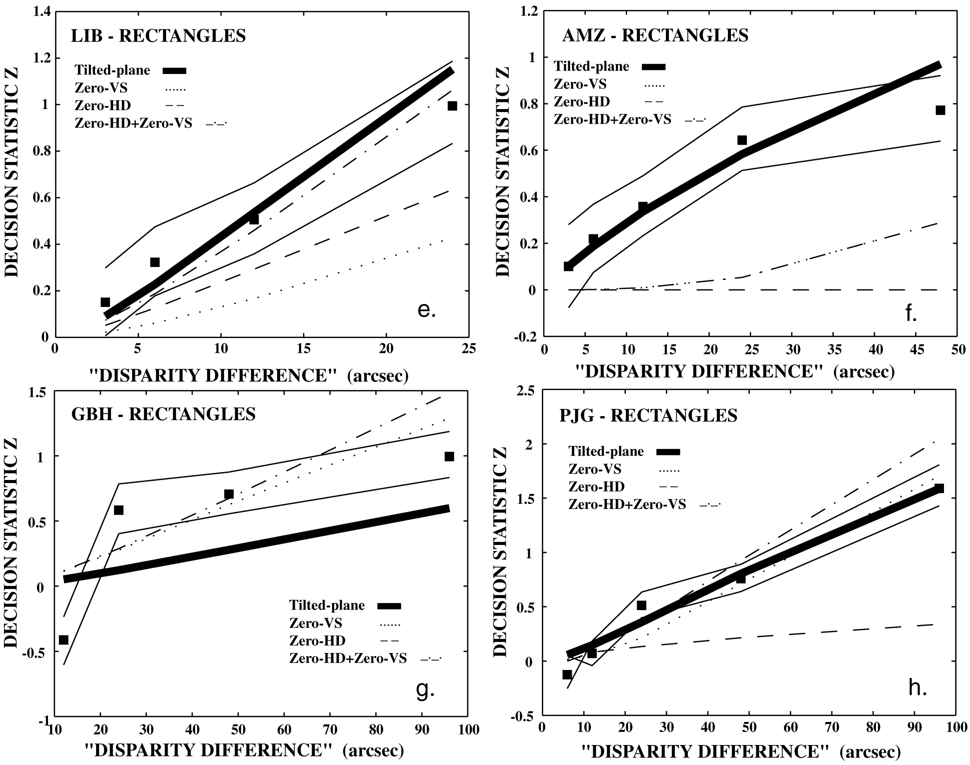


Figure 10. (*Continued*).

and

$$\omega_{\mathrm{P}} = \frac{\sigma_{\mathrm{D}}^2}{\sigma_{\mathrm{D}}^2 + \sigma_{\mathrm{P}}^2}. \tag{13}$$

When these weights are substituted into equation (9), the denominator in the slanted-plane condition becomes:

$$\sigma_{\mathrm{C}} = \frac{\sqrt{2}\sigma_{\mathrm{D}}\sigma_{\mathrm{P}}}{\sqrt{\sigma_{\mathrm{D}}^2 + \sigma_{\mathrm{P}}^2}}. \tag{14}$$

Now equation (9) may be expressed solely in terms of means and variances of the component cues:

$$z_{\mathrm{C}} = \frac{\sigma_{\mathrm{P}}}{\sqrt{\sigma_{\mathrm{D}}^2 + \sigma_{\mathrm{P}}^2}} \frac{(\mu_{\mathrm{Db}} - \mu_{\mathrm{Df}})}{\sqrt{2}\sigma_{\mathrm{D}}} + \frac{\sigma_{\mathrm{D}}}{\sqrt{\sigma_{\mathrm{D}}^2 + \sigma_{\mathrm{P}}^2}} \frac{(\mu_{\mathrm{Pb}} - \mu_{\mathrm{Pf}})}{\sqrt{2}\sigma_{\mathrm{P}}}. \tag{15}$$

When weights have been substituted (from equations (12) and (13)) and when z_{D} and z_{P} are substituted (from equations (1) and (3)), the z-value for the slanted-plane condition is given by:

$$z_C = \sqrt{\omega_D} z_D + \sqrt{\omega_P} z_P, \tag{16}$$

where $\omega_D + \omega_P = 1$. Using the assumptions that led to equation (10) and assuming optimal weighting, z_H is given by:

$$z_H = \sqrt{\omega_D} z_D. \tag{17}$$

Similarly, using the assumptions that lead to equation (11) and again assuming optimal weighting, z_V is given by:

$$z_V = \sqrt{\omega_P} z_P. \tag{18}$$

The z_D cannot be measured easily, because any stimulus with vertical extent must have some perspective even if the perspective does not change from interval to interval. The z-value for the underlying perspective cue, z_P, is also not directly observable, but one possible way of solving equation (18) is to obtain results from the zero-horizontal-disparity condition measured monocularly. If one assumes that the disparity system does not interfere with perspective judgements when one eye is closed, then, when viewing is monocular, $\omega_D = 0$ and $\omega_P = 1$. The results from a monocular zero-horizontal-disparity condition are presented in Figs 7 and 8.

By measuring monocular discrimination in the zero-horizontal-disparity condition to determine z_M, it was hoped to get the z-value for perspective, z_P, without interference from the disparity mechanism. We examine two possible relations between z_P and z_M: $z_P = \sqrt{2} z_M$ if the eyes operate independently, and $z_P = z_M$ for perfectly correlated eyes.

Let us now return to equation (18) and substitute for z_P. If the eyes were perfectly correlated ω_P becomes:

$$\omega_P = \frac{z_V^2}{z_M^2}, \tag{19}$$

or, if the eyes were independent:

$$\omega_P = \frac{z_V^2}{2 z_M^2}. \tag{20}$$

Substituting each z_V and z_M ($= z_P$) into equation (20) allows weights for perspective to be derived, but only if $z_V \leqslant z_P$ (because $0 \leqslant \omega_P \leqslant 1$). Table 1a shows the calculation assuming perfect correlation across the eyes. Across all the observers, only 14 of 51 weights lie in the *a-priori* limited range [0, 1] and of these (shown in bold face) only three weights had their 95% confidence limits within that range. (A short explanation of how the 95% confidence intervals were calculated is given in Appendix A.)

Table 1b shows the calculation assuming independent eyes. Across all the observers still only 17 of 51 weights lie in the *a-priori* limited range [0, 1] and of these (shown in bold face) only seven weights had their 95% confidence limits

Table 1a.
Each configuration (column C) 'equivalent disparity differences' (column D) for four observers with the corresponding proportion correct (P for z_V) in binocular viewing, and proportion correct, for z_M in monocular viewing assuming perfectly correlated eyes

O	C	D	P for z_V	z_V	P for z_M	z_M	$\omega_P = z_V^2/z_M^2$	95% CI
LIB	L	6	0.44	−0.15	—	—		
		12	0.58	0.20	—	—		
		24	0.56	0.15	0.44	−0.15	—	
		48	0.62	0.30	0.54	0.10	—	
		96	0.78	0.77	0.56	0.15	—	
		144	*0.90**	1.28	0.72	0.58	1.55	—
		192	0.96	1.75	0.94		—	
	R	3	0.48	−0.05	—	—		
		6	0.52	0.05	—	—		
		12	0.61	0.27	0.42	−0.20	—	
		24	0.67	0.43	0.58	0.20	—	
		48	0.80	0.84	0.82	0.91	0.85	0.19–4.01
		96	0.99	2.32	0.84	0.99	—	
		192	—	—	1.00	Inf		
AMZ	L	24	0.52	0.05	—	—		
		36	—	—	0.54	0.10		
		48	0.62	0.30	0.66	0.41	0.55	0.00–26.51
		72	*0.59**	0.22	0.67	0.43	0.26	0.00–9.34
		96	0.64	0.35	0.73	0.61	0.34	0.01–2.14
		144	*0.77**	*0.73**	0.83	0.95	0.59	0.16–1.84
		192	0.85	1.03	0.96	1.75	**0.34**	**0.02–0.67**
		288	1.00	Inf	0.96	1.75	—	
ave ω_P							0.37	
	R	12	—	—	0.62	0.30		
		24	0.44	−0.15	0.70	0.52	—	
		48	0.65	0.39	0.66	0.41	0.87	0.03–13.8
		72	0.81	0.87	*0.77**	0.74	—	
		96	0.88	1.17	0.81	0.87	—	
		144	*0.99**	2.32	0.80	0.84	—	
		192	1.00	Inf	0.88	1.17	—	
		288	—	—	1.00	Inf	—	
PJG	L	12	0.52	0.05	—	—		
		24	0.59	0.20	—	—		
		48	0.62	0.30	0.49	−0.02	—	
		96	0.72	0.58	0.73	0.61	0.90	0.11–6.05
		120	*0.79**	0.80	0.75	0.67	—	
		144	*0.83**	0.95	0.96	1.75	**0.30**	**0.00.–0.78**
		192	0.87	1.12	—	—		
		288	1.00	Inf	—	—		
ave ω_P							0.32	

Table 1a.
(*Continued*)

O	C	D	P for z_V	z_V	P for z_M	z_M	$\omega_P = z_V^2/z_M^2$	95% CI
	R	...6	0.44	−0.15	—	—		
		12	0.50	0.00	0.56	0.15	—	
		24	0.60	0.25	0.61	0.27	0.82	0.00–783.2
		48	0.79	0.80	0.68	0.46	—	
		96	0.96	1.75	0.92	1.40	—	
		144	—	—	1.00	Inf	—	
		192	0.96	1.75	1.00	Inf	—	
GBH	R	12	—	—	0.56	0.15		
		24	0.58	0.20	0.64	0.35	0.32	0.0–5e+14
		48	0.76	0.70	0.84	0.99	0.50	0.0–2.08
		64	*0.80**	0.84	0.92	1.40	**0.36**	**0.11–0.97**
		96	0.88	1.17	0.94	1.55	0.57	0.01–1.57
		192	1.00	Inf	—	—		
ave ω_P							0.42	

NOTE: 'L' indicates data with Lines, 'R,' data with Rectangles. Values in italics were interpolated from the best fitting psychometric functions. The table also shows corresponding z-values for binocular and monocular viewing: z_V, z_M and $\sqrt{2}z_M$. Where possible, the last two columns give the weights for perspective, ω_P, with their corresponding 95% confidence intervals, again for perfectly correlated and independent eyes, respectively. Weights shown in **bold** type lie within the *a priori* [0, 1] range.

within that range. It is difficult to see how the confidence intervals could easily be reduced. Each block of monocular and binocular trials was usually repeated 3 or 4 times (75–100 observations) and a very large number of additional repetitions would be necessary to reduce most 95% confidence limits to lie in the [0, 1] range and then only for the rather few conditions in which the mean values of the weights fall within [0, 1].

On the basis of the results we might, however, venture a tentative interpretation: the two observers with poor stereoacuity in the zero-vertical-scaling condition with rectangles (AMZ and PJG) appear to give more weight to perspective cues with closed configurations than with the open configuration. Between-observer comparisons do not support the hypothesis that the group of observers who might have been expected to depend heavily on perspective cues actually weight perspective more than the disparity-dependent observer. No definite conclusions could be drawn, however, because of the large variability associated with the estimated weights.

It seems that for whatever reason, monocular performance does not provide a useful estimate of sensitivity to perspective cues. It is possible that one or more of the assumptions underlying the cue-combination analysis employed may not be correct, or not correct for some observers. In a recent study of slant perception with linear perspective and texture cues to depth, it has been reported that both the optimality of cue combination, and even the degree of correlation of cues, vary

Table 1b.
Each configuration (column C) 'equivalent disparity differences' (column D) for four observers with the corresponding proportion correct (P for z_V) in binocular viewing, and proportion correct, for z_M in monocular viewing assuming independent eyes, for $\sqrt{2}z_M$

O	C	D	P for z_V	z_V	P for z_M	$\sqrt{2}z_M$	P for $\sqrt{2}z_M$	$\omega_P = z_V^2/z_M^2$	95% CI
LIB	L	6	0.44	−0.15	—	—	—		
		12	0.58	0.20	—	—	—		
		24	0.56	0.15	0.44	−0.21	0.42	—	
		48	0.62	0.30	0 54	0.14	0.55	—	
		96	0.78	0.77	0.56	0.21	0.58	—	
		144	_0.90_*	1.28	0.72	0.82	0.79	—	
		192	0.96	1.75	0.94	2.19	2.19	0.63	0.02–17.68
	R	3	0.48	−0.05	—	—	—		
		6	0.52	0.05	—	—	—		
		12	0.61	0.27	0.42	−0.28	0.39		
		24	0.67	0.43	0.58	0.28	0.61	—	
		48	0.80	0.84	0.82	1.29	0.90	0.42	0.11–1.25
		96	0.99	2.32	0.84	1.40	0.91	—	
		192	—	—	1.00		—		
AMZ	L	24	0.52	0.05	—	—	—		
		36	—	—	0.54	0.14	0.55		
		48	0.62	0.30	0.66	0.58	0.71	0.19	0.0–7.46
		72	_0.59_*	0.22	0.67	0.62	0.73	0.01	0.0–2.10
		96	0.64	0.35	0.73	0.90	0.81	**0.08**	**0.01–0.82**
		144	_0.77_*	_0.73_*	0.83	1.29	0.91	**0.20**	**0.09–0.84**
		192	0.85	1.03	0.96	2.47	0.99	**0.16**	**0.01–0.64**
		288	1.00	Inf	0.96	2.47	0.99		
ave ω_P								0.15	
	R	12	—	—	0.62	0.43	0.66		
		24	0.44	−0.15	0.70	0.74	0.77		
		48	0.65	0.39	0.66	0.66	0.74	0.29	0.01–2.06
		72	0.81	0.87	_0.77_*	0.90	0.81	—	
		96	0.88	1.17	0.81	1.19	0.84	—	
		144	_0.99_*	2.32	0.80	1.66	0.86	—	
		192	1.00	Inf	0.88	1.15	0.84	—	
		288	—	—	1.00	—	—		
PJG	L	12	0.52	0.05	—	—	—		
		24	0.59	0.20	—	—	—		
		48	0.62	0.30	0.49	−0.03	0.49		
		96	0.72	0.58	0.73	0.86	0.80	0.45	0.06–2.13
		120	_0.79_*	0.80	0.75	0.95	0.83	0.71	0.17–2.54
		144	_0.83_*	0.95	0.96	2.47	0.99	**0.15**	**0.01–0.56**
		192	0.87	1.12	—	—	—		
		288	1.00	Inf	—	—			
ave ω_P								0.59	

Table 1b.
(*Continued*)

O	C	D	P for z_V	z_V	P for z_M	$\sqrt{2}z_M$	P for $\sqrt{2}z_M$	$\omega_P = z_V^2/z_M^2$	95% CI
	R	…6	0.44	−0.15	—	—	—		
		12	0.50	0.00	0.56	0.21	0.58		
		24	0.60	0.25	0.61	0.39	0.65	0.41	0.0–41.40
		48	0.79	0.80	0.68	0.66	0.74	—	
		96	0.96	1.75	0.92	1.98	0.97	0.78	0.02–21.54
		144	—	—	1.00	—	—		
		192	0.96	1.75	1.00	—	—		
GBH	R	12	—	—	0.56	0.21	0.58		
		24	0.58	0.20	0.64	0.50	0.69	0.16	0.00–9.25
		48	0.76	0.70	0.84	1.40	0.91	0.25	**0.06–0.92**
		64	*0.80**	0.84	0.92	1.98	0.97	0.18	**0.01–0.54**
		96	0.88	1.17	0.94	2.19	0.98	0.29	**0.01–0.81**
		192	1.00	Inf	—	—	—		
ave ω_P								0.22	

NOTE: 'L' indicates data with Lines, 'R,' data with Rectangles. Values in italics were interpolated from the best fitting psychometric functions. The table also shows corresponding z-values for binocular and monocular viewing: z_V, z_M and $\sqrt{2}z_M$. Where possible, the last two columns give the weights for perspective, ω_P, with their corresponding 95-% confidence intervals, again for perfectly correlated and independent eyes, respectively. Weights shown in **bold** type lie within the *a priori* [0, 1] range.

across observers (Oruç *et al.*, 2003). The large individual differences found in cue-combination studies suggest that human observers differ in their cue-weighting strategies and it may be that there is no single model to account for all behaviour, especially when cues to depth are few and in conflict.

6. ISOLATING DISPARITY CUES

Finally an attempt to isolate disparity cues was made. Any stimulus with non-zero vertical extent would have a perspective cue even if that cue did not change from one observation interval to another. Thus one possibility would be to use points to measure disparity sensitivity. It seems very likely, however, that disparity sensitivity with point stimuli would underestimate disparity sensitivity with lines (McKee, 1983).

We have taken a different approach and assume that, in the absence of vertical scaling, sensitivity to horizontal disparity with lines gives a direct estimate of sensitivity to horizontal disparity that is very little contaminated by perspective cues and is appropriate for our stimuli. There are several observations to justify such an assumption: (1) Comparison of depth-discrimination performance with lines in the slanted plane and zero-vertical-scaling conditions shows no difference in three out of four observers. (The exception is an observer with very poor disparity sensitivity.) This suggests that, with lines, the perspective cues would be given almost no weight except for observers who are stereo weak. (2) The reason why observers with good disparity sensitivity would give perspective cues with lines negligible weight is that the magnitude of the differences based on foreshortening with lines at depths corresponding to disparity thresholds are too small to be detectable; that is, they fall so far down the psychometric function for either binocular or monocular viewing that the probability of correctly detecting the difference in line length at the threshold in the slanted-plane condition would be no better than guessing. Thus any use of perspective cues with lines in the slanted-plane condition seems unlikely. In that case, zero weight for perspective would be assigned on the basis of normal experience with line stimuli. (Observers can, of course, discriminate foreshortening differences in the absence of horizontal disparity but they require differences that are 8 times larger than those yielding 75% correct responses when disparity is present. Differences that large can be discriminated with sequentially presented stimuli where disparity is not present (Ono, 1967) so that depth-sensitive mechanisms may not even be involved in the judgment.)

If we take the discrimination performance using lines with no foreshortening as a measure of disparity sensitivity, z_{D}, we can solve equation (17) for ω_{D} by taking the ratio $z_{\mathrm{H}}^2/z_{\mathrm{D}}^2$. The results of these calculations are shown in Table 2 where, for five different observers and a range of disparities, ω_{D} values are given together with their 95% confidence intervals. (The results for an additional observer, TCC, who has good stereoscopic vision are included. Like observer LIB, her performance when discriminating depth with rectangles in the absence of vertical scaling was only a factor of 3 worse than when the cues were consistent.) In Table 2, values for which the 95% confidence interval lies within the *a priori* range are shown in bold-face type.

Weights for disparity when the observers are viewing the closed form range from 0.28 for LIB, who has the most acute disparity sensitivity among observers, to 0.02 for GBH who has very poor disparity sensitivity. The remaining observers appear to have intermediate weights and the weights given to disparity appear to decrease as the observers' disparity sensitivity decreases, at least when we consider the weights having 95% confidence intervals within the [0, 1] range. However, even the 95% confidence intervals that are less than the *a priori*-constrained range are large.

7. SUMMARY

We have measured stereoacuity with the same observers in four different conditions and combined a detection-theory analysis of our two-alternative forced-choice tasks with an analysis of the way in which the observers might combine information from different cues to depth. Our experimental results are consistent with earlier findings (McKee, 1983; Westheimer, 1979) in that relative depth estimation with rectangles was markedly worse than with lines when horizontal disparity is introduced without changing vertical scaling. However, stereo thresholds for 'rectangles' improved dramatically when perspective cues consistent with horizontal disparity

Table 2.
The proportion correct for z_H and the proportion correct for z_D for five observers (column O) and a range of disparities (column D)

O	D	P for z_H zero-VS rectangles	P for z_D zero-VS rectangles	$\omega_D = z_H^2/z_D^2$	95% CI
LIB	3	0.32	0.42	—	
	6	0.56	0.68	0.01	0.0–4.97
	12	0.63	0.76	0.22	0.0–1.64
	24	0.75	0.90	**0.28**	**0.057–0.89**
	48	0.88	1.00	—	
	96	0.96	—	—	
TCC	3	—	0.52	—	
	6	0.52	0.70	0.009	0.0–2.10
	12	0.65	0.85	**0.13**	**0.0–0.70**
	24	0.68	0.92	**0.11**	**0.01–0.45**
	48	0.93	1.00		
AMZ	3	—	0.54	—	
	6	—	0.72	—	
	12	0.49	0.64	—	
	24	0.67	0.82	0.23	0.01–1.98
	48	0.62	0.94	**0.04**	**0.0–0.45**
	96	0.64	—	—	
	192	0.56	—	—	
PJG	6	0.50	0.49	—	
	12	0.52	0.56	0.11	0.01–33379
	24	0.58	0.72	0.11	0.00–3.09
	48	0.61	0.82	0.09	0.00–1.30
	96	0.70	0.88	0.19	0.02–1.30
	192	0.58	0.93	**0.02**	**0.00–0.35**
GBH	12	0.54	0.62	0.10	0.00–5.7e+14
	24	0.52	0.56	0.10	0.00–9.8e+14
	48	0.58	0.66	0.24	0.00–4.2e+14
	96	0.54	0.78	**0.02**	**0.00–0.77**
	192	0.76	0.88	0.36	0.04–1.64

NOTE: Where possible, the last two columns show weights for disparity, ω_D, and the corresponding 95% confidence intervals. Weights shown in **bold** type lie within the *a priori* [0, 1] range.

were present. (When only perspective cues were manipulated, all observers performed better with 'rectangles' than lines by a factor of between 2 and 3.) Substantial observer variability was found in this as in other studies. It appears that some observers are disparity-dependent and others rely more on perspective cues when estimating relative depth. Furthermore, the observers may switch the weighting they give to different cues between stimulus conditions. We infer from our results that: (1) observers use many sources of information in making relative depth judgments; (2) they sometimes use sources of information that, although they may be helpful in normal visual experience, hurt their performance in some relative-depth-judgment tasks; (3) one such task is depth judgment of virtual images in which vertical scaling (and hence perspective) is not adjusted to correspond to horizontal disparity; (4) different stimulus configurations produce changes in the way observers combine information from different cues — in particular, cues based on perspective are given greater weight in depth judgments with closed figures than with pairs of lines. Estimates of the weights assigned to different cues often have confidence intervals that are very large indeed.

Acknowledgements

This research was supported by the Wellcome Trust and AMZ was supported by grants from the British Federation of Women Graduates and the Overseas Research Student Award Scheme. This work was part of the material submitted by AMZ for the degree of Doctor of Philosophy in the University of Oxford (Zalevski, 2004). We should like to thank Mrs. Sue Fowles at the Dyslexia Research Centre, The University Laboratory of Physiology, University of Oxford for measuring the observers' phorias. We are grateful to Mrs. L. Browning, Miss T. Cristescu, and Ms P. Gallop for many hours of patient and careful observing. We are also grateful to Dr. F. A. Wichmann and two reviewers for helpful suggestions.

REFERENCES

Allison, R. S. and Howard, I. P. (2000). Temporal dependencies in resolving monocular and binocular cue conflict in slant perception, *Vision Research* **40**, 1869–1885.

Andersen, E. E. and Weymouth, F. W. (1923). Visual perception and the retinal mosaic: 1. Retinal mean local sign: an explanation of the fineness of binocular perception of distance, *Amer. J. Phys.* **64**, 561–594.

Backus, B. T. and Banks, M. S. (1999). Estimator reliability and distance scaling in stereoscopic slant perception, *Perception* **28**, 217–242.

Berry, R. N. (1948). Quantitative relations among Vernier, real depth, and stereoscopic depth acuities, *J. Exper. Psych.* **38**, 708–721.

Bourdon, B. (1900). L'acuité steréoscopique, *Revue Philosophique* **49**, 74–78.

van Ee, R. and Erkelens, C. J. (1998). Temporal aspects of binocular slant perception: an evaluation and extension of Howard and Keneko's theory, *Vision Research* **38**, 3871–3882.

van Ee, R., Banks, M. S. and Backus, B. T. (1999). An analysis of binocular slant contrast, *Perception* **28**, 1121–1145.

van Ee, R., van Dam, L. C. J. and Erkelens, C. J. (2002). Bi-stability in perceived slant when binocular disparity and monocular perspective specify different slants, *Journal of Vision* **2**, 597–607.

van Ee, R., Adams, W. J. and Mamassian, P. (2003). Bayesian modelling of perceived slant in bi-stable stereoscopic perception, *J. Opt. Soc. Amer.* **20**, 1398–1406.

Epstein, W. (1966). Perceived depth as a function of relative height under three background conditions, *J. Exper. Psychol.* **72**, 335–338.

Epstein, W. and Landauer, A. H. (1969). Size and distance judgments under reduced conditions of viewing, *Perception and Psychophysics* **6**, 169–272.

Ernst, M. O. and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion, *Nature* **415**, 429–433.

Ernst, M. O., Banks, M. S. and Bülthoff, H. H. (2000). Touch can change visual slant perception, *Nature Neuroscience* **3**, 69–73.

Gillam, B. (1968). Perception of slant when perspective and stereopsis conflict: experiments with aniseikonic lenses, *J. Exper. Psychol.* **78**, 299–305.

Gillam, B. and Ryan, C. (1992). Perspective, orientation disparity, and anisotropy in stereoscopic slant perception, *Perception* **21**, 427–439.

Gillam, B. and Ryan, C. (1993). Stereoscopic slant reversals: a new kind of 'induced' effect, *Perception* **22**, 1025–1036.

Gogel, W. C. (1969). The absolute and relative size cues to distance, *Amer. J. Psychol.* **82**, 228–234.

Green, D. M. and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. Wiley, New York.

Hillis, J. M., Ernst, M. O., Banks, M. S. and Landy, M. S. (2002). Combining sensory information: mandatory fusion within, but not between, senses, *Science* **298**, 1627–1630.

Hillis, J. M., Watt, S. J., Landy, M. S. and Banks, M. S. (2006). Slant from texture and disparity cues: optimal cue combination, *Journal of Vision* **4**, 967–992.

Howard, H. J. (1919). A test for the judgment of distance, *Amer. J. Ophthalmol.* **2**, 656–675.

Howard, I. P. (2002). Depth perception, in: *Stevens Handbook of Experimental Psychology*, Pashler, H. and Yantis, S. (Eds), Vol. VI, pp. 77–120. Wiley, New York.

Jacobs, R. A. (1999). Optimal integration of texture and motion cues to depth, *Vision Research* **39**, 3621–3629.

Knill, D. C. and Saunders, J. (1999). Humans optimally weight stereo and texture cues to estimate surface slant, *Journal of Vision* **2**, 3621–3629.

Knill, D. C. and Saunders, J. (2003). Do humans optimally integrate stereo and texture information for judgments of surface slant, *Vision Research* **43**, 2539–2558.

Kumar, T. and Glaser, D. A. (1992). Shape analysis and stereopsis for human depth perception, *Vision Research* **32**, 499–512.

Landy, M. S., Maloney, L. T., Johnson, E. B. and Young, M. (1995). Measurement and the modelling of depth cue combination: in defence of weak fusion, *Vision Research* **35**, 389–412.

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

Mitchison, G. J. and Westheimer, G. (1984). The perception of depth in simple figures, *Vision Research* **24**, 1063–1073.

Ogle, K. N. (1953). Precision and validity of stereoscopic depth perception from double images, *J. Opt. Soc. Amer.* **43**, 906–913.

Ono, H. (1967). Difference threshold for stimulus length under simultaneous and nonsimultaneous viewing conditions, *Perception and Psychophysics* **2**, 201–207.

Oruç, I., Maloney, L. T. and Landy, M. S. (2003). Weighted linear cue combination with possibly correlated error, *Vision Research* **43**, 2451–2468.

Papoulis, A. and Pillai, S. U. (2002). *Probability, Random Variables and Stochastic Processes*, 4th edition. McGraw-Hill, Boston.

Richards, W. (1970). Stereopsis and stereoblindness, *Experimental Brain Research* **10**, 380–388.

Richards, W. (1971). Anomalous stereoscopic depth perception, *J. Opt. Soc. Amer.* **61**, 410–414.

Rosas, P., Wichmann, F. A. and Wagemans, J. (2004). Some observations on the effects of slant and texture type on slant-from-texture, *Vision Research* **44**, 1511–1535.

Rosas, P., Wagemans, J., Ernst, M. O. and Wichmann, F. A. (2005). Texture and haptic cues in slant discrimination: reliability-based cue weighting without statistically optimal cue combination, *J. Opt. Soc. Amer. A* **22**, 801–809.

Ryan, C. and Gillam, B. (1994). Cue conflict and stereoscopic surface slant about horizontal and vertical axes, *Perception* **23**, 645–658.

Sato, M. and Howard, I. P. (2001). Effects of disparity-perspective cue conflict on depth contrast, *Vision Research* **41**, 415–426.

Schlosberg, H. (1950). A note on depth perception, size constancy, and related topics, *Psychol. Rev.* **57**, 314–317.

Sivia, D. S. (1996). *Data Analysis: A Bayesian Tutorial*, 1st edition. The Clarendon Press, Oxford.
Stevens, K. A. and Brookes, A. (1988). Integrating stereopsis with monocular interpretations of planar surfaces, *Vision Research* **28**, 371–386.
Strang, G. (1988). *Linear Algebra and Its Applications.* Harcourt Brace Jovanovich, San Diego, CA.
Vollmerhausen, R. H. and Driggers, R. G. (2000). *Analysis of Sampled Imaging Systems.* SPIE Press, Washington, DC.
Werner, H. (1937). Dynamics in binocular depth perception, *Psychological Monographs* **49**, 1–127.
Westheimer, G. (1979). The spatial sense of the eye, *Investig. Ophthalmol. Visual Science* **18**, 893–912.
Westheimer, G. and McKee, S. P. (1978). Stereoscopic acuity for moving retinal images, *J. Opt. Soc. Amer.* **68**, 450–455.
Wichmann, F. A. and Hill, N. J. (2001a). The psychometric function: I. Fitting, sampling, and goodness-of-fit, *Perception and Psychophysics* **63**, 1293–1313.
Wichmann, F. A. and Hill, N. J. (2001b). The psychometric function: II. Bootstrap-based confidence intervals and sampling, *Perception and Psychophysics* **63**, 1314–1329.
Young, M. J., Landy, M. S. and Maloney, L. T. (1993). A perturbation analyssis of deppth perception from combinations of texture and motion cues, *Vision Research* **33**, 2685–2696.
Youngs, W. M. (1976). The influence of perspective and disparity cues on the perception of slant, *Vision Research* **16**, 79–82.
Zalevski, A. M. (2004). Some determinants of the loss in relative depth with closed configurations. D.Phil. thesis, Oxford.

APPENDIX A. CALCULATION OF 95% CONFIDENCE LIMITS FOR WEIGHTS

The binomial probabilities associated with the number of trials and estimated probabilities of correct responses associated with a given ‘disparity’ in the binocular zero-horizontal-disparity condition and in the monocular zero-horizontal-disparity condition were determined. (The associated z-values enter into the numerator and denominator of equations (19) or (20).) Performance in the two tasks is assumed to be independent so that the product of the binomial probabilities is the joint probability of getting a particular number correct in the condition associated with the numerator and a particular number correct in the condition associated with the denominator. Then the region containing 0.95 of the resulting (2-D) probability mound was found. Next we calculated the ratios of the numerators, z_{V}^2, to the denominator, z_{M}^2, associated with each possible pair of number correct. Finally, the smallest and largest ratios in the region containing 0.95 of the probability mound was taken as the estimate of the 95% confidence limits for the ratio.