

Stereo vision requires an explicit encoding of vertical disparity

Ignacio Serrano-Pedraza

Institute of Neuroscience, Newcastle University,
Newcastle upon Tyne, UK



Jenny C. A. Read

Institute of Neuroscience, Newcastle University,
Newcastle upon Tyne, UK



Vertical disparities influence the perception of 3D depth, but little is known about the neuronal mechanisms underlying this. One possibility is that these perceptual effects are mediated by an explicit encoding of two-dimensional disparity. Recently, J. C. A. Read and B. G. Cumming (2006) pointed out that current psychophysical and physiological evidence is consistent with a much more economical one-dimensional encoding. Almost all relevant information about vertical disparity could in theory be extracted from the activity of purely horizontal-disparity sensors. Read and Cumming demonstrated that such a 1D system would experience Ogle's induced effect, a famous illusion produced by vertical disparity. Here, we test whether the brain employs this 1D encoding, using a version of the induced effect stimulus that simulates the viewing geometry at infinity and thus removes the cues which are otherwise available to the 1D model. This condition was compared to the standard induced effect stimulus, presented on a frontoparallel screen at finite viewing distance. We show that the induced effects experienced under the two conditions are indistinguishable. This rules out the 1D model proposed by Read and Cumming and shows that vertical disparity, including sign, must be explicitly encoded across the visual field.

Keywords: binocular vision, vertical disparity, induced effect

Citation: Serrano-Pedraza, I., & Read, J. C. A. (2009). Stereo vision requires an explicit encoding of vertical disparity. *Journal of Vision*, 9(4):3, 1–13, <http://journalofvision.org/9/4/3/>, doi:10.1167/9.4.3.

Introduction

Because our eyes are offset horizontally, objects at different distances have horizontal disparities between their images on the two retinæ. Even in the absence of other depth cues, horizontal disparity suffices to produce a powerful impression of 3D depth. However, it has been known since the nineteenth century (Helmholtz, 1925) that vertical disparities also occur and can influence perception. A famous example is Ogle's induced effect (Ogle, 1938, 1964), in which subjects view a frontoparallel plane with a meridional size lens placed in front of one eye, so as to magnify its image in a vertical direction. This produces the impression that the plane has been rotated about a vertical axis, so that it is now closer to the viewer on the side of the magnified eye. What are the neuronal processes that underlie such percepts? At first sight, it seems obvious that the brain must explicitly encode two-dimensional disparity. It is now thoroughly documented that early visual cortex contains disparity-tuned neurons tuned to a range of horizontal disparities in close agreement with the range of perception (Cumming & DeAngelis, 2001; Parker, Cumming, & Dodd, 2000). Analogously, it is often assumed that neurons must be tuned to a range of vertical disparities, mirroring the range of vertical disparities which influence perception. In this view, the distribution of preferred 2D disparities within

visual cortex must resemble that sketched in Figure 1A. Here, each dot represents the preferred disparity of a different neuron, with in general both a horizontal and a vertical component. However, relatively few studies have examined this distribution, and those that have are difficult to interpret, as we discuss below.

Recently, Read and Cumming (2006) argued that the documented perceptual effects of vertical disparity do not necessarily require the 2D distribution of Figure 1A. Even if all neurons were tuned to zero vertical disparity, as sketched in Figure 1B, their finite receptive field size means that they would continue to respond in the presence of small amounts of vertical disparity, consistent with the fact that subjects can still perceive depth from horizontal disparity in this situation (Stevenson & Schor, 1997). Furthermore, such a population also implicitly encodes the magnitude—though not the sign—of vertical disparity. The key insight is that, for sensors tuned to zero vertical disparity, a non-zero vertical disparity is equivalent to a reduction in binocular correlation. Let us idealize disparity-tuned neurons as detectors of interocular correlation (Qian & Zhu, 1997). For a stimulus with uniform horizontal disparity and no vertical disparity, the sensor tuned to the disparity of the stimulus will report a correlation of 1. As vertical disparity is introduced, this sensor will report progressively less correlation, but it will continue to report more correlation than its colleagues. The steepness with which reported correlation declines

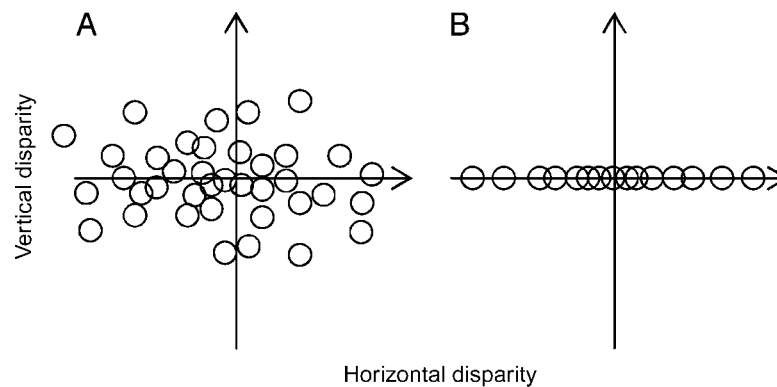


Figure 1. Hypothetical distributions of disparity tuning. Circles show preferred 2D disparity of a neuron in early visual cortex. (A) 2D distribution: The population includes neurons tuned to a range of both horizontal and vertical disparities. The distribution is shown concentrated on zero horizontal disparity, to account for the higher stereoacuity close to fixation, and also on zero vertical disparity, to account for the predominance of vertical disparities close to zero in normal viewing. (B) 1D distribution postulated by Read and Cumming (2006). The neurons are now located along the epipolar lines of primary position.

with vertical disparity V depends on the receptive field size σ .

If the stimulus correlation is C , stimulus horizontal and vertical disparity are H and V , respectively, then a sensor tuned to a horizontal disparity H_{pref} will report an effective interocular correlation of $C_{\text{eff}} = C \exp(-0.25 [(H - H_{\text{pref}})^2 + V^2] / \sigma^2)$ (Read & Cumming, 2006). Thus, the response of a single neuron in this model confounds horizontal disparity, vertical disparity, and interocular correlation, but the three can be distinguished by their different effects on the population as a whole. Horizontal disparity can be read off from the preferred disparity of the maximally responding sensors in the population, while the magnitude of vertical disparity can be deduced from the effective interocular correlation sensed by these maximally responding sensors. Vertical disparity can be distinguished from a genuine reduction in stimulus interocular correlation because vertical disparity affects predominantly the smallest receptive fields, whereas reductions in interocular correlation affect all scales equally. However, only the magnitude of vertical disparity, $|V|$, can be deduced from the activity in this 1D population, not the vertical disparity V itself. Read and Cumming (2006) showed that this suffices to explain illusions such as the induced effect. In fact, they pointed out that under most circumstances the fully signed vertical disparity, V , can be deduced from a knowledge of how the magnitude $|V|$ varies across the retina. In normal viewing, the pattern of vertical disparity across the retina is highly constrained by viewing geometry. Figures 2A and 2B show vertical disparity fields for two example eye positions. In each case, vertical disparity is zero along the horizontal retinal meridian and along a vertical line whose position depends on the gaze azimuth. The convergence angle controls the rate at which vertical disparity magnitude increases away from the “cross” formed by these two lines. Vertical disparity is positive in the 1st and 3rd, and negative in the 2nd and 4th, quadrants of this cross.

Figures 2C and 2D show the effective interocular correlation sensed by the population tuned to the stimulus horizontal disparity, and Figures 2E and 2F show the vertical disparity magnitude reconstructed from this effective correlation. Although this only gives the magnitude, not the sign of vertical disparity, the sign can be deduced from the position relative to the “cross” of zero vertical disparity, as indicated by the symbols. Thus, the sign of vertical disparity anywhere in the retina can usually be deduced from the overall pattern of vertical disparity magnitude.

The suggestion that the brain encodes 2D disparity with a 1D population is at first sight counter-intuitive, but nevertheless worth consideration, since such a scheme would be highly efficient. In natural viewing, disparities are overwhelmingly horizontal (Hibbard, 2007; Read & Cumming, 2004). Thus, the 2D distribution of Figure 1A seems wasteful: it requires the brain to build and maintain a population of sensors tuned to vertical disparities that hardly ever occur. If these cells resembled conventional energy-model units, they would fire at half their maximal rate to even binocularly uncorrelated stimuli. Thus, although these cells would rarely reach peak firing rate (because their preferred non-zero vertical disparities are relatively rare), they would still incur substantial energetic costs. Even if a threshold was imposed to ensure they were silent until exposed to their preferred vertical disparity, the brain would still incur costs in maintaining the population, and in giving valuable cortical space over to it. Surely these costs would have to bring a real benefit. Traditionally, this has been assumed to be the measurement of vertical disparity. But Read and Cumming (2006) showed that this information, up to sign, could be produced with the much more economical 1D distribution shown in Figure 1B. This raises the question of what benefits a 2D distribution might provide to merit the costs—a question we return to in the Discussion.

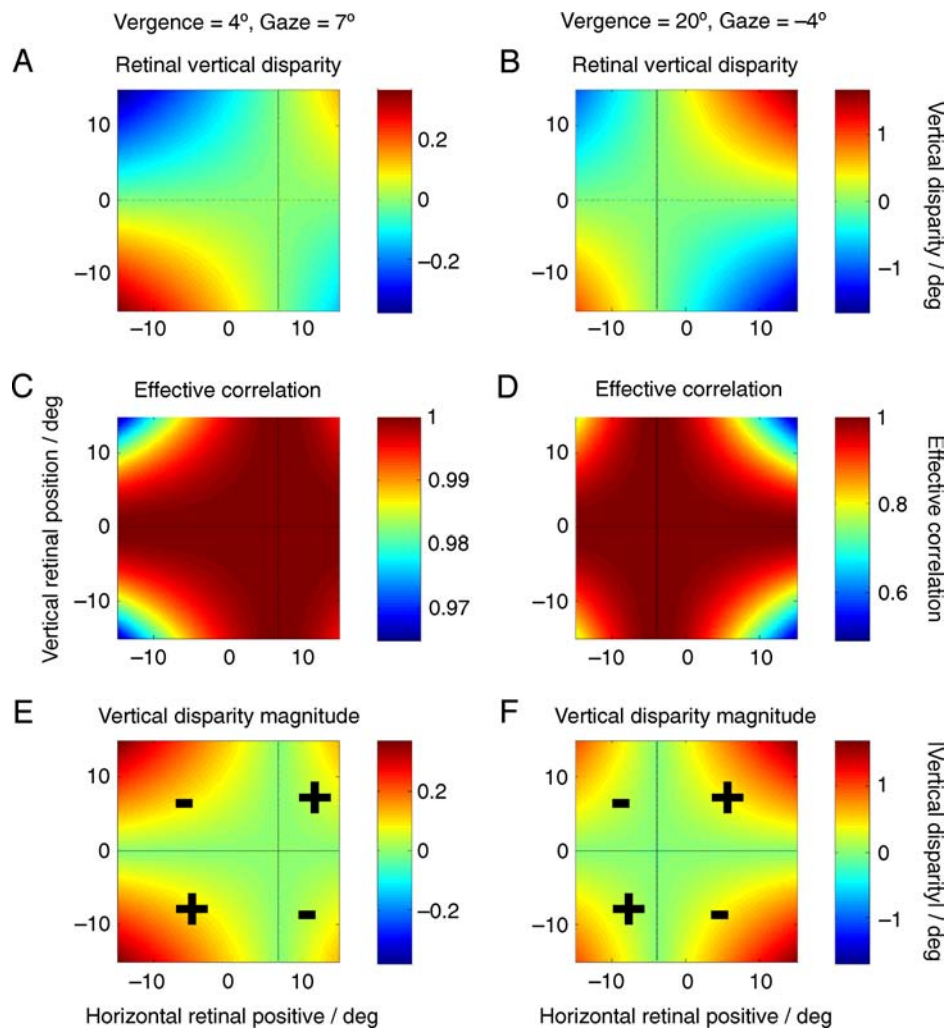


Figure 2. (AB) Vertical disparity field for two different viewing positions. (CD) The effective interocular correlation sensed by neurons tuned to the horizontal disparity of the stimulus, $C_{\text{eff}} = \exp(-0.25V^2 / \sigma^2)$. (EF) Magnitude of vertical disparity deduced from this activity, $|V| = 2\sigma\sqrt{-\ln C_{\text{eff}}}$. The cross-shaped locus of zero vertical disparity, or equivalently of unit effective correlation, is marked with a thin black line. The sign of vertical disparity at each point in the retina can be deduced from the position relative to this cross, as indicated by the + and – symbols.

Ultimately, the question of how disparity tuning is distributed in visual cortex is one for physiology. However, the current physiological literature is inconclusive. Some workers have reported cells tuned to clearly non-zero vertical disparities, supporting a 2D distribution (Durand, Celebrini, & Trotter, 2007; Durand, Zhu, Celebrini, & Trotter, 2002; Gonzalez, Relova, Perez, Acuña, & Alonso, 1993; Trotter, Celebrini, & Durand, 2004), while others have reported essentially no cells tuned to disparities significantly different from zero (Cumming, 2002; Gonzalez, Justo, Bermudez, & Perez, 2003; Maunsell & Van Essen, 1983; Poggio, 1995). One obvious reason for conflict concerns the range of eccentricities used. Vertical disparities introduced by unusual gaze postures increase with eccentricity (Rogers & Bradshaw, 1993), so detectors designed to detect such patterns would be expected to be found predominantly at large eccentricities (in contrast to detectors designed to drive corrective eye movements,

see below). This may be, for example, why Durand et al. (2002), studying eccentricities $>10^\circ$, reported cells tuned to a wider range of vertical disparities than Cumming (2002), studying $2\text{--}9^\circ$. However, the same factor also applies to a possible artefact. To appreciate this, it becomes necessary to define more precisely what we mean by “vertical disparity.”

Here and previously (Read & Cumming, 2004, 2006), we have adopted the definition of Longuet-Higgins (1982), using a Cartesian coordinate system fixed *on the retina*. This is convenient, given that cells in early visual cortex encode visual information in retinotopic coordinates. In this system, the directions “horizontal” and “vertical” on the retina are defined when the eyes are in *primary position*, i.e., looking straight ahead to infinity (Figure 3A). With the eyes in primary position, the two retinal images of an object, such as the black dot at the corner of the square in Figure 3A, differ only in

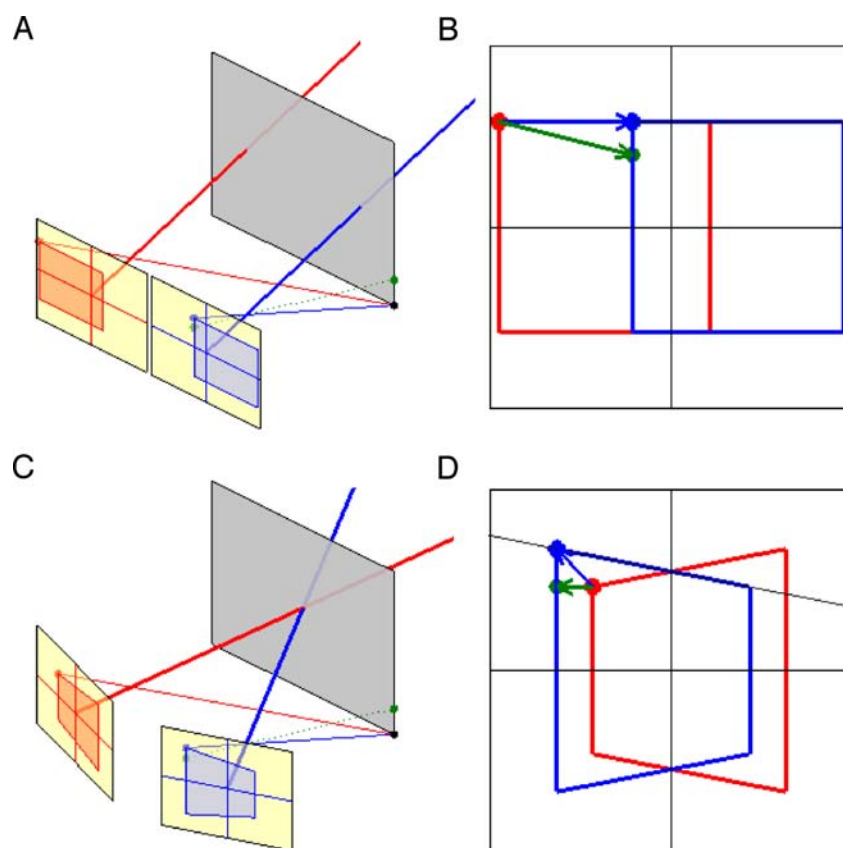


Figure 3. Retinal images produced by a square stimulus, viewed (AB) with the eyes in primary position, and (CD) converged so as to fixate at the center of the square. Diagram BD shows the two retinal images superimposed. The vector indicates the disparity of the dot on the bottom-right corner of the square. Viewed with convergence, it has a vertical component on the retina (CD). A stimulus with artificial (non-epipolar) vertical disparity is also shown. We envisage an experimental situation, in which polarizing filters or similar are used to ensure only the right eye views the black dot, while the left eye views the green dot. Clearly, this stimulus has vertical disparity on the screen; in primary position, it also has vertical disparity on the retina (green vector in B). However, when the eyes converge, the experimentally applied vertical disparity cancels out the vertical disparity which would normally be experienced at this eccentricity, resulting in zero vertical disparity on the retina (green vector in D). For clarity, we have used planar retinas. Since there is a one-to-one mapping between these planes and the real retinas, this does not affect the argument or involve any loss of generality (see Figure 3 of Read & Cumming, 2006). The stimulus is drawn at 8.6 cm from the observer, and the two eyes' nodal points are 6.25 cm apart, so vergence in CD is 40° . The green dot is 1.55 cm above the black dot on the screen, giving an on-screen vertical disparity of 7.6° if we define this as the angle between the lines joining the two dots to the cyclopean point midway between the two nodal points.

their horizontal coordinate in this coordinate system. Thus, whatever an object's position in space, it can have only horizontal disparity on the retina (blue vector in Figure 3B). When the eyes move away from primary position, this is no longer the case. An example is shown in Figures 3C and 3D, where the eyes are converging at 40° . Now, the images of the black dot differ both in their horizontal and vertical coordinates (blue vector in Figure 3D). In other words, the object has a vertical disparity on the retina. However, most physiologists have used “vertical disparity” to refer to vertical displacements on the computer screen used to display the stimuli. This produces a *non-epipolar disparity*, i.e., one which could not be produced by any real object, given the current position of the eyes, but which can be produced experimentally. So for example we might arrange matters

such that the left eye views the black dot at the bottom-right corner of the screen in Figure 3C, but the right eye views the dot color-coded green. Since the two dots are directly above one another on the screen, they have a purely vertical disparity on the screen. But as the green vector in Figure 3D shows, they project to the same vertical position on the retina. Thus, experimentally adding in vertical disparity *on the screen* has produced a vertical disparity *on the retina* of zero.

Suppose then that the physiologist was probing a system like that in Figure 1B, containing only pure horizontal-disparity detectors. When the animal views a frontoparallel screen at a distance of 50 cm, as in Durand et al. (2002), cells at 10° eccentricity experience retinal vertical disparities of up to 0.07° , cells at 20° up to 0.27° , and cells at 30° up to 0.61° (assuming a monkey interocular

distance of 4 cm; smaller values produce smaller estimates). Adding the appropriate *on-screen* vertical disparity would remove this vertical disparity on the retina, enhancing the cells' response. Thus, the fact that Durand et al. (2002) reported preferred vertical disparities up to 0.6° (their Figure 3A) does not enable us to rule out the possibility that all cells were tuned to zero vertical disparity on the retina. To convert a cell's preferred on-screen vertical disparity into its preferred retinal disparity requires a knowledge of its precise location in the visual field, but this is not usually provided. Thus, existing physiological studies tell us little about the distribution of preferred disparities in visual cortex, especially in the periphery where the distinction between screen and retinal vertical disparity is most crucial.

The psychophysics literature is also inconclusive on this point, because studies have not been designed to answer this question. Read and Cumming (2006) probed whether the visual system confounds interocular correlation and vertical disparity, as horizontal-disparity sensors would do. They were not able to demonstrate this, but the failure was not conclusive, because the mapping between correlation and vertical disparity depends on receptive field size and hence on spatial scale. Vertical disparity reduces the effective interocular correlation most markedly at the finest spatial scales, with progressively less effect at lower spatial frequencies. Genuine changes in interocular correlation would affect all scales equally. A sufficiently sophisticated read-out of a 1D population of disparity sensors could potentially use this to distinguish between the two and prevent noise being misinterpreted as vertical disparity.

As noted above, most investigations of vertical disparity have presented stimuli on a frontoparallel screen viewed at a finite distance, so that the screen is at an angle to the optic axis of each eye. As Figures 3C and 3D illustrate, such stimuli produce a vertical disparity field on the retina even when no vertical disparity is applied on the screen. When on-screen vertical disparity is introduced, the interaction between the two sources of vertical disparity produces a characteristic pattern of vertical disparity, enabling the vertical disparity applied to the stimulus to be read off from a purely 1D population of disparity detectors (Read & Cumming, 2006). This means that most psychophysical effects of vertical disparity could in theory be supported by the population of Figure 1B. For example, consider the short-latency corrective vertical vergence movements elicited by stimuli which simulate the effect of a vertical vergence misalignment (Busetini, Fitzgibbon, & Miles, 2001). In this situation, the sign of the vertical misalignment can be deduced from the pattern of the unsigned vertical disparity field. When the eyes are correctly aligned, there should be no vertical disparity along the horizontal meridian of the retina. When the eyes are misaligned, so that they are not fixating on a common point in space, the locus of this zero vertical disparity line shifts either above or below the horizontal meridian,

depending on the sign of the vertical disparity. Thus, a system with no vertical disparity detectors could still sense the misalignment and respond so as to eliminate it (see Figures 13 and 14 of Read & Cumming, 2006). This strategy would fail for stimuli at infinity, but it has not been demonstrated that vertical vergence corrections can still be made under these circumstances.

A few researchers have used haploscopes in which the screen is perpendicular to each optic axis (Backus, Banks, van Ee, & Crowell, 1999; Banks, Hooge, & Backus, 2001). This simulates viewing at infinite distance and produces no vertical disparity on the retina (Figures 3A and 3B) unless vertical disparity is applied on the screen. When the induced effect is applied on this apparatus, the same pattern of vertical disparity magnitude is produced on the retina irrespective of which eye's image is vertically magnified: only the *sign* of the pattern inverts depending on whether the image is expanded or compressed. The 1D population of Figure 1B cannot encode this sign. Yet, the classic induced effect is still experienced in this apparatus, with the direction of perceived slant depending on the sign of magnification (Backus et al., 1999). This at last is evidence that the visual system can measure the sign of vertical disparity, yet even this does not absolutely prove the existence of a 2D distribution of disparity detectors across the visual field. Since the stimuli were presented for several seconds, it remains possible that a signed measurement of vertical disparity is made only at the fovea, and a map of stimulus vertical disparity is built up by fixating different regions of the visual field. It is suggestive in this context that several authors report that vertical disparity illusions depend on long presentations, building up gradually over time (Allison, Howard, Rogers, & Bridge, 1998; Kaneko & Howard, 1997; Ogle, 1938; Westheimer, 1984). The analogous possibility for horizontal disparity was considered when stereopsis was first discovered: "It may be supposed that ... [horizontal disparity] is appreciated by successively directing the point of convergence of the optic axes successively to a sufficient number of its points to enable us to judge accurately of its form" (Wheatstone, 1838). Wheatstone (1838) and countless others have presented compelling evidence that horizontal disparity is not measured solely via vergence, for example the fact that we can perceive multiple horizontal disparities even when a stimulus is presented too briefly to allow eye movements. Only one study to date has used short presentations in this apparatus (Banks et al., 2001). The results suggest that the sign of retinal vertical disparity is detected, but the study was not designed to address this, and this conclusion was not explicitly drawn by the authors.

In this study, therefore, we compared two versions of the induced effect. In each case, the stimulus was presented on a frontoparallel screen viewed at a distance of 165 cm and appeared for 200 ms, too briefly to allow eye movements. We interleaved two different versions of

the induced effect. In the “standard” condition, the image presented to one eye was vertically magnified, simulating the effect of the meridional size lens placed at axis 180° used by Ogle (1938). In the “infinite-distance” condition, the physical screen was still at 165 cm, but additional vertical disparity was added, designed to cancel out the retinal vertical disparity introduced by the finite viewing distance, and thus reproduce the stimulus of Backus et al. (1999), simulating a screen viewed at infinity with a cylindrical lens over one eye. We find that even with this short-duration stimulus, the classic induced effect persists. Both types of stimuli produce the classic induced-effect slant illusion, with no discernible change in threshold, even when presented for only 200 ms. We conclude that the visual system must depend on a genuinely 2D distribution of disparity detectors, which provide it with a signed measure of vertical disparity across the visual field.

Methods

Apparatus

Stimuli were presented on a rear projection screen. Each eye’s image was presented on a separate FX2+ Projection Design DLP (Digital Light Processing) projector with a resolution of 1400 × 1050 pixels (horizontal × vertical). The projectors were held rigidly on an adjustable frame supplied by Virtalis (Manchester, UK). The images were aligned by displaying identical sets of one-pixel-wide gridlines on each projector, red on one and green on the other. The projectors’ physical position, zoom, focus, and vertical lens shift were adjusted until the combined image appeared as a single set of yellow gridlines, indicating that the displays were in alignment. Alignment was better than one pixel over almost the whole screen and was nowhere worse than two pixels. Polarizing filters ensured that each eye saw only one projector’s image. The interocular cross talk, measured with a Minolta LS-100 photometer, was less than 1%. The projection screen was frontoparallel to the observers, who viewed it at a distance of 165 cm using a head and chin rest (UHCOTech HeadSpot). The long viewing distance minimized the vertical disparity introduced by viewing geometry even in the “standard” condition (see Figure 6) and hence minimized any conflict between the distance cues provided by vertical disparity and those provided by accommodation and vergence angle. The projected image was 127 cm × 95 cm (42° × 32°).

Stimuli

In the “standard” condition, we displayed the classic induced effect. White disks, 3 pixels in diameter, were

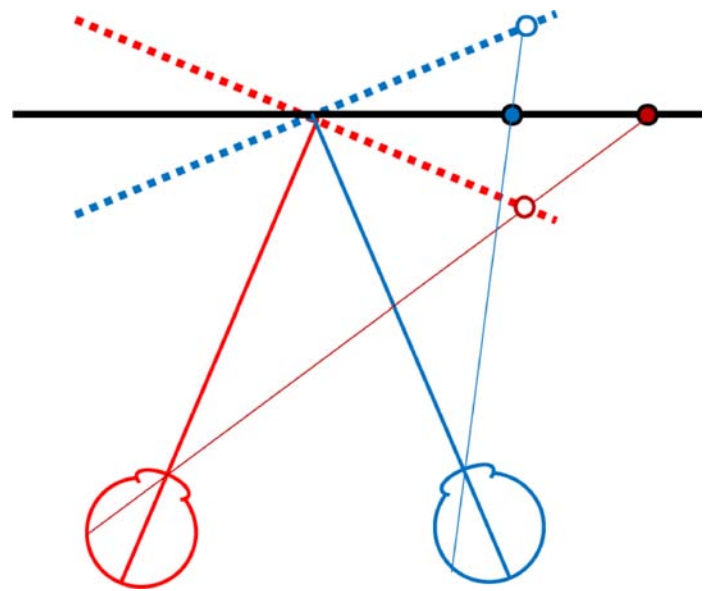


Figure 4. Generating the “infinite-distance” induced effect. The eyes fixate the physical projection screen (heavy black line). The heavy red and blue lines show the optic axes of the left and right eyes, respectively. The dotted lines show the virtual screens, orthogonal to the optic axis, on which we imagine displaying the stimulus. The lighter red and blue lines show how we calculate where to place a dot on the physical projection screen (filled disks) so as to simulate a dot on the virtual screen (open disks).

distributed uniformly and randomly across a black background. The same pattern was displayed to both eyes, except that the vertical position of each dot relative to the observer’s eye position was magnified by a constant factor in one eye relative to the other. In the “infinite-distance” condition, we simulated displaying this classic stimulus on screens orthogonal to each eye’s optic axis, shown with dotted lines in Figure 4. The disks in Figure 4 show how each dot on the virtual screen (open circles) was projected onto the physical screen.

If (X'_L, Y'_L) are the coordinates of a point on the virtual screen perpendicular to the left eye, then this point projects to coordinates (X_L, Y_L) on the physical screen (Figure 5), where

$$\begin{aligned} X_L &= X'_L Z \sec \theta / (Z + X'_L \sin \theta); \\ Y_L &= Y'_L Z / (Z + X'_L \sin \theta); \end{aligned} \quad (1)$$

Z is the viewing distance (165 cm), θ is half the vergence angle, and the origin of both coordinate systems is the fixation point. The analogous equations for the right eye are

$$\begin{aligned} X_R &= X'_R Z \sec \theta / (Z - X'_R \sin \theta); \\ Y_R &= Y'_R Z / (Z - X'_R \sin \theta). \end{aligned} \quad (2)$$

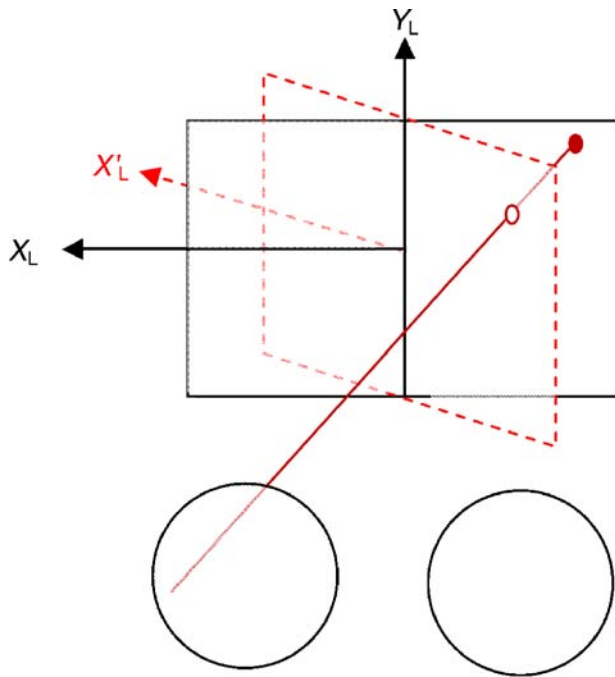


Figure 5. Coordinate system used to represent points on the virtual and physical screens. Y is upwards, X is leftwards. Open red dot shows a point on the left eye's virtual screen (red dashed lines) with coordinates (X'_L, Y'_L) . This projects to the red dot on the physical screen (black lines), with coordinates (X_L, Y_L) .

These equations are derived in [Appendix A](#). They require that the optic plane is perpendicular to the screen, i.e., is horizontal. This was achieved by using a laser spirit-level (Laserliner, Autocross Laser ACL 2) to project a horizontal plane of light passing through the position of the observer's eyes in the head rest, and adjusting the position of the fixation cross to be in the same plane.

The equations also assumed that the projection to the screen is linear. We used the Laserliner to project vertical and horizontal lines onto the screen and verified that rows and columns of pixels were straight and orthogonal. We also projected white squares onto the screen at different locations and measured their physical size using a laser distance meter (Leica Disto A3). In this way, we verified that a given number of pixels projected to the same distance on the screen, both vertically and horizontally and independent of position on the screen, to within measurement error.

The half-vergence angle θ depends on the interocular distance I of the observer: $\theta = \arctan(I / (2Z))$. We used a value of 1.1° for all observers, corresponding to $I = 6.5$ cm. Measured distances for our observers ranged from 6.2 to 6.5 cm, only slightly larger than the error on the measurement. For an observer whose interocular distance is in fact 5.0 cm (the bottom end of the adult distribution; [Dodgson, 2004](#)), running the experiment with $I = 6.5$ cm introduces a maximum error of about a pixel at the edge of the image where the correction is largest. This is similar to the alignment error between the projectors.

The two types of stimuli, and the resulting vertical disparity fields, are shown in [Figure 6](#). In the “standard” condition (6AE), one eye's image is simply stretched vertically on the screen with respect to the other. This produces a pattern of on-screen vertical disparity, which is independent of horizontal location on the screen (6BF). However, even for this large viewing distance (165 cm), the vertical disparity field on the retina is asymmetric, increasing from right to left (6CG). [Read and Cumming \(2006\)](#) showed by simulations that this asymmetry could be detected by the 1D distribution of disparity sensors shown in [Figure 1B](#), even though this population is blind to the sign of vertical disparity (6DH). In the “infinite-distance” condition (right two columns), a more complicated pattern of horizontal and vertical on-screen disparity is applied. This produces an asymmetric pattern of on-screen vertical disparity which exactly cancels out the viewing geometry and results in a pattern of vertical disparity on the retina which is symmetric and independent of horizontal position. Applying the magnification to the left eye instead of the right would simply invert the sign of this retinal vertical disparity field, leaving the magnitude unchanged ([Figure 6K](#) vs. [6O](#)). This change could not be detected by the 1D population of [Figure 1B](#).

The subjects were the two authors, plus four observers unaware of the experimental hypothesis and new to psychophysical observation (3 male, 1 female, all aged between 16 and 18 years). Subjects were introduced to the experiment by being shown long-duration stimuli with horizontal magnification (the geometric-effect) for both conditions (standard and infinite distance) and asked to report the direction of perceived slant by indicating whether the left or right side of the stimulus appeared closer to them. After they had learnt to report the direction of perceived slant in these stimuli, stimulus duration was reduced to 200 ms. When the subjects had practiced with these stimuli, the two conditions of induced-effect stimuli were then interleaved with the geometric-effect stimuli, without informing the subject of the change and without error feedback. The two conditions, “standard” and “infinite distance” were always randomly interleaved throughout the experiment. In between stimulus presentations, subjects viewed a fixation X flanked by vertical and horizontal Nonius lines to ensure correct vergence.

Control experiment

In a control experiment with 2 observers, we examined a second way of producing the “infinite-distance” condition. Here, the fixation crosses presented to each eye were offset horizontally by the observer's interocular distance (measured individually for each observer), so that the optic axes were parallel. To aid fusion, a random-dot field was presented in between trials, with the same horizontal displacement, using yellow dots to indicate that no judgment of slant was required. With this full-screen

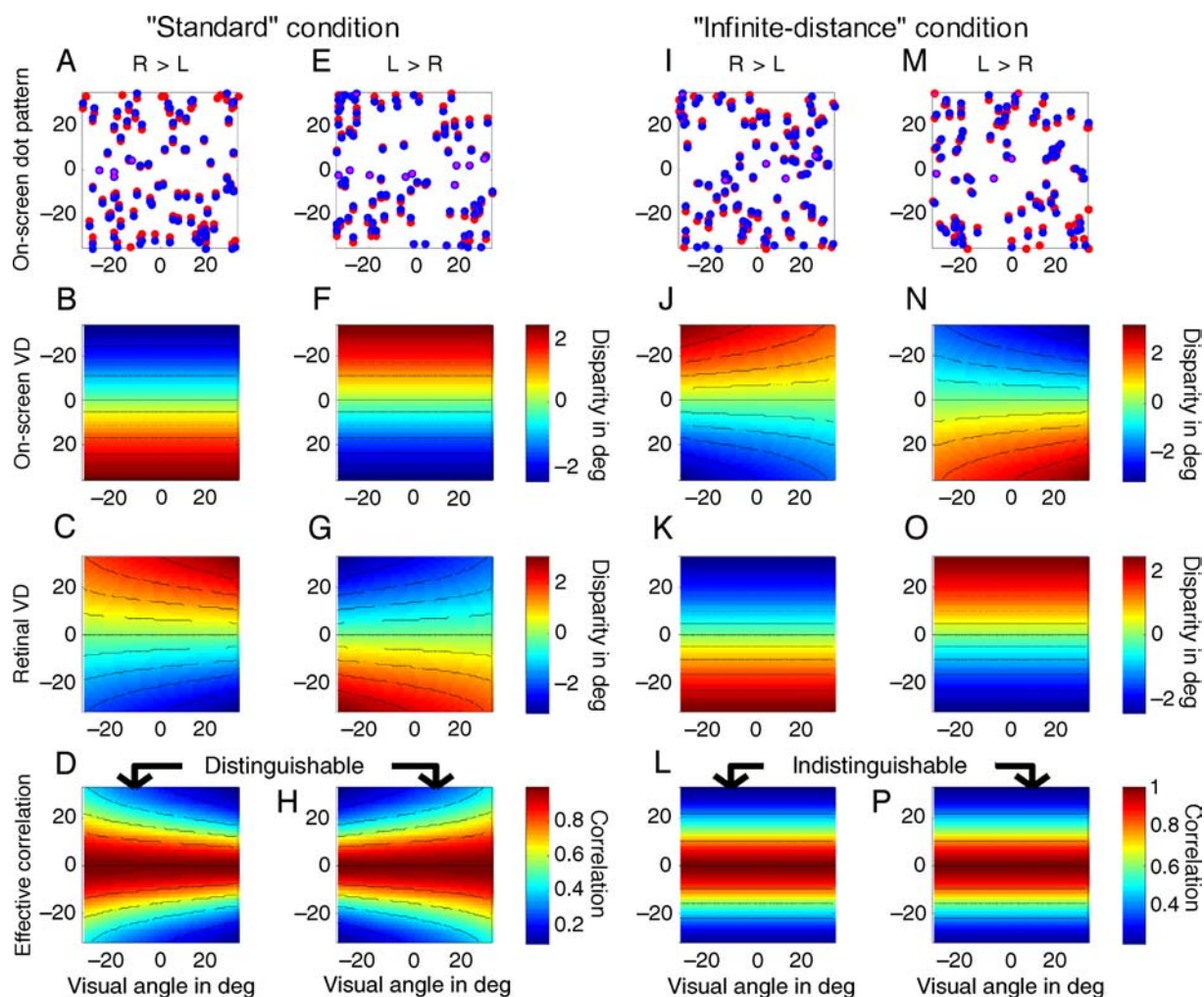


Figure 6. Induced effect stimulus for the two conditions, the “standard” induced effect (A–H) and the “infinite-distance” condition (I–P). Each condition is shown for the two signs of magnification: right eye stretched vertically by 5%, left eye compressed vertically by 5% (A–D, I–L) and left eye stretched 5%, right eye compressed vertically 5% (E–H, M–P), for an overall magnification of 10%. Horizontal and vertical axes of each plot show position in the visual field in degrees visual angle. Top row AEIM: random-dot pattern on the physical projection screen (red dots = left eye, blue dots = right eye; corresponding dots are linked by a purple line). The apparatus is adjusted so that observer’s eyes are level with 0 on the screen. Second row BFJN: Vertical disparity V on the screen. Third row CGKO: Resulting pattern of vertical disparity on the retina, given the viewing geometry. This is calculated for a viewing distance of 165 cm and an interocular distance of 6.3 cm. Fourth row DHLP: Effective interocular correlation sensed by correlation detectors with zero vertical disparity on the retina, calculated from $C_{\text{eff}} = \exp(-V^2 / 2\sigma^2)$, for a receptive field size of $\sigma = 1^\circ$. The effective correlation depends only on the magnitude, not the sign, of retinal vertical disparity, but in the standard induced effect the two signs of magnification can nevertheless be distinguished by the pattern of correlation across the retina. In the infinite-distance condition, the two signs of magnification produce identical patterns of correlation, and so cannot even in principle be distinguished by this population of correlation detectors.

stimulus, observers had no problem maintaining a single image of the fixation cross. Trial images were then the classic induced effect stimulus, i.e., with one eye’s image simply vertically magnified, again presented with the images offset horizontally. Once again, trial images were presented for 200 ms. If the observer fixates this stimulus correctly, it will produce the same retinal disparity field as the previous “infinite-distance” condition. In this experiment, the “standard-distance” condition could not be interleaved, as this would have required changes in vergence from trial to trial.

Data analysis

Psychometric functions were fitted as a cumulative Gaussian function of magnification, using a maximum likelihood fit assuming simple binomial statistics. The standard deviation, σ , of the cumulative Gaussian was taken as the threshold. Confidence intervals on the fitted threshold were obtained by boot-strap resampling (Wichmann & Hill, 2001). Briefly, we simulated each experiment by using the fitted psychometric function to the original data (see Figure 8) as the model for the

observer and using the same number of samples per point as in the original data. A new psychometric function was fitted to each set of simulated data and the value of the threshold (σ') was recorded. Vertical and horizontal black lines crossing the data points in Figure 9 represent the central 95% range of the distribution of 2000 simulated thresholds σ' .

Results

Figure 8 shows psychometric functions for the four conditions, 2 types of magnification (geometric-effect and induced-effect) \times 2 viewing geometries (standard and infinite distance). The data are plotted in separate panels for each of the six subjects. Each panel shows the proportion of right responses as a function of the magnification factor.

For all observers, for the geometric-effect condition (red lines and data points), and for both types of viewing geometry, performance is effectively perfect for the highest and lowest magnification factors. The slopes (σ) estimated from the fitted psychometric functions for both viewing geometries are similar. These results confirm that the geometric effect produces a strong percept of a slanted surface, in both viewing conditions (“standard” vs. “infinite distance”).

For the induced-effect condition (blue lines and data points), the thresholds were usually greater than for the induced effect (especially in subject ISP). However, once again, there was no evidence that thresholds depended on the viewing condition. Subjects were able to report the perceived direction of slant and hence discriminate which eye was magnified, even in the infinite-distance condition.

Figure 9 plots the thresholds in the “infinite-distance” condition as a function of those in the “standard” condition. Squares show results for the induced-effect stimulus (vertical magnification) and circles those for the geometric effect (horizontal magnification). All points lie close to the identity line, and there is no evidence that thresholds are elevated in the “infinite-distance” condition, in either the geometric or induced effect.

Our simulation of “infinite distance” suffers from several potential sources of error. Failure to align the observer’s eyes exactly with fixation, variation in observers’ interocular distances, small misalignments between the two images, and geometrical nonlinearities all mean that the correction will not be perfect. However, under the model proposed by Read and Cumming (2006), one would expect the slant illusion to be substantially weakened by the applied correction, even if inaccuracies in the correction meant that the illusion was not abolished completely. Thus, the fact that thresholds are completely unchanged is strong evidence that the model of Read and Cumming is not correct. As a control, observers JCC and

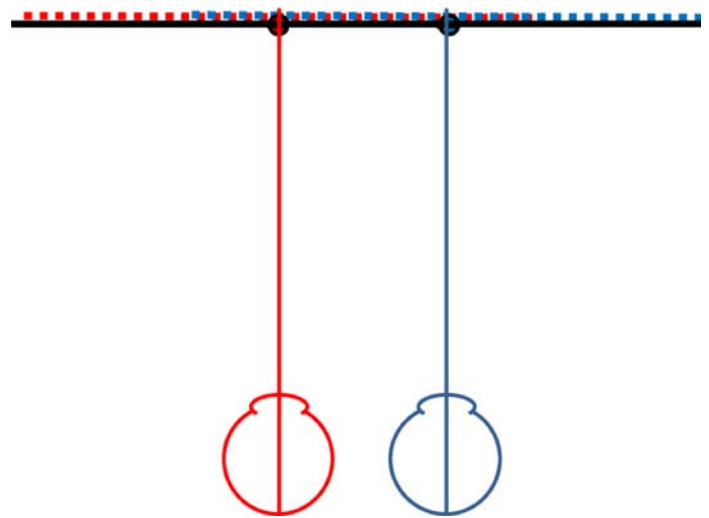


Figure 7. Alternative means of generating the “infinite-distance” induced effect, used in control experiments. The red and blue dots indicate the position of the fixation cross presented to the left and right eye respectively. The red and blue dotted lines indicate the horizontal extent of the random dot fields presented to left and right eyes. If the observer fixates the crosses correctly, they will adopt primary position, suitable for viewing an infinite-distance stimulus.

JLH also carried out the task using a different way of simulating infinite distance (Methods, Figure 7). In this experiment, the fixation crosses presented to the two eyes were offset horizontally by the observer’s individual interocular distance, so that the optic axes were parallel. Once again, both subjects perceived the induced effect, with no significant change in threshold.

Discussion

In the “infinite-distance” condition, the pattern of vertical disparity magnitude is the same on the retina irrespective of which eye’s image is magnified and which compressed; only the sign inverts. Yet, subjects still clearly perceive the induced effect under these circumstances (Figures 8 and 9). They are able to discriminate which eye is magnified, via the effect on the sign on perceived slant. There was considerable inter-subject variation in the strength of the percept produced by the induced effect and hence in the reliability of the discrimination. Observer ISP, for example, experienced the induced effect only very weakly, and was never able to rise above 70% correct for any magnification factor (Figures 8 and 9). Critically, however, whatever the strength of the perception, there was no evidence that it was weaker in the “infinite-distance” condition (Figure 9). It is hard to envisage any read-out of a 1D population which would not produce a weaker illusion in this

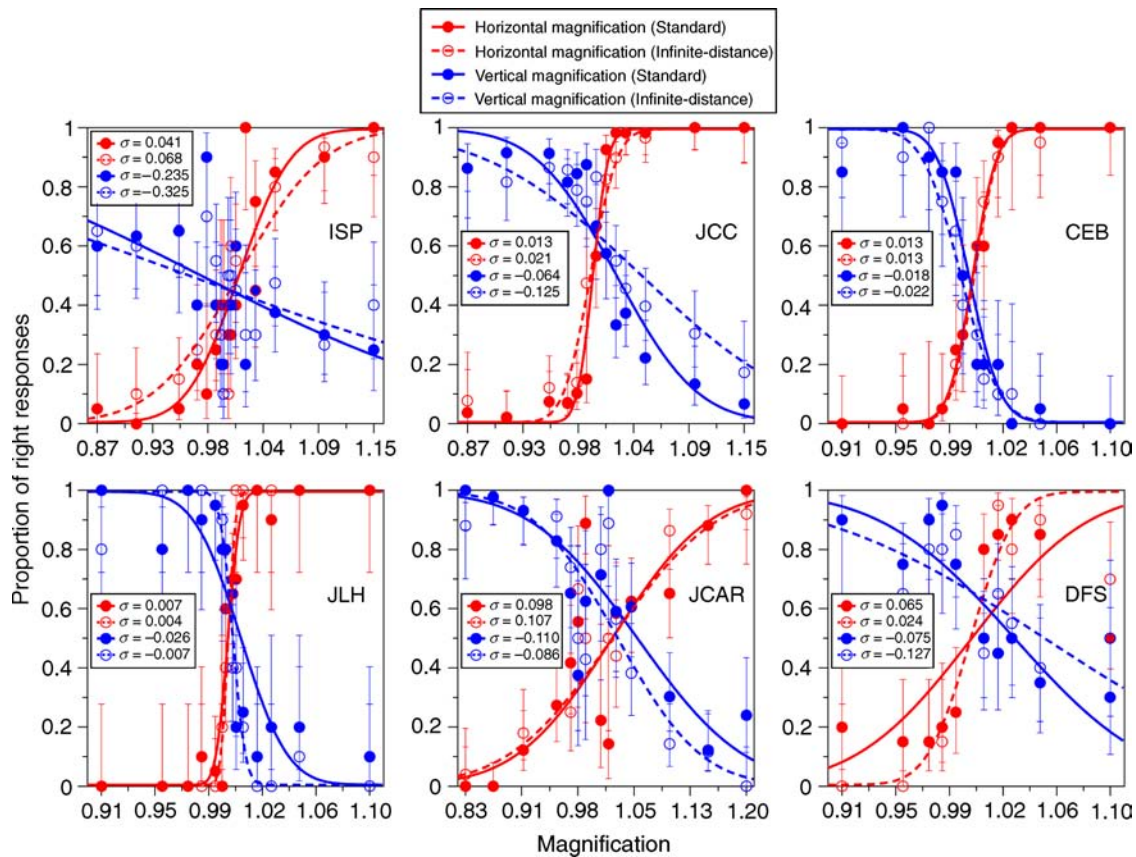


Figure 8. Results of the experiment for the six subjects. The subjects were asked to discriminate which side of the screen (left or right) was closer to them in a one-interval forced-choice task. For magnification factors (vertical or horizontal) higher than 1 the image was bigger for the left eye and vice versa. Each panel shows the proportion of right responses as a function of the magnification factor for each subject. Red lines and data points show the results for the geometric-effect condition (horizontal magnifications). Blue lines and data points show the results for the induced-effect condition (vertical magnifications). Filled dots and solid lines show the results for the “standard” viewing geometry and unfilled dots and dashed lines show the results for the “infinite-distance” viewing geometry. Solid and dashed lines are Gaussian psychometric functions fitted to the experimental data by maximum likelihood. On the left part of each panel the values of the 84% thresholds (σ) obtained from the fitted functions are showed. Error bars show the 95% confidence limits assuming binomial variability, the limits were obtained using the score confidence interval (Agresti & Coull, 1998).

condition. Even if our simulation of infinite-distance viewing geometry was not entirely successful, so that some cues remained, it would surely have had to lower the reliability with which a 1D population could support discrimination. We conclude that visual perception measures the sign as well as to the magnitude of vertical disparity, and thus that it depends on a 2D encoding. The model proposed by Read and Cumming (2006) is not in fact used by the visual system.

Why does the visual system not take advantage of what we have argued would be a more efficient encoding? One possibility is the difficulty of stereo correspondence. Once the correct horizontal disparity is known, then vertical disparity can be deduced even from a 1D population, by the reduction in experienced correlation, but of course this depends critically on knowing what the correct horizontal disparity is. For a uniform-disparity stimulus, this is trivial, even if the stimulus contains both vertical and horizontal disparity: as described above, it is simply the

preferred disparity of the sensor reporting maximal correlation. However, in a realistic visual scene, containing a multitude of different disparities, stereo correspondence is a very serious challenge. The sensor reporting the highest correlation at a particular scale is not necessarily that tuned to the correct stimulus disparity. One can imagine that this problem is still worse in a 1D population, where the sensor tuned to the correct horizontal disparity is handicapped further by not being tuned to the correct vertical disparity (if this is non-zero). This reduces the effective binocular correlation it reports and thus might make it less likely to win out over sensors tuned to false matches which randomly happen to give high correlation in this particular image.

More speculatively, a 2D population would in theory enable the brain to simplify stereo correspondence by taking account of eye posture. Imagine a population of neurons all tuned to the same position in the visual field, the same horizontal disparity but a range of vertical

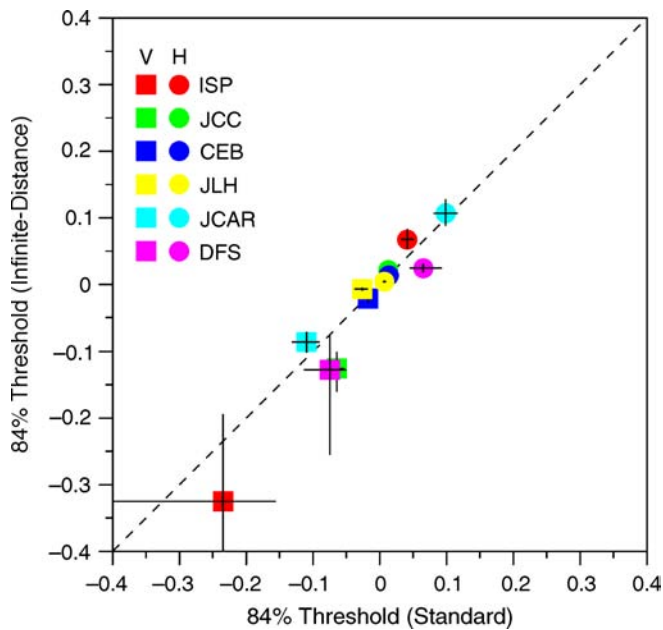


Figure 9. Thresholds (σ) for the “infinite-distance” condition as a function of those for the “standard” condition. The diagonal dashed line shows the identity line. Squares represent thresholds for the induced-effect (vertical magnification) stimulus and circles those for the geometric-effect (horizontal magnification) stimulus. Error bars were obtained by bootstrap resampling as described in the [Methods](#).

disparities. For a given binocular eye posture (convergence, gaze angle, and elevation), only one vertical disparity is *epipolar*, i.e., consistent with the binocular geometry. Thus, there is only one cell in this population which can be reporting the correct binocular correspondence at this point in the visual field; activity in the others must reflect false matches. The correspondence problem would be made easier to solve if neurons tuned to epipolar disparities were somehow boosted. This would only be possible if neurons were available tuned to a range of vertical disparities (or if neurons dynamically adjusted their disparity tuning so as to ensure they were tuned to the current epipolar geometry). Using epipolar geometry to aid stereo correspondence is routine in multiple-camera machine vision (Hartley & Zisserman, 2000), but it must be said there is as yet no evidence for it in human stereopsis.

Although our results show that visual perception measures the sign of vertical disparity, they say little about the scale on which this map is constructed. Strictly, all this study really shows is that vertical disparity sign is available at at least two positions in the visual field, say above and below fixation. Further psychophysical work, for example extending the work of Kaneko and Howard (1997) to short durations, will be needed to ascertain the degree of detail with which vertical disparity is encoded. Since ecological vertical disparity varies slowly and predictably across the visual field (see e.g., Figure 9 of Read & Cumming, 2006), it would be economical for the brain to concentrate

computational resources on encoding horizontal disparity. Vertical-disparity detectors may be few and far between in the visual field, even if the total output of this sparse population is important in interpreting activity in the much more numerous horizontal-disparity detectors. This means that, despite the results of this paper, physiologists may struggle to find vertical disparity detectors, simply because they are not very numerous. Whatever the outcome, it will be important for future physiology studies to report vertical disparity not only in screen coordinates, but also in one of the retinal coordinate systems which are in common use (of which the Longuet-Higgins Cartesian system adopted here is just one).

To our knowledge, this is the first published paper demonstrating that the induced effect can be reliably perceived at short durations. Thus, while the strength of the percept may indeed build up over time (Kaneko & Howard, 1997; Ogle, 1938; Westheimer, 1984), long presentations are not necessary to produce the effect.

Conclusions

Visual perception uses an explicit, signed encoding of two-dimensional disparity.

Appendix A

Derivation of Equations 1 and 2

In [Figure A1](#), the white dot shows the position of an image on the virtual screen (dotted red line), with horizontal and vertical coordinates X' and Y' , respectively, on the virtual screen. The red dot shows where the corresponding image has to be drawn on the physical screen (black line). It has coordinates (X, Y) on the physical screen. We now derive [Equation 1](#) relating (X_L, Y_L) to (X'_L, Y'_L) for the left eye. To do this, it will be useful to introduce a head-centered coordinates system (X_H, Y_H, Z_H) as indicated in [Figure A1](#). X_H and Y_H are parallel to the coordinate axes on the physical screen; the Y_H therefore comes “out of the paper” towards the reader. The virtual white dot therefore has head-centered coordinates $X_H = X'_L \cos \theta$, $Y_H = Y'_L$, and $Z_H = Z + X' \sin \theta$, where Z is the distance to the screen and θ is half the vergence angle. The nodal point of the left eye has coordinates $X_H = Z \tan \theta$, $Y_H = Z_H = 0$. A line passing through the nodal point and the virtual image has vector equation

$$\begin{pmatrix} X_H \\ Y_H \\ Z_H \end{pmatrix} = \begin{pmatrix} Z \tan \theta \\ 0 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} X'_L \cos \theta - Z \tan \theta \\ Y'_L \\ Z + X'_L \sin \theta \end{pmatrix}. \quad (\text{A1})$$

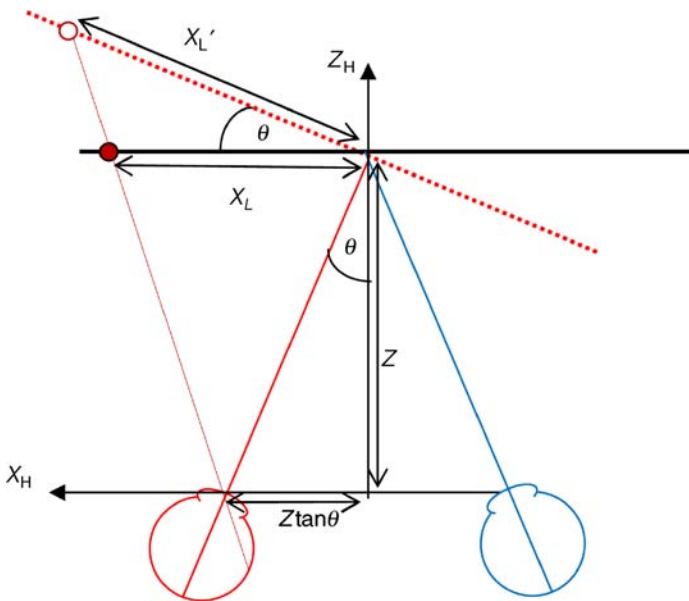


Figure A1. As Figure 4, showing distances and angles used in the derivation of Equation 1.

This line intersects the screen, which is the plane $Z_H = Z$, at $\lambda = Z / (Z + X'_L \sin \theta)$. Substituting this value of λ into Equation A1 gives us the coordinates of the red dot on the physical screen:

$$X_L = Z \tan \theta + \frac{Z(X'_L \cos \theta - Z \tan \theta)}{Z + X'_L \sin \theta} = \frac{X'_L Z \sec \theta}{Z + X'_L \sin \theta}$$

$$Y_L = \frac{Y'_L Z}{Z + X'_L \sin \theta}, \quad (\text{A2})$$

which is Equation 1. For the right eye, the nodal point is at $X_H = -Z \tan \theta$. A similar derivation then yields Equation 2.

Acknowledgments

This research was supported by Royal Society University Research Fellowship UF041260 and MRC New Investigator Award 80154 to JCAR.

Commercial relationships: none.

Corresponding author: Jenny Read.

Email: j.c.a.read@ncl.ac.uk.

Address: Henry Wellcome Building, Framlington Place, Newcastle upon Tyne, NE2 4HH, UK.

References

- Agresti, A., & Coull, B. (1998). Approximate is better than “exact” for interval estimation of binomial proportions. *American Statistician*, 52, 119–126.
- Allison, R. S., Howard, I. P., Rogers, B. J., & Bridge, H. (1998). Temporal aspects of slant and inclination perception. *Perception*, 27, 1287–1304. [PubMed]
- Backus, B. T., Banks, M. S., van Ee, R., & Crowell, J. A. (1999). Horizontal and vertical disparity, eye position, and stereoscopic slant perception. *Vision Research*, 39, 1143–1170. [PubMed]
- Banks, M. S., Hooge, I. T., & Backus, B. T. (2001). Perceiving slant about a horizontal axis from stereopsis. *Journal of Vision*, 1(2):1, 55–79, <http://journalofvision.org/1/2/1/>, doi:10.1167/1.2.1. [PubMed] [Article]
- Busettini, C., Fitzgibbon, E. J., & Miles, F. A. (2001). Short-latency disparity vergence in humans. *Journal of Neurophysiology*, 85, 1129–1152. [PubMed] [Article]
- Cumming, B. G. (2002). An unexpected specialization for horizontal disparity in primate primary visual cortex. *Nature*, 418, 633–636. [PubMed]
- Cumming, B. G., & DeAngelis, G. C. (2001). The physiology of stereopsis. *Annual Review of Neuroscience*, 24, 203–238. [PubMed]
- Dodgson, N. (2004). *Variation and extrema of human interpupillary distance*. Paper presented at the Proceedings of SPIE, San Jose, California.
- Durand, J. B., Celebrini, S., & Trotter, Y. (2007). Neural bases of stereopsis across visual field of the alert macaque monkey. *Cerebral Cortex*, 17, 1260–1273. [PubMed] [Article]
- Durand, J. B., Zhu, S., Celebrini, S., & Trotter, Y. (2002). Neurons in parafoveal areas V1 and V2 encode vertical and horizontal disparities. *Journal of Neurophysiology*, 88, 2874–2879. [PubMed] [Article]
- Gonzalez, F., Justo, M. S., Bermudez, M. A., & Perez, R. (2003). Sensitivity to horizontal and vertical disparity and orientation preference in areas V1 and V2 of the monkey. *Neuroreport*, 14, 829–832. [PubMed]
- Gonzalez, F., Relova, J. L., Perez, R., Acuña, C., & Alonso, J. M. (1993). Cell responses to vertical and horizontal retinal disparities in the monkey visual cortex. *Neuroscience Letters*, 160, 167–170. [PubMed]
- Hartley, R., & Zisserman, A. (2000). *Multiple view geometry in computer vision*. Cambridge, UK: Cambridge University Press.
- Helmholtz, H. v. (1925). *Treatise on physiological optics*. Rochester, NY: Optical Society of America.

- Hibbard, P. B. (2007). A statistical model of binocular disparity. *Visual Cognition*, *15*, 149–165.
- Kaneko, H., & Howard, I. P. (1997). Spatial limitation of vertical-size disparity processing. *Vision Research*, *37*, 2871–2878. [PubMed]
- Longuet-Higgins, H. C. (1982). The role of the vertical dimension in stereoscopic vision. *Perception*, *11*, 377–386. [PubMed]
- Maunsell, J. H., & Van Essen, D. C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey: II. Binocular interactions and sensitivity to binocular disparity. *Journal of Neurophysiology*, *49*, 1148–1167. [PubMed]
- Ogle, K. N. (1938). Induced size effect I: A new phenomenon in binocular vision associated with the relative size of the images in the two eyes. *Archives of Ophthalmology*, *20*, 604.
- Ogle, K. N. (1964). *Researches in binocular vision*. New York: Hafner.
- Parker, A., Cumming, B., & Dodd, J. (2000). Binocular neurons and the perception of depth. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences* (pp. 263–277). Cambridge, MA: The MIT Press.
- Poggio, G. E. (1995). Mechanisms of stereopsis in monkey visual cortex. *Cerebral Cortex*, *5*, 193–204. [PubMed]
- Qian, N., & Zhu, Y. (1997). Physiological computation of binocular disparity. *Vision Research*, *37*, 1811–1827. [PubMed]
- Read, J. C. A., & Cumming, B. G. (2004). Understanding the cortical specialization for horizontal disparity. *Neural Computation*, *16*, 1983–2020. [PubMed] [Article]
- Read, J. C. A., & Cumming, B. G. (2006). Does depth perception require vertical-disparity detectors? *Journal of Vision*, *6*(12):1, 1323–1355, <http://journalofvision.org/6/12/1/>, doi:10.1167/6.12.1. [PubMed] [Article]
- Rogers, B. J., & Bradshaw, M. F. (1993). Vertical disparities, differential perspective and binocular stereopsis. *Nature*, *361*, 253–255. [PubMed]
- Stevenson, S. B., & Schor, C. M. (1997). Human stereo matching is not restricted to epipolar lines. *Vision Research*, *37*, 2717–2723. [PubMed]
- Trotter, Y., Celebrini, S., & Durand, J. B. (2004). Evidence for implication of primate area V1 in neural 3-D spatial localization processing. *The Journal of Physiology*, *98*, 125–134. [PubMed]
- Westheimer, G. (1984). Sensitivity for vertical retinal image differences. *Nature*, *307*, 632–634. [PubMed]
- Wheatstone, C. (1838). On some remarkable, and hitherto unobserved, Phenomena of Binocular Vision. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences*, *128*, 371–394.
- Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Perception & Psychophysics*, *63*, 1314–1329. [PubMed]