The contribution of vergence change to the measurement of relative disparity

Benjamin T. Backus

Daniel Matza-Brown

The relative disparity between two objects in a scene can in principle be measured directly from the retinal images, without knowledge of eye position. But relative disparity increment thresholds are lowest when the relative disparity is small and the objects are not widely separated in the visual field: thus, some relative disparities are easier for the visual system to measure than others. We consider, after others, a second method by which the visual system could measure relative disparity, based on change in vergence (“delta vergence” or DV). The DV mechanism could be more reliable than the retinal mechanism when visual targets are widely separated in visual direction or depth. We used a cue-conflict paradigm to measure the extent to which perceived depth depends on DV. As target separation increased, so did reliance on DV. As intertarget disparity increased, reliance on DV increased for one observer but not for two others.

Keywords: binocular vision, sequential stereopsis, relative disparity, vergence, depth

Introduction

The usefulness of binocular vision for depth perception derives from the fact that visual targets at finite distance in front of the head have different head-centric visual directions at each eye. This difference in visual direction is called the target’s “binocular parallax” (Foley, 1978) or “headcentric disparity” (Erkelens & van Ee, 1998) in the vision literature, or its “vergence” in the photogrammetry and computer vision literatures. Along any line of sight, targets with large headcentric disparity are closer to the head than targets with small headcentric disparity.

If one wants to know which of two targets is closer, and perhaps by how much, the relative disparity—or difference in the objects’ headcentric disparities—is a useful signal. In principle, relative disparity can be measured directly from the retinal images, without knowledge of eye position. In practice, however, sensitivity to change in relative disparity declines as the targets are separated from one another within the visual field (Andersen & Weymouth, 1923; Rady, 1955; Wright, 1951), as the magnitude of the relative disparity increases (Lasley, Kivlin, Rich, & Flynn, 1984), or as the targets’ common absolute retinal disparity (the “disparity pedestal”) increases (Blakemore, 1970; Ogle, 1953). These falloffs in sensitivity can be attributed to inefficiencies in the visual system. Although the system could have been built to make precise comparisons of absolute retinal disparity (across large expanses of visual field, between very different absolute retinal disparities, or between similar absolute retinal disparities evoked by targets that are both far from the horopter), it appears not to have been so built.

A second, less-studied method exists by which the visual system could measure relative disparity, that we will call the “delta vergence” (DV) method. This method is based not on simultaneously available retinal disparities but rather on the change in vergence eye posture across sequential fixations (Brenner & van Damme, 1998; Enright, 1991a, 1991b, 1996; Frisby, Catherall, Porrill, & Buckley, 1997; Rady, 1955; Taroyan, Buckley, Porrill, & Frisby, 2000; Wright, 1951). Observers prefer to look back and forth between targets when asked to estimate depth intervals between them (Enright, 1991a). One reason for this could be to make several retinal disparity measurements, but it could also be that doing so allows observers to overcome limitations in their ability to use retinally measured disparity.

Wright (1951) and Rady and Ishak (1955) estimated the relative contributions of the vergence-based and retinal-disparity based mechanisms, across a wide range of target separations, by asking observers to make relative depth judgments when eye movements between the targets were either allowed or not allowed. Both studies found that as target separation increased, so did reliance on vergence eye posture. However, they came to different conclusions about the importance of vergence: only Wright concluded that vergence normally plays a major role in depth perception. Neither study considered in detail how vergence might actually be used by the visual system. Recent discussions of possible mechanisms can be found in Enright (1991a), Brenner and van Damme (1998), and Taroyan et al. (2000).
Does the visual system regularly make use of the DV method for measuring relative disparity? We asked observers to make perceptual depth matches, using a simple cue-conflict stimulus in which the retinal relative disparity (RRD) between two objects differed from the change in vergence eye posture (DV) required to fixate the objects successively. We therefore made a direct measurement of the relative contributions of RRD and DV to perceived depth, in a situation where observers were free to choose when to make eye movements. As objects were more widely separated within the visual field, observers’ depth settings did indeed show a transition from reliance on RRD to reliance on DV, in agreement with Wright (1951) and Rady and Ishak (1955). However, there were large individual differences, and this might account for the differences between the conclusions reached in those studies.

To summarize, RRD and DV are two signals that the visual system could use to estimate the disparity between a pair of targets in the world. Under normal conditions, RRD and DV are both equal to the actual disparity between the targets. In a laboratory setting, we can create artificial stimuli in which these signals differ, in order to determine which signal the visual system relies on.

**Sequential Stereopsis: Previous Work**

Enright (1989; 1991a) used the term “sequential stereopsis” to describe a mechanism by which the visual system could compare the distances of targets in space, separated by a visual angle. He supposed that the visual system could execute precise “isovergence saccades” between the targets’ locations. After such a saccade, the absolute retinal disparity in the second target’s image signals its headcentric disparity relative to initial fixation. Enright (1991a) noted that accurate fixation of the second target was not required to set its disparity equal to the first: targets that were flashed for as little as 50 ms sufficed for trained observers to null depth differences to within about 3 arc min across 14 deg of visual space. Enright therefore reasoned that change in vergence, per se, is not necessary for the comparison of depth across space, and that the absolute disparities of targets can be compared not only over time (Foley, 1976) but also over time with a saccadic eye movement between stimuli.

Enright appears to have reasoned incorrectly when he described the putative mechanism to be dependent on isovergence saccades. Brenner and van Damme (1998) found that observers were able to use vergence to adjust the distance of a target to one half or twice the distance of a standard, with a precision (trial-by-trial SD) that was not much worse than setting the target to the same distance as the standard (not surprisingly, mean settings were not as accurate in the half-distance condition). They concluded that vergence change is likely measured with precision by the visual system; in their task, measuring the change, per se, was not sufficient for accurate performance without accurate knowledge of the eye’s absolute vergence, and they suggested that errors in absolute calibration could account for observers’ biases. Taroyan et al. (2000) also found evidence that the visual system is able to precisely control vergence change, not only for isovergent saccades, but also for arbitrary changes.

Brenner and van Damme (1998) and Taroyan et al. (2000) chose not to give this generalized version of “sequential stereopsis” a new name. The term now has two meanings in the literature, and we will continue to use it for both of these meanings: sequential stereopsis refers to the specific mechanism proposed by Enright (1989, 1991a), in which the visual system determines which of two targets is closer by comparing their absolute retinal disparities across an isovergence; and it also refers more generally to phenomena wherein the visual system is able to estimate depth intervals by measuring the disparities of two targets, collected one after the other, and taking vergence into account.

We use the term “delta vergence mechanism” to describe a particular hypothetical mechanism that measures relative disparity by measuring changes in vergence. We suggest that its output is a measurement of relative disparity per se, rather than a distance interval estimate, precisely because the latter would require knowledge of the absolute vergence posture of the eyes (Brenner & van Damme, 1998), and because relative disparity is a useful signal in its own right, that might best be calibrated independently of perceived distance or absolute vergence eye posture.

Like Brenner and van Damme (1998), we are unable to distinguish two possible variants of the DV mechanism: one in which relative disparity is measured by integrating all changes in vergence needed to achieve fixation of the second target, versus one in which the absolute retinal disparities of the targets are measured before and after the eye movement, and added to the vergence change during the eye movement. However, Enright’s observations with flashed stimuli suggest that the latter is possible. For simplicity, we assume accurate initial fixation of the first target (before the eye movement), so its absolute retinal disparity is zero.

**Vergence During Saccades**

Version and vergence during a non-isovergent saccade are to a large extent independently controlled (Collewijn, Erkelens, & Steinman, 1997). The saccade to a visual target includes not only a preprogrammed change in version (Robinson, 1964), but also a preprogrammed change in vergence (Semmlow, Hung, Horng, & Ciuффреда, 1993). Suppose that the visual system knows, with good precision, the magnitude of the vergence component in a saccade. How might it use this information to measure the relative disparity between two targets?
A paradox is that in order to plan the vergence component of a saccade in the first place, the absolute retinal disparity of the new target must be estimated. Sheliga and Miles (2003) report in this issue that perceived depth contributes to such an estimate, and of course the retinal disparity of the new target can also be measured from the retinal images. If the estimate is sufficiently accurate and precise, however, then the visual system already knows the relative disparity between the two targets; in this case, there would be no need to monitor vergence change during the change of fixation from one target to the other. However, the system may know actual changes in vergence eye posture better than it knows the change required to fixate a new target. First, there is some evidence that vergence eye movements are characterized by an initial, preprogrammed component that achieves much but not all of the required vergence change, followed by a second, slower, feedback-controlled component that closes the remaining fixation disparity (Semmlow, Hung, Horng, & Ciuffreda, 1994). In essence, the system can be described as habitually relying on the feedback-controlled phase to improve on its inaccurate initial guess. Second, the fact that perceived depth is more accurate when eye movements are allowed is itself evidence that measurements of actual vergence change are of higher quality than the retinal measurements of disparity that would be needed to plan accurate saccades: observers are better able to null the depth interval between widely separated targets when they are allowed to make eye movements (Enright, 1991a, 1991b), and perceived depth for large disparity intervals is poor when eye movements between near and far targets are disallowed (Foley & Richards, 1972). For large intertarget separations and intertarget disparities, the visual system appears not to rely on retinal disparity alone to generate perceived depth, so it is likely that retinally measured disparity cannot support the planning of accurate vergence eye movements either.

The Delta Vergence Mechanism

We suppose that in the case of large intertarget separation or disparity, the DV mechanism can measure the magnitude of actual vergence change better than it can predict the magnitude of change needed to fixate the new target. It would operate as follows (see Figure 1):

1. A saccade is planned to fixate the new target. The plan specifies changes in both version and vergence eye posture, but the specified change of vergence does not correspond very well to the change needed to fixate the target.

2. The saccade is executed, and the specified change of vergence is stored in memory.

3. The residual disparity of the target is estimated, and added to the specified change. The residual disparity might be measured from absolute retinal disparity (because separation and disparity are now small), or from the vergence eye movements used to ultimately fixate the target.

The DV mechanism for estimating disparity is presumably limited by the visual system’s ability to measure changes in vergence. If both the separation and the disparity between two targets are small, then the retinal method of measuring relative disparity will be more reliable than the DV method. It is possible that both the retinal and DV methods become less reliable as intertarget separation or disparity is increased. However, when observers are asked to adjust the distance to a target to half or double the distance of a standard, the variance in their settings is not much more than for a simple distance match (Brenner & van Damme, 1998). This result suggests that the DV mechanism’s reliability does not fall off with intertarget disparity as quickly as the RRD mechanism’s reliability.

In the experiments described below, we measured the extent to which the visual system relies on DV as opposed to RRD, as intertarget separation and disparity were varied.

Figure 1. The delta vergence (DV) mechanism. The relative disparity between points A and B, or \( \Delta \mu \), is estimated from a corollary discharge signal that accurately estimates the change in vergence eye posture during the eye movement, added to a measurement of the residual disparity. Quantities estimated by the visual system are indicated with hat symbols (\(^\hat{\cdot}\)).
General Methods

Observers

All observers had normal vision or vision corrected to normal with contact lenses. They included the authors, an experienced adult psychophysical observer who was naïve to the details of the experiment, and 13 undergraduates from the University of Pennsylvania. Only some observers passed the initial screening tests for Experiments 1 and 2; no further data were collected from observers who did not pass. Observers participated in accordance with a University of Pennsylvania Institutional Review Board’s approved protocol.

Apparatus & Stimulus Construction

Stimuli were displayed in a dark room using a haploscope, with each eye viewing a separate image in a mirror. Four-bit monochrome images were presented using a pair of 19-in. Clinton Medical monochrome CRT displays, each with a screen resolution of 1280 x 1024 pixels. Visual targets were dots of radius 2.7 mm with raised cosine luminance profiles. The dots were anti-aliased to allow interpixel positioning, and shown against a black background. The displays were refreshed at 75 Hz. Displays were spatially calibrated by aligning a lattice of dots on the screen to a calibration grid located 40 cm directly in front of the viewer (seen though half-silvered mirrors).

The observer’s gaze was tracked using an SR Research Eyelink I at a sample rate of 250 Hz. The eye tracker’s cameras were mounted under the haploscope mirrors, and were pointed at the observer’s eyes through the mirrors. The observer’s head was kept still using a bite bar. A real-time recalibration procedure was used to correct for slight shifts in head position during the experiment; it assumed that observers fixated the targets, as per the instructions.

The software to run the experiments was written using Matlab 6.5, the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) and the Eyelink Toolbox (Cornelissen, Peters, & Palmer, 2002).

Experiment 1: Depth Nulling

Does the visual system rely increasingly on DV, as target separation increases? To answer this question, we used a cue conflict stimulus, in which the intertarget disparity measured by the RRD mechanism differed by a constant amount from the disparity measured by the DV mechanism. Observers nulled the apparent depth between two vertically separated targets by adjusting the horizontal disparity of the bottom target. When looking at the top target, its distance from the observer was 40 cm. But when a gaze-shift to the bottom target was detected, both the top and bottom targets were displaced in depth by a constant disparity. Thus, the observer might do the task by nulling the instantaneous intertarget disparity, showing reliance on RRD, or by nulling the disparity between the top and bottom targets across changes of fixation, showing reliance on DV, or some mixture of these two strategies.

Experiment 1 Methods

Stimulus

The stimulus contained two binocularly viewed, vertically separated target dots. At the start of each trial, the top target was projected onto a frontoparallel plane 40 cm in front of the observer, equidistant from the two eyes. The top target served as the standard (reference) throughout the experiment. The bottom target was projected directly below the top target, and the vertical separation between the two targets was either 2.5, 5, 10, or 20 deg, equidistant above and below the horizontal plane. The bottom target acted as the probe, and its disparity was controlled by the observer. Conditions were blocked by vertical separation to facilitate eye-tracking.

Observers were instructed to look only at the targets, and not at the space between them. The stimulus was turned off whenever the observer’s gaze was in the middle third of the space between the two targets (Figure 2, Movie 1). The stimulus changed according to the observer’s fixation. When the observer fixated the top target, one stimulus was shown, and when the observer fixated the bottom target a different stimulus was shown. When the observer fixated the top target, its simulated distance was 40 cm. When the observer fixated the bottom target, however, the top target was sometimes at a different distance. The difference in the top target’s distance during top and bottom fixation represented the conflict size on that trial (Figure 3). The conflict size did not change during the trial.

Task

On each trial, the observer used a numeric keypad to adjust the disparity of the bottom target relative to the top target. Therefore, as the observer’s fixation changed from one target to the other, both the bottom and top targets’ headcentric disparities shifted by the same amount—equal to the trial’s conflict size—and their relative disparity remained unchanged. Observers were instructed to set the bottom target to be the same distance from their eyes as the top target, regardless of the apparent elevation of the target ensemble.

Screening procedures

An initial screening procedure was performed to determine if the observer could see stereoscopic depth, and to estimate the observer’s stereoaucuity at the four separation values. During the screening procedure, the
Figure 2. The stimulus for Experiment 1. The red images represent the stimulus visible to the left eye, and the green images represent the stimulus visible to the right eye. The red horizontal line indicates the height of the observer’s gaze. Panel A shows the stimulus that was visible when the observer fixated the top target, and Panel C shows the stimulus visible when the observer fixated the bottom target. When fixation was between the two targets, no stimulus was visible (Panel B). Panel D summarizes the three zones within the stimulus.

Movie 1. Illustration of the stimulus. The red horizontal line indicates the height of the observer’s gaze. The first 4 s of the movie show the subject looking back and forth between the two targets. The next 6 s show the observer adjusting the disparity of the bottom target.

Figure 3. A side view of the stimulus in Experiment 1 (not to scale). Top panel: the observer fixates the top target. The red arrow shows the depth interval as specified by the targets’ relative retinal disparity. This quantity was adjusted by the observer during the depth-setting procedure. Bottom panel: the observer fixates the bottom target. During the fixation shift, the absolute headcentric disparity of both targets changes by an amount equal to the trial’s conflict size. The retinal disparity between the two targets, however, stays constant during the shift. Crossed disparity is negative, and this figure illustrates a negative conflict (because the targets become more crossed when the observer’s gaze shifts from the top target to the bottom target).
task was identical to that in the other sections of the experiment, except that the screening stimuli were not gaze-contingent and their conflict sizes were 0 (no conflict). Observers made 7 settings at each of the 4 vertical separations, in a single mixed block of 28 trials. The SDs of the settings were examined to determine the observer’s ability to perform the task, with maximum acceptable SDs set to 0.1 deg for the 2.5-deg vertical separation, 0.2 deg for the 5-deg separation, 0.3 deg for the 10-deg separation, and 0.4 deg for the 20-deg separation. Under these criteria, 3 of 14 observers failed to set the targets to be equidistant. Another observer passed the initial screening session, but refused to use the bite bar. These 4 observers were excluded from the rest of the experiment.

Settings in the first screening were used to estimate the four SDs at the four different vertical target separations. The conflict sizes for the rest of the experiment were based on these SDs (Table 1): the five stimuli used (at a given separation) had conflict sizes of -3σ, -1.5σ, 0 (no conflict), +1.5σ, and +3σ. For each vertical separation condition, observers made 7 settings at each of the 5 conflict sizes. On any given trial, the potential retinal disparity setting that the observer could make ranged from -30σ to +30σ, with the trial’s initial retinal disparity value selected randomly from within that range.

Table 1. SDs: Experiment 1 Screening.

<table>
<thead>
<tr>
<th>Observer</th>
<th>Vertical separation (deg)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.5</td>
</tr>
<tr>
<td>BTB</td>
<td>0.021</td>
</tr>
<tr>
<td>DMB</td>
<td>0.020</td>
</tr>
<tr>
<td>N1</td>
<td>0.018</td>
</tr>
<tr>
<td>N2</td>
<td>0.039</td>
</tr>
<tr>
<td>N3</td>
<td>0.031</td>
</tr>
<tr>
<td>N4</td>
<td>0.042</td>
</tr>
</tbody>
</table>

SDs σ in Experiment 1 were estimated from 7 settings at each vertical separation. Conflict sizes used in the rest of Experiment 1 ranged from -3σ to +3σ.

A second screening procedure employed the eye-tracker. The observer’s head was kept in place with a bite bar, because head movement caused the eye-tracker to fall out of calibration. The procedure tested whether the observer’s eyes could be tracked reliably, and also tested the observer’s ability to use DV disparity. The stimuli from the second screening were identical to the stimuli from the experiment’s 10-deg vertical separation condition, except that each target was monocularly presented until it was fixated, at which point it became binocular. When the top target was fixated, the bottom target was visible only in the left eye; when the bottom target was fixated, the top target was visible only in the left eye. In this second screening condition, then, retinal disparity information was not simultaneously available for both targets at once, so the observer was forced to use DV to perform the task.

In the second screening procedure, we failed to obtain reliable eye-tracking for 4 of the 10 remaining naïve observers, because of inadequate pupil recognition or excessive head movement. Of the 6 naïve observers who completed the second screening, 2 were unable to perform the task, and did not continue with the rest of the experiment. Thus, Experiment 1 employed 4 naïve observers and the 2 authors.

Analysis in terms of weights

We model settings as being based on an estimate of the relative disparity between the top (fixed) and bottom (moveable) target, where that estimate is a simple linearly weighted average of the relative disparities measured by the DV and RRD mechanisms, respectively. In this model, observers adjusted the disparity of the bottom target until the weighted average indicated zero relative disparity between the targets. For the cue conflict stimuli, DV and RRD make different predictions about what stimulus will make the targets appear equidistant. However, there could also be a systematic bias in observers’ settings, so rather than interpret settings directly, we used, as our measure of relative weight, the slope of the line that related settings to the stimulus conflict magnitude. This analysis is illustrated in Figure 4,
for two possible outcomes: complete reliance on RRD or DV, respectively. The observer’s actual disparity settings can be described either in terms of RRD or DV; we arbitrarily chose to use RRD.

**Experiment 1 Results**

For each condition of Experiment 1, the observer’s settings (measured in terms of RRD) were plotted as a function of conflict size (Figure 5). A line was fit to these data using simple regression. The weight given to DV was determined by the negative of the slope of this line. The SE of the slope was based on the assumption that slopes were normally distributed. For each observer, we then plotted the weight given to DV as a function of the vertical separation between the targets (Figure 6). The graphs for all six observers are plotted on the same axes in Figure 7.

Each observer gave greatest weight to DV when the target separation was 20 deg (largest separation tested). However, the reliance on DV varied greatly across observers. Observers N1 and N4 showed no use of DV for vertical separations less than 20 deg. Observers’ weights for DV at 20-deg separation ranged from 0.1 (not significant) to 0.8.

--

Figure 5. Data from the 10-deg vertical separation condition of Experiment 1 (observer BTB). Data points are settings (in terms of RRD) of the bottom target’s depth relative to the top target. The curve connects mean values for five different levels of conflict.

Figure 6. DV weight as a function of vertical target separation for one observer (BTB). SEs are roughly equal across conditions because the magnitude of each condition’s cue conflict was based on the observer’s SD for that condition.

Figure 7. Experiment 1 data for all six observers. DV weight is plotted as a function of the vertical separation of the targets. SEs for slope estimates are similar for all observers, because the magnitudes of the cue conflict were based on observers’ SDs.
Experiment 2: Depth Matching

Does the visual system rely increasingly on DV, as intertarget disparity increases? To answer this question, we asked observers to match a depth interval defined by two targets on the left side of the display using two-target cue-conflict stimulus on the right side of the display. When a gaze-shift to the bottom target (in the right target pair) was detected, both the top and bottom targets (on the right) were displaced by a constant disparity offset. The observer might do the task by matching the instantaneous intertarget disparity, showing reliance on RRD; or, he might match the disparity between the top and bottom target across changes of fixation, showing reliance on DV.

Experiment 2 Methods

Stimulus

The stimulus in Experiment 2 consisted of two vertically separated target pairs (each similar to the target pair in Experiment 1), separated horizontally by 15 deg of visual angle. The left target pair served as the standard (or reference) throughout the experiment, while the right target pair acted as the probe. The standard pair contained no conflict, while the probe pair contained conflict sizes that were determined during special no-conflict sessions that preceded sessions in the main experiment.

When the observer was looking at the standard, the probe was visible only in the left eye. Likewise, when the observer was looking at the probe, the standard was visible only in the left eye (occasionally it was also hidden by the blind spot), as shown in Figure 8. This prevented the observer from retinally comparing the disparity of the probe relative to the stimulus, and encouraged the observer to do the task by matching the apparent depth interval of the probe to the remembered depth interval in the standard.

During fixation of the top target of either stimulus (as in Experiment 1), the top target was projected onto a frontoparallel plane 40 cm in front of the observer. The standard target pair had a relative disparity of either –2, –1, 0, 1, or 2 deg. For each trial, the standard and probe pairs had an equal vertical separation. Vertical separation was 2.5, 7, or 20 deg (blocked). Only the probe could have a nonzero conflict size.

Task

The observer was instructed, for each trial, to look first at the depth between the targets in the standard target pair, then to set the depth between the targets in the probe to be the same as in the standard.

To prevent the observer from simply nulling the apparent depth between the two bottom targets, the observer was required to fixate the top target first, whenever he switched gaze from standard to probe or from probe to standard. The stimulus was never presented binocularly until the top target had been fixated.

Screening procedure

Before each block of a given vertical separation, a screening procedure was run (a) to determine if the observer could perform the interval-matching task, and (b) to estimate the observer's SDs in matching depth. The screening procedure was identical to the subsequent experimental condition, except that the probe always had zero conflict. Observers made 7 settings at each of the 5 relative disparity values, in a single mixed block of 35 trials. The SDs of the settings were examined to determine the observer’s ability to perform the task. If the settings were not significantly greater for larger depth intervals ($p < .05$), the observer was not asked to complete the rest of the experiment.

For each vertical separation, the results from the screening procedure were used to estimate the five SDs at the 5 different relative disparity values (Table 2). Observer’s responses in Experiment 1 were linear functions of conflict magnitude, so we collected data at two conflict sizes, namely ±3 SDs of the screening setting distribution, to increase the power of the experiment.

![Figure 8. The stimulus for Experiment 2. Red shows which dots were visible to the left eye, and green shows which dots were visible to the right eye. The short vertical line segment indicates the left-right position of the observer’s gaze. Panel A: the 6 regions of the display (see text). Panel B: observer is looking at the left targets; right targets are monocular. Panel C: observer is looking at the right targets; left targets are monocular.](http://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/933554/ on 06/16/2017)
Observers made 10 settings (5 per conflict value) at each of the 5 relative disparity values, in a single mixed block of 50 trials.

Table 2. SDs: Experiment 2 Screening.  

<table>
<thead>
<tr>
<th>Vertical separation (deg)</th>
<th>Disparity interval (deg)</th>
<th>Observer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>BTB</td>
</tr>
<tr>
<td>-2</td>
<td>2.5</td>
<td>0.20</td>
</tr>
<tr>
<td>-1</td>
<td></td>
<td>0.03</td>
</tr>
<tr>
<td>0</td>
<td></td>
<td>0.02</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>0.08</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>0.18</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.18</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.18</td>
</tr>
</tbody>
</table>

SDs $\sigma$ in Experiment 2 were estimated from 7 settings for each of 5 depth intervals at each of several vertical separations. Conflict sizes used in Experiment 2 were $-3\sigma$ and $+3\sigma$.

Four observers from Experiment 1 agreed to participate in Experiment 2. In the screening for the 7-deg vertical separation condition, observer N2’s settings for the 2-deg depth interval were not significantly different from his settings for the 1-deg depth interval, so N2 did not participate in the rest of the experiment. Observers BTB, DMB, and N1 were able to perform the task, and were observers for Experiment 2.

**Experiment 2 Results**

DV weights and SEs were calculated in the same way as in Experiment 1. All observers showed greater reliance on DV for blocks with larger target separation, as in Experiment 1. The new factor in Experiment 2 was the depth interval, which varied randomly among 5 values within a given block. Figure 9 plots the results. Observer BTB showed a clear effect in the predicted direction, when the target separation was 2.5 deg: at larger depth intervals, there was greater reliance on DV. At 20-deg separation, BTB was saturated and relied mostly on DV for all depth intervals. Observer DMB failed to show a meaningful effect of depth interval at any separation. Observer N1 showed greater reliance for the depth interval of 2 deg than at 0 or 1 deg, but the overall pattern fails to provide strong evidence of an increase in weight for DV as depth interval increases.

**Discussion**

Observers made significant use of DV, and their reliance on DV relative to RRD increased with target separation, as predicted, in both the depth nulling task (Experiment 1) and the depth matching task (Experiment 2). One observer in Experiment 2 also showed clear evidence of increasing reliance on DV as the depth interval increased; evidence from the other two observers in that experiment was not as compelling.

Why might an observer who showed greater use of DV as target separation increased fail to show greater use of DV as intertarget disparity increased? One possibility is that those observers’ ability to use RRD did indeed fall
off as intertarget disparity increased, but that their ability to use DV fell off as well, at a similar rate.

**Phenomenology of the Cue-Conflict Stimulus: A Role for Adaptation?**

An important aspect of the piloting work for these experiments was choosing the magnitude of the cue conflict. We always first measured the observer's ability to null (or match) the depth interval between the top and bottom target, using a stimulus with zero conflict between DV and RRD. In pilot work, we found that so long as the cue conflict was three or fewer SDs in the observer’s depth settings, observers did not notice that the target pair jumped in space during changes of fixation. When observers were debriefed after the experiment, they reported being unaware that the targets jumped in depth during eye movements.

Even the authors (who knew the design of the experiment) were unable to reliably classify trials as having cue conflicts or not. This may seem surprising in the case of a three SD jump. Why might this be the case? We suspect this occurred because the visual system interprets failure to achieve post-eye-movement target fixation as indicating that the eye movement system needs to be recalibrated. The vergence system is subject to continuous recalibration, as shown by the fact that the gain (of the vergence eye movement required to achieve fixation, relative to the retinally measured demand measured just before the eye movement) can be modified through exposure to stimuli that demand gains other than unity to achieve post-saccadic fixation (Erkelens, Collewijn, & Steinman, 1989). If the system falls out of calibration easily, then small discrepancies between DV and RRD should not be interpreted as reliable indicators that objects jumped in depth at the same time as the eye movement, but should instead be interpreted by the system as an error signal for recalibration.

For large cue conflicts (e.g., six or more SDs), the targets *did* appear to jump during saccades between them. In that case, it was possible—indeed it became necessary—to choose whether to do the task by nulling depth while maintaining fixation on a single target, or by nulling the apparent distance between the top target when it was fixated and the bottom target when it was fixated.

**Interobserver Variability**

The second screening procedure for Experiment 1 required the use of DV alone to null an apparent depth interval. SDs for these settings varied by a factor of 4, even for the observers selected to continue in the experiment. Other observers were unable to do the task at all. Why did some observers have much more difficulty than others, when RRD was eliminated from the displays? In short, we don’t know. The DV mechanism exploits, we believe, precise knowledge of planned changes in vergence eye posture; these may be qualitatively different from a continuous, feedback-driven mechanism that is also used to adjust vergence eye posture (Semmlow et al., 1993). If so, and if only the former method gives rise to a useful DV signal, then perhaps observers who show little reliance on DV are those who rely to a greater extent on the second method to achieve binocular fixation.

In Experiment 1 (depth nulling), no observer showed a decline in the use of DV as target separation increased; all were either increasing or not statistically distinguishable from flat (see Figure 7). However, in Experiment 2, only observer BTB showed a robust increase in reliance on DV as intertarget disparity increased (i.e., as the depth interval being matched increased). Two explanations seem possible for the lack of effect in the other observers. First, it could be that for those observers, RRD and DV became harder to measure at similar rates, as relative disparity increased. This would predict the absence of change in their relative weights. Data from Blakemore (1970) show that at an eccentricity of 10 deg, disparity increment thresholds increased fivefold as pedestal disparity (crossed or uncrossed) increased from 0 to 100 arc min; this increase was even larger at eccentricities of 0 or 5 deg. Perhaps DV also becomes harder to use as relative disparity increases. For example, the SD for measuring DV might reflect Weber’s Law, in which case large vergence changes would be measured with less precision than small ones; from this logic, we would expect transition from use of RRD to use of DV to be most dramatic when the vertical separation of the targets is small, not large, but there is no evidence for this in our data. Furthermore, Brenner and van Damme (1998) reported similar variance in half-distance, equal-distance, and double-distance settings in a task that required the use of DV. Thus, more data are needed to resolve this question.

Second, it could be that observers’ fixation strategies changed as a function of intertarget separation. A systematic decrease in the rate of fixation change with increasing disparity would be consistent with greater reliance on RRD. The data do not support this idea, as we discuss in the next section.

It is possible that some of the difference in the results obtained by Wright (1951) and by Rady and Ishak (1955) was a consequence of differences in the actual reliance of their observers on DV.

**Uncontrolled Fixation Shifts and Unlimited Viewing Time**

In order to best estimate the extent to which vergence eye posture is used during natural vision, we allowed observers to choose their own fixation shift timing. We also did not limit the amount of time for each setting. Some observers made many more changes of fixation than others. Indeed, it was possible to do the task without making any changes of fixation at all, in which
case depth settings in Experiments 1 and 2 would of course be based exclusively on RRD. We analyzed the actual fixation strategy employed by each observer, by counting the number of fixation changes per trial and by calculating the rate at which fixation moved between the targets. Figure 10 plots DV weight against fixation frequency for each of the six observers in the 20-deg target separation condition. These data show that differences in fixation strategy cannot account for the different weights accorded to DV by different observers. The same lack of correlation characterizes the other target separations; the r-squared statistic was 0.005, 0.12, and 0.006, for the 5, 10, and 20 deg separations, respectively.

Given that observers must rely on RRD if they are not allowed to make fixation shifts, and that DV requires fixation shifts, one might predict that the depth setting for an individual trial would reflect the number of times the observer looked back and forth between targets on that trial. We would predict that the settings for trials with few fixation shifts would null RRD, and those with frequent fixation shifts would null DV. Yet the data showed that neither fixation shift frequency nor total fixations per trial correlated across trials with reliance on DV, for any observer. Figure 11 shows representative data for one observer. A few observers made more fixation shifts, or more frequent fixation shifts, as a function of the vertical separation of the targets, but these differences were idiosyncratic and there was no pattern across observers.

Why didn’t the number of fixation shifts predict the depth setting on a given trial? Perhaps the reliability of DV measurements varies significantly from one fixation shift to the next. This is plausible for the DV mechanism, because accurate vergence eye movements give rise to more precise DV estimates than inaccurate vergence eye movements, and whether the movement is accurate or not depends quite a bit on chance. In that case, observers may simply have made fixation shifts on each trial until they reached a criterion level in the reliability of their depth setting. It seems likely that fixation shifts impose only a small cost on the measurement of RRD; in any case, trial duration was under the observer’s control. Paradoxically, we would therefore predict that if viewing duration were fixed, DV weight would be highest for trials with the fewest fixation shifts, where the observer garnered reliable DV measurements early in the trial. One might also control the number of fixation shifts, and measure the actual vergence changes. We would predict the greatest reliance on DV on those trials for which the vergence component of the eye movement was most accurate.

![Figure 10](http://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/933554/)  
Figure 10. Each observer’s DV weight for the 20-deg separation condition of Experiment 1, as a function of mean frequency of fixation shift. Vertical error bars are from Figure 7. Horizontal error bars are SEs for the mean across trials.

![Figure 11](http://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/933554/)  
Figure 11. Scatterplot of DV weight versus fixation shift frequency in Experiment 1, observer N1. Each symbol represents one setting. Different symbols show data for the 4 different vertical separations. In this plot, DV weight was computed for each single trial, as the negative of RRD setting, divided by the conflict size (data with a 0-deg conflict size are excluded).
Effects of Stimulus Configuration

Our stimuli consisted of isolated luminous targets seen in the dark. This stimulus has energy at very low spatial frequencies that would be well suited to monitoring by the RRD mechanism. It seems likely we would have observed smaller RRD weights had we used targets defined by high spatial frequency patterns (Enright, 1996; Taroyan et al., 2000). In that case, the absolute retinal disparity of each target would have been visible only when the target was near fixation. Indeed, we would predict that as low spatial frequency content of the targets is progressively eliminated, the weight given to the DV mechanism would increase accordingly.

Our stimuli would also be expected to create afterimages. Could this have played a role in our findings? In particular, did the visual system make use of relative disparity between the currently fixated target, and an afterimage of the previously fixated target? The position of the afterimage in each eye is retinotopic, so the afterimage has fixed absolute retinal disparity. Therefore, to use the afterimage for depth perception, the visual system would have to measure the change in vergence between fixations. Thus, in our experiments, afterimages might in principle have contributed to the use of DV, but not to the use of RRD. It is unlikely that use of DV requires the presence of afterimages, given that reliance on a vergence-based mechanism occurs for high spatial frequency, low-contrast targets (Enright, 1996; Taroyan et al., 2000).

DV, Vertical Disparity, and Slant Perception

Sequential stereopsis has been considered a possible contributor to perceived slant for stereoscopically rendered surfaces. Van Ee and Erkelens (1999) found that large scanning eye movements had no effect on slant estimates for cue-conflict surfaces in which vergence changes specified a slant of 0 deg, and horizontal and vertical disparity together specified a nonzero slant (and evoked the percept of nonzero slant). This was evidence that vergence change did not contribute to slant perception. Berends, Zhang, and Schor (2003) report in this issue that observers better discriminated slant when changes of fixation were allowed than when they were not allowed, but only for stimuli in which disparity noise was added to the positions of dots representing the surface, and only when fixation was shifted in the direction of the disparity gradient. These results suggest that slant is usually estimated using a retinal measurement of the horizontal disparity gradient, but that the DV mechanism limits performance in slant discrimination tasks when the gradient is difficult to measure directly from the retinal images.

In both of those studies, the vertical disparity gradient provided a highly reliable signal that could be used to interpret horizontal disparities as slant (Backus, Banks, van Ee, & Crowell, 1999; Gårding, Porrill, Mayhew, & Frisby, 1995), so perhaps it is not surprising that vergence change across the surface contributed little to perceived slant. In Experiments 3a and 3b of their study, in which eye movements were allowed, Berends et al. (2003) found that when the visual system’s ability to measure the vertical disparity gradient was progressively reduced (either by adding vertical disparity noise or by decreasing the height of the stimulus), slant discrimination thresholds increased, and then leveled off. Thus, performance became limited by a factor other than vertical disparity. In the absence of reliable vertical disparity information, slant can be estimated in several ways: from DV; from the horizontal disparity gradient and static eye position; and from the prior probability distribution and nonstereo cues, both of which indicated zero slant. In our laboratory, we have often observed that allowing eye movements makes it easier to develop a stable percept of stereoscopically rendered nonzero slant; this benefit is clearest for stimuli that do not contain usable vertical disparities. Berends et al. (2003) argue persuasively that the improvement in stereo slant discrimination that can occur when eye movements are permitted does not derive from improved retinal measurement of the horizontal disparity gradient. Therefore, we interpret the effect of eye movements in our laboratory as being due to the additional weight of evidence provided to the system by DV, and DV may also account for the plateau observed by Berends et al.

Limitations in the Measurement of Retinal Disparity

Why is there a need for the DV mechanism at all? The mechanism that measures RRD is likely limited by the neuronal connections that are required for the comparison of corresponding image features on the two retinae (Enright, 1991a). Neurons in visual cortex respond selectively to stimulation of particular retinal locations, and binocular cells typically respond to simultaneous stimulation of corresponding locations—or nearly corresponding locations, in the case of neurons that respond to nonzero disparity—in the two retinae. Measurement of large absolute disparities requires that discrepant locations be compared between the retinæ; measurement of relative disparity for widely separated targets requires the comparison of absolute disparities across large regions of the visual field. In either case, long distance connections are required between regions of cortex that represent different parts of the visual field, and providing for all possible pairwise combinations of such measurements, at all spatial scales, is surely infeasible.

Thus, it might be desirable for the visual system to find some other method for, such as DV, measuring relative disparity that is not limited by the density of connections between corresponding locations in cortex.
Conclusions

We directly measured observers’ relative reliance on two distinct methods (based on vergence change or retinal disparity alone) by which the depth interval between a pair of isolated targets can be either nulled (Experiment 1) or matched (Experiment 2). The signal that the visual system had to measure to do these tasks was relative disparity between the targets. As the visual angle separating the targets increased, depth nulling was based increasingly on vergence change (or delta vergence, DV). As the relative disparity between the targets was increased, depth matching was based increasingly on DV for some observers, but not for others. These results extend previous work that found vergence to play a significant role in depth nulling tasks.

Acknowledgments

We thank James Hillis for serving as an observer, and other members of the Penn Perception group for helpful discussion. This research was supported by National Institutes of Health Grant EY-013988 (BTB). Commercial relationships: none.

References


