

# Binocular global motion perception is improved by dichoptic segregation when stimuli have high contrast and high speed

Lanya T. Cai

Graduate Center for Vision Research, SUNY College of Optometry, New York, NY, USA



Alexander E. Yuan

Graduate Center for Vision Research SUNY College of Optometry, New York, NY, USA



Benjamin T. Backus

Graduate Center for Vision Research, SUNY College of Optometry, New York, NY, USA



The brain combines information from the two eyes during vision. This combination is obligatory to a remarkable extent: In random-dot kinematograms (RDKs), randomly moving noise dots were similarly effective at preventing observers from seeing the motion of coherently moving signal dots, independent of whether the signal and noise were presented to the same eye or segregated to different eyes. However, motion detectors have varied binocularity: Neurons in visual brain area V1 that encode high contrast, high speed stimuli may be less completely binocular than neurons that encode low contrast, low speed stimuli. Also, neurons in MT often have unbalanced inputs from the two eyes. We predicted that for high contrast, high speed stimuli only, there would be a benefit to segregating the signal and noise of the RDK into different eyes. We found this benefit, both when performance was measured by percent coherence thresholds and when it was measured by luminance contrast ratio (signal-dot-contrast to noise-dot-contrast) thresholds. Thus, for high contrast, high speed stimuli, binocular fusion of local motion is not complete before the extraction of global motion. We also replicated a cross-over interaction: At high speed, global motion extraction was generally more efficient when dot contrast was high, but at low speed it was more efficient when dot contrast was low. We provide a schematic model of binocular global motion perception, to show how the contrast-speed interaction can be predicted from neurophysiology and why it should be exaggerated for segregated viewing. Our data bore out these predictions. We conclude that different neural populations limit performance during binocular global motion perception, depending on stimulus contrast and speed.

## Introduction

Single vision requires that a binocular visual system must suppress one eye or merge (fuse) the inputs from the two eyes. In humans fusion is often obligatory: It is difficult to determine the eye of origin for visual stimuli (Enoch, Goldmann, & Sunga, 1969; Barbeito, Levi, Klein, Loshin, & Ono, 1985), and noise in one eye makes it difficult to see signal in the other (Legge, 1979; Hess, Hutchinson, Ledgeway, & Mansouri, 2007).

Here we study the rules by which the local motions of dots, seen by the two eyes, respectively, are combined to arrive at a global motion percept in people with normal binocular vision. Binocular combination of *static* stimuli has been studied in some detail (Ding & Sperling, 2006; Meese, Georgeson, & Baker, 2006; Baker, Meese, Mansouri, & Hess, 2007; Ding, Klein, & Levi, 2013; Ding & Levi, 2015; Ding & Levi, 2016; Georgeson, Wallis, Meese, & Baker, 2016; Yehezkel, Ding, Sterkin, Polat, & Levi, 2016; Ding & Levi, 2017), but the binocular combination of motion signals during global motion perception has not, and global motion perception does not necessarily use the same mechanisms of binocular combination as static pattern perception. Random-dot motion displays have been used not only to measure normal binocular global motion perception, but also to treat interocular imbalance in people with amblyopia (e.g., Thompson & Liu, 2006; Hess, Mansouri, & Thompson, 2010; J. Li et al., 2011; Huang, Zhou, Lu, & Zhou, 2011; J. Li et al., 2013). It is therefore important to understand how motion signals combine binocularly in normal vision.

Hess et al. (2007) previously compared performance across monocular, segregated (“dichoptic”—signal and noise dots in different eyes), and duplicate (“binocu-

Citation: Cai, L. T., Yuan, A. E., & Backus, B. T. (2019). Binocular global motion perception is improved by dichoptic segregation when stimuli have high contrast and high speed. *Journal of Vision*, 19(13):10, 1–17, <https://doi.org/10.1167/19.13.10>.



lar”—identical display in the two eyes) viewing conditions. There were no significant differences between these viewing conditions (except at low contrast, which was predicted from binocular contrast summation but not motion combination per se). In other words, the randomly moving noise dots were equally effective for masking signal dots whether the noise and signal were in the same eye or different eyes. Hess et al. (2007) therefore concluded that the spatial integration needed for global motion perception occurs after binocular fusion. However, for reasons described in the next section, we predicted that noise dots would be more effective when shown to the same eye as signal dots, *but only when stimuli had both high contrast and high speed*. In three experiments we tested this conjecture and found it was correct. The situation therefore requires additional explication.

Thresholds in random-dot kinematogram (RDK) global motion stimuli can be measured for either of two variables: percent coherence or signal-to-noise contrast ratio. We measured both, in order to (a) provide redundancy to make the findings more robust, and (b) provide normative data for both types of stimulus. In Experiments 1 and 3, we used the same contrast for signal dots and noise dots and measured the threshold for global percent coherence under varied conditions of contrast, speed, and binocular viewing. In Experiment 2, we did the reverse: We fixed global percent coherence and measured the threshold ratio for signal-dot-contrast to noise-dot-contrast under viewing conditions used in Experiment 1. Additionally, we ran a duplicate condition as part of Experiment 2 to confirm that our stimuli were of sufficiently high contrast to be out of the regime of binocular contrast summation. In this condition, each dot was presented binocularly (thus doubling the overall number of dots in the display).

The simplest question we address is whether global motion mechanisms can exclude noise from one eye, which would be indicated by lower thresholds during segregated viewing (signal and noise to different eyes). We addressed this question by comparing performance in the segregated condition to performance in the monocular and binocular conditions. Anticipating our result, those conditions gave rise to similar performance, except when dots had high contrast and high speed, in which case segregated viewing was beneficial. Several additional viewing conditions (Table 1 and Figure 2) allowed us to refine our interpretation of this result.

Aside from global motion mechanisms, the lower thresholds during segregated viewing could have been attributed to interocular gain control mechanisms linked to the total number of dots in each eye. The segregated viewing condition usually had fewer dots in the signal eye and more dots in the noise eye, whereas the binocular viewing condition always had the same

number of dots in the two eyes. In Experiment 3, both a “segregated signal-balanced” condition and a “segregated signal-imbalanced” condition were added to test the effects of total number of dots in each eye.

## Model of binocular combination for RDK stimuli

Koch, Jin, Alonso, and Zaidi (2016) reported a functional difference between two topologically identified groups of neurons in cat primary visual cortex. Contrast saturation and cross-orientation suppression were found to be stronger within iso-orientation domains, where neurons are largely binocular, than at pinwheel centers, where they are largely monocular. In the Discussion we show how it follows that neurons in the iso-orientation domains (ocular dominance columns) are suited to the representation of high contrast, high speed stimuli, whereas neurons in the pinwheels are better suited to low contrast, low speed stimuli. We now consider the implications for global motion perception, under the assumption that both classes of neuron contribute to global motion perception, with performance in a motion discrimination task being limited by one or the other class of neuron depending on the stimuli.

Accordingly, in Figure 1 we propose a six-stage schematic model of binocular combination for global motion perception, which we will call the Cai-Backus binocular combination (CBBC) model. The input in the model is a dichoptic RDK stimulus, and the output is a decision to respond “up” or “down” in a 2AFC task. First, the retinae transduce the stimulus from light to neural representation. Second, monocular gain controls, based on stimulus luminance and contrast, act within monocular representations at the retina, LGN, and area V1 to modulate the strength of visual signals used in the third stage: extraction of local motion from monocular signals, which is done in parallel by two mechanisms in area V1. We identify these two mechanisms primarily with two types of domain in area V1, respectively: monocular pinwheels (high speed, high contrast) and binocular nonpinwheels (low speed, low contrast).

In the fourth stage, regional motion is extracted by neurons in area MT/V5, using outputs from the two local motion mechanisms in V1. In fact, V2 and MT cells also receive monocular input directly from LGN, and MT receives binocular input from V2 (Felleman & Van Essen, 1987; Born & Bradley, 2005). What is important here, however, is that multiple mechanisms with different response characteristics encode the motion, and because known properties of the two V1 mechanisms are qualitatively consistent with the data,

Experiment	S-to-N contrast ratio	Percent coherence	Weber contrast	Viewing conditions	Dot motion	Frame rate (Hz)	Stimulus duration (ms)	Dot speed (°/s)	Number of observers
1	1	S.V.	1.0, 0.1	M, S, B	Two-frame	30	300	1.2 3.3 8.1	7
2	S.V.	8	n/a	M, S, B, D	Continuous	60	300	4.2 8.4 12.6 16.8	7
3	1	S.V.	1.0	M, S, B, SSB, SSI	Two-frame	30	300	8.1	10

Table 1. Design parameters of Experiments 1 through 3. Notes: S.V. = staircase variable, M = monocular, S = segregated, B = balanced, D = duplicate, SSB = segregated-signal balanced, and SSI = segregated-signal imbalanced.

we refer to them by their V1 correlates (pinwheel and nonpinwheel). To convert regional motion estimates into a response, the model has a fifth stage that constructs a global motion percept and a sixth stage that decides to respond “up” or “down.”

The local and regional motion detectors in this model, that are identified with direction-selective cells in areas V1 and MT/V5 of the primate visual system, respectively (Maunsell & Van Essen, 1983a, 1983b; Movshon & Newsome, 1996), are the distinguishing features of the model. After monocular gain control based on luminance and contrast, local motion is

extracted in the monocular pinwheel and binocular nonpinwheel domains in V1, whose outputs are sent to area MT where regional motion detection occurs. Importantly, the regional motion mechanisms, identified with MT neurons, have varied ocular dominance. Additional computation is then necessary to report whether the net motion of the entire stimulus is upward or downward; we assume this decision is based on the location of peak activity within a bank of global motion detectors that we identify with neurons in primate visual area MST, each of which inherits its direction tuning from its afferent regional-motion

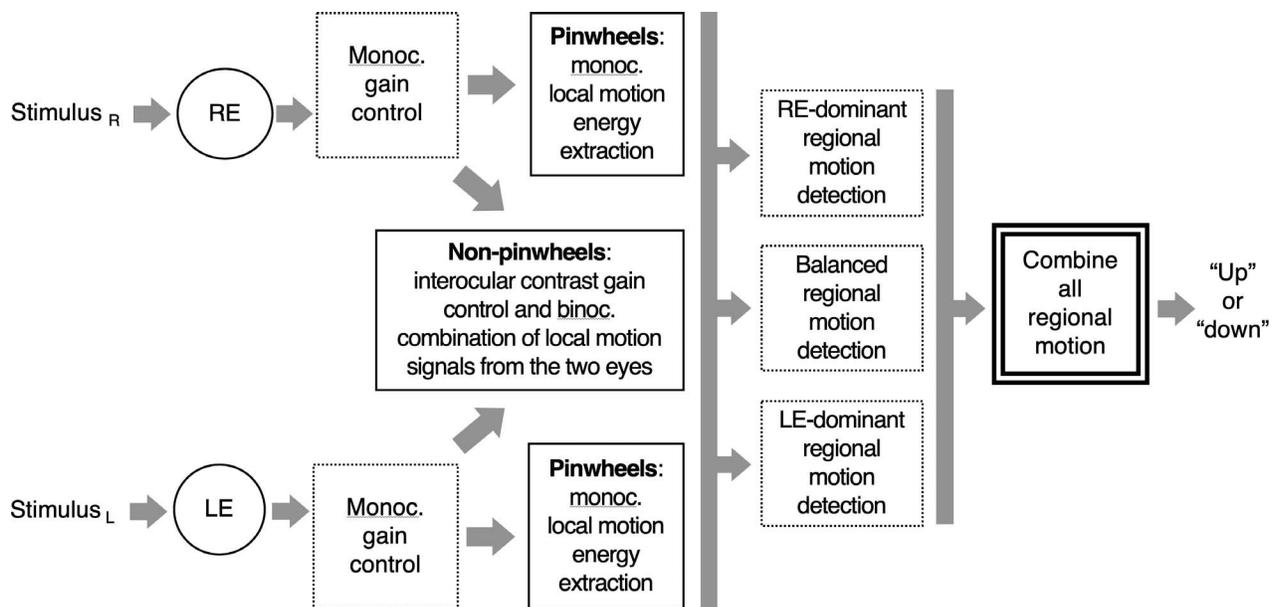


Figure 1. Schematic “CBBC” model of binocular global motion perception for random-dot kinematograms. The model has six cascading stages and produces a decision to respond “up” or “down” on a given trial of the experiment. Stages represent transformation of the visually acquired information from one representation to another that can be identified with neuroanatomical structures in the visual system; (a) stimuli enter the eyes; (b) monocular gain control, based on luminance and contrast, modulates signal strength in retina, LGN, and V1; (c) monocular local motion extraction occurs in parallel along two processing streams: monocular pinwheel domains and binocular nonpinwheel domains in V1; (d) regional motion is detected by mechanisms with varied ocular dominance (MT); (e) regional motions are combined to estimate global motion (MST); and (f) a decision is made to respond “up” or “down” based on the estimate of global motion.

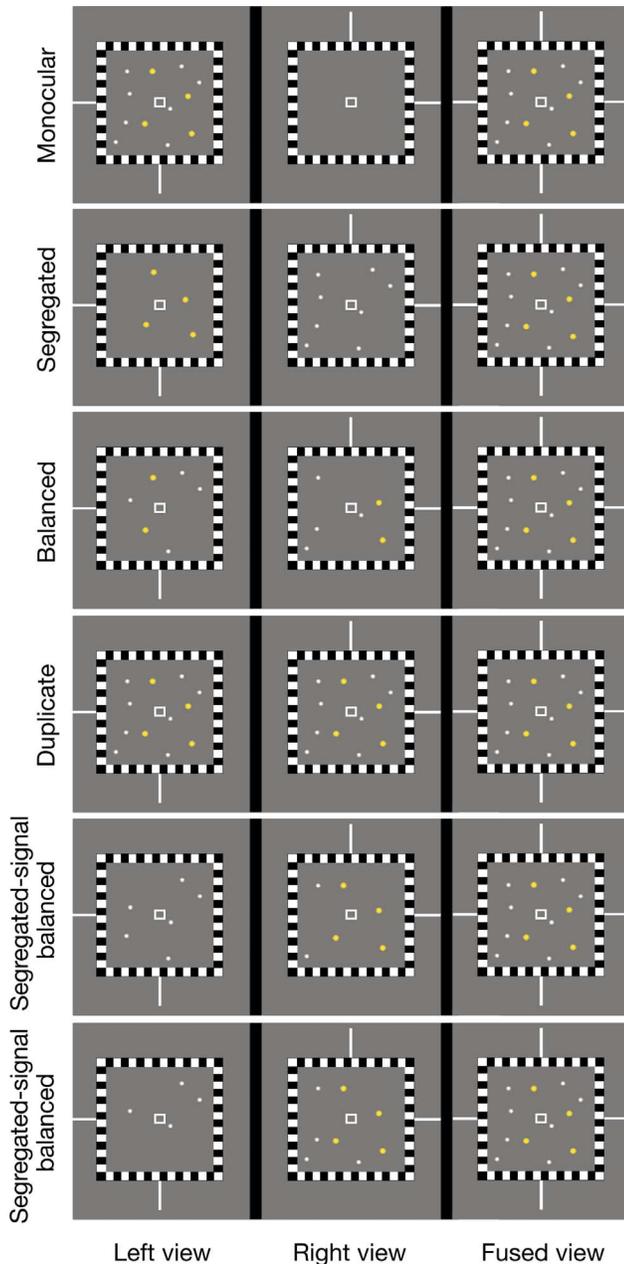


Figure 2. The distribution of signal and noise dots in the two eyes, for the six viewing conditions used variously in the three experiments of this study. For illustration, the large yellow dots and small white dots represent signal and noise, respectively, but in the actual experiment all the dots were identical in size and chromaticity.

subunits (Ungerleider & Desimone, 1986; Van Essen & Gallant, 1994; Gilaie-Dotan, 2016).

The CBBC model makes it explicit that there could be imbalance between the eyes at the input to and within (MT) regional motion detectors and shows how this imbalance could interact with two classes of (V1) local-motion-selective detectors. The model does not make a specific prediction; rather, it reveals that a

segregation effect at high speed and high contrast is possible.

We assume that the local motion signals used by observers in our experiments were extracted monocularly, by mechanisms that have relatively small receptive fields. Even though local motion can also be extracted after binocular fusion (Shadlen & Carney, 1986; Carney & Shadlen, 1992; Carney & Shadlen, 1993; Hayashi, Nishida, Tolia, & Logothetis, 2007; Hayashi, Miura, Tabata, & Kawano, 2008), this postfusion mechanism has low efficiency. Since motion signals were available to monocular mechanisms in our stimuli, motion would have been extracted by the more efficient monocular mechanisms so that performance at threshold would have used monocularly extracted local motion signals. The strength of the monocular local motion signals in the CBBC model is modulated by monocular and binocular gain control processes, both of which must be taken into account to interpret the data (see Discussion).

## General methods

### Observers

Fourteen human observers (four females and 10 males) between 20 and 53 years old with normal or corrected-to-normal vision participated in the study (mean age  $\pm$  SD = 29.4  $\pm$  6.5). All observers were students or faculty in the graduate programs at the SUNY College of Optometry; their vision status was self-reported. All observers passed the Randot stereogram test at 20 arcsec, confirming binocular vision. Observers S1 through S7 participated in Experiment 1; observers S1, S2, S4, S5, S6, S8, and S9 participated in Experiment 2; observers S1, S2, S4, S6, S8, and S10 through S14 participated in Experiment 3. Three of the observers (S1, S2, and S3) were authors of the present study, and the others were naïve to the purpose of the experiments.

### Apparatus

Visual stimuli for the left and right eyes were displayed side by side in a mirror stereoscope on an HP 23-in. LED-backlit LCD monitor. Stimuli were controlled by MATLAB (MathWorks, Natick, MA) with the Psychophysics Toolbox Version 3 extensions (Brainard, 1997; Pelli, 1997; Kleiner, Brainard, & Pelli, 2007) on a PC running under the Windows 7 operating system. The display was refreshed at 60 Hz, so the time between video frames was 16.7 ms. The monitor resolution was 1920  $\times$  1080 pixels. The left and right

halves of the monitor were only visible to the left and right eyes, respectively. To help lock fusion, stimuli for each eye was displayed in a  $10^\circ$  aperture surrounded by a high contrast checker frame (see Figure 2), and observers were instructed to maintain the alignment of nonius lines while fixating. The viewing distance was 114 cm. The lowest and the highest luminance levels of the monitor ( $0.4 \text{ cd}\cdot\text{m}^{-2}$  and  $228.3 \text{ cd}\cdot\text{m}^{-2}$ ) were defined as black and white, respectively. The background color was set to gray at 25% of white ( $57 \text{ cd}\cdot\text{m}^{-2}$ ) for the entire study. Head position of the observer was stabilized using a chinrest. Observers were instructed to respond to the visual stimulus on each trial by pressing one of two buttons on a number pad. A pair of speakers provided auditory feedback to indicate if the response was correct. To prevent visual distraction, the experiment room and furniture were black, and the room light was off during the experiment.

## Procedure

The procedures complied with The Declaration of Helsinki and were approved by the Institutional Review Board at the SUNY College of Optometry. The general purpose of the study was explained to each observer, and written consent was obtained prior to the observer's first experimental session. The study consisted of three experiments, all using random-dot kinematograms stimuli. On a given trial in any experiment, "signal dots" all moved in the same direction, either up or down. "Noise dots" either (a) moved in random directions at the same speed as signal dots, in the case of "continuous motion", or (b) were single dots displayed for one stimulus frame, in the case of "two-frame motion." In two-frame motion, the maximum lifetime of each dot was 67 ms (two stimulus scenes consisted of four display frames). Dot positions were not updated on each refreshed frame of the 60 Hz display, but rather on every other frame, so dot lifetime was either 33 ms (stationary noise dots lasting two display frames) or 67 ms (signal dots lasting four display frames, that moved between frames 2 and 3).

The different viewing conditions distributed signal and noise dots differently to the two eyes, as illustrated in Figure 2. Six viewing conditions are schematized in Figure 2; they differed from one another by adhering to the following constraints:

- Monocular (M): all 100 dots displayed to just one eye.
- Segregated (S): all signal dots displayed to one eye and all noise dots to the other eye.
- Balanced (B): both signal and noise dots distributed equally to both eyes; 50 dots total in each eye.
- Duplicate (D): all 100 dots displayed to each eye; the two eyes had duplicate scenes.

- Segregated-signal balanced (SSB): signal and noise dots in one eye, but only noise dots in the other eye, with 50 dots total in each eye.
- Segregated-signal imbalanced (SSI): signal and noise dots in one eye, but only noise dots in the other eye, with 76 and 24 dots in these two eyes, respectively.

Table 1 shows how the experiments differed from one another in terms of the viewing conditions tested and other parameters such as the type of dot motion, and the staircase variable used to measure threshold. For both percent coherence (Experiments 1 and 3) and signal-to-noise contrast ratio (Experiment 2) measurement, each threshold was measured by an 80-trial 3-down-1-up staircase. The stimulus duration of each trial was always 300 ms. Signal motion was upward or downward, at random, on a given trial. The observer's task was to respond seeing upward or downward global motion of the stimulus by pressing buttons. Then from each staircase, a threshold could be estimated by fitting a probit psychometric function that maximized likelihood of the staircase data.

In the case of percent coherence threshold measurement, to compensate for bias in responding, the fraction coherence threshold was an average of two percent-coherence values: that for 82% upward responding and that for 82% downward responding. These values were obtained by fitting a probit curve that fitted the probability of answering "upward motion" as a function of fraction coherence ranged from  $-1$  (when all dots moved downward) to  $1$  (when all dots moved upward). In the case of signal-to-noise contrast ratio threshold measurement, the psychometric function was fitted as the probability of correct answer, and threshold was estimated as the signal-to-noise contrast ratio at 82% correct.

In the monocular, segregated, segregated-signal balanced, and segregated-signal imbalanced conditions, both situations of signal-in-left-eye and signal-in-right-eye were tested with intermingled trials. This design discouraged a closing-one-eye strategy in binocular tasks. During data analysis, thresholds from signal-in-left-eye and signal-in-right-eye were averaged within each viewing condition.

Each threshold was measured twice in Experiments 1 and 3, or four times in Experiment 2, and the results were averaged.

## Experiment 1

### Methods

In Experiment 1, we measured the threshold for global percent coherence at two levels of dot contrast

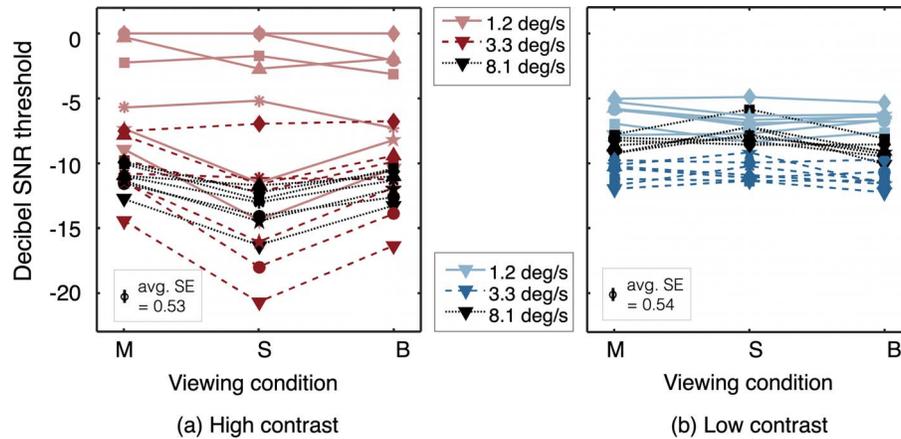


Figure 3. Threshold SNR comparison between monocular, segregated, and balanced conditions. A more negative threshold SNR reflects lower coherence threshold and better performance. The two panels plot data when all dots in the display had either high contrast (a) or low contrast (b). Each series connects three data points that are the threshold SNRs from one observer when tested at a given dot displacement. Data from each observer are shown with a unique marker shape. Colored solid lines, dashed lines, and black dotted lines plot data for dot displacements of  $1.2^\circ/\text{s}$ ,  $3.3^\circ/\text{s}$ , and  $8.1^\circ/\text{s}$ , respectively.

(Weber contrast 1.0 and 0.1), three levels of dot displacement (8, 12, and 20 pixels per frame, equivalent to  $1.2^\circ/\text{s}$ ,  $3.3^\circ/\text{s}$ , and  $8.1^\circ/\text{s}$ ), and in three viewing conditions (monocular viewing with stimulus in the left or the right eye, segregated viewing with signal in the left or the right eye and noise in the other eye, and balanced viewing). Combining across eye of presentation there were  $2 \times 3 \times 3 = 18$  conditions. The signal-to-noise contrast ratio, defined as the contrast of the signal dots divided by the contrast of the noise dots, was set to 1.

With long dot lifetimes, single dots or a small group of dots can be attended to and tracked over time, which could make it easier to identify the motion direction of signal dots. To eliminate selective tracking as a possible strategy, we used two-frame dot motion, for which the maximum possible percent coherence, defined as the percentage of dots on a refresh frame that are replaced by a paired dot in the defined motion direction, was 50%. In this two-frame motion stimulus, any signal dot would have disappeared before it could be tracked using attention. We tested one signal-dot displacement magnitude in a given session, and two contrast levels (blocked), while the viewing conditions were randomly intermixed.

## Results

Figure 3 plots individual observers' coherence thresholds from Experiment 1 as a function of viewing condition.

Panels (a) and (b) show data from the high and low contrast levels, respectively. For the monocular and

segregated conditions, data from situations of signal in left eye and signal in right eye were averaged.

Panel (a) shows that performance in the segregated condition was better as compared to the monocular and balanced conditions at high contrast, but not at low contrast. This benefit of segregation was most evident for the two larger displacements (speeds). Thresholds of slow dots ( $1.2^\circ/\text{s}$ ) are also generally higher in Panel (a) than in Panel (b), whereas thresholds of medium and fast dots ( $3.3^\circ/\text{s}$  and  $8.1^\circ/\text{s}$ ) are generally higher in Panel (b). These results agree with Seitz, Pilly, and Pack (2008), who reported that increasing the luminance contrast caused an improvement in performance (direction judgments became more accurate) for large two-frame displacements (fast dots), but worsened performance for small displacements (slow dots).

To ascertain the reliability of this difference, we did a repeated measures ANOVA on threshold SNRs with Dot Displacement, Contrast, and Viewing Condition as factors. The first order effects were significant for Dot Displacement,  $F(2, 12) = 49.6$ ,  $p < 0.0001$ , and Viewing Condition,  $F(4, 24) = 11.5$ ,  $p < 0.0001$ , but not for Contrast,  $F(1, 6) = 1.06$ ,  $p = 0.34$ . The interaction between Dot Displacement and Viewing Condition was not significant,  $F(8, 48) = 1.27$ ,  $p = 0.28$ . However, the interaction between Contrast and each of the other two factors was significant,  $F(2, 12) = 12.6$ ,  $p = 0.001$  for Contrast and Dot Displacement;  $F(4, 24) = 10.2$ ,  $p < 0.0001$  for Contrast and Viewing Condition. The three-way interaction between Dot Displacement, Contrast, and Viewing Condition was also significant,  $F(8, 48) = 5.02$ ,  $p = 0.0001$ .

We found that large individual differences are a particular characteristic of the low speed, high contrast condition. For example, in Figure 3a some data have a

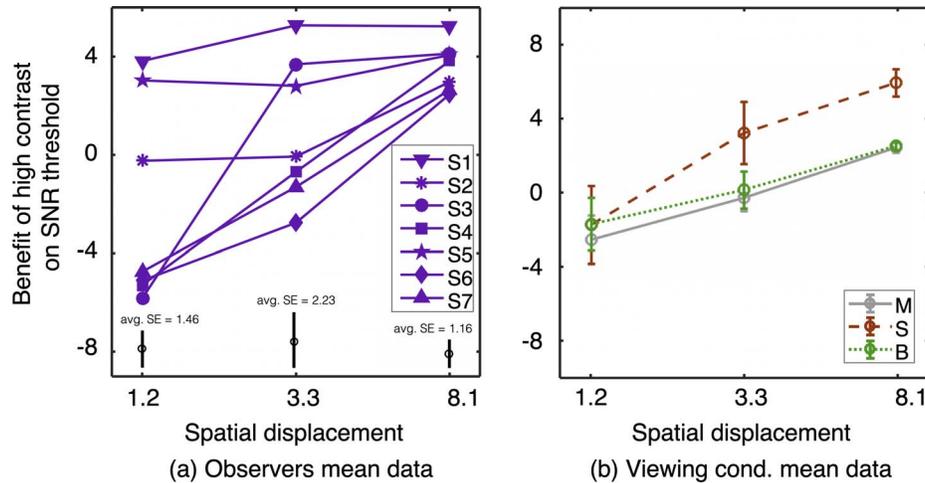


Figure 4. Difference of threshold SNR between low and high contrasts as a function of displacement. A positive difference reflects a benefit (better performance) at high contrast. (a) Observers' average performance across viewing conditions. (b) Viewing conditions' average performance with the standard error across observers. The three colors and line styles (gray solid line, red dashed line, and green dot line) correspond to the three groups of viewing condition (“M” for monocular, “S” for segregated, and “B” for balanced). Error bars in panel (a) show the mean of the subject SEs at each spatial displacement. Error bars in panel (b) are SE across subjects.

decibel SNR of 0 for threshold: The observer was at the maximum SNR, with a value of 1. It was therefore reasonable to redo ANOVA excluding the lowest speed condition. After this adjustment, the first order effects were significant for Viewing Condition,  $F(4, 24) = 9.40$ ,  $p = 0.0001$ , and Contrast,  $F(1, 6) = 10.8$ ,  $p = 0.017$ , but not for Dot Displacement,  $F(1, 6) = 4.83$ ,  $p = 0.07$ . The interaction between Dot Displacement and Viewing Condition was not significant,  $F(4, 24) = 1.84$ ,  $p = 0.154$ . The interaction between Contrast and each of the other two factors was still significant,  $F(1, 6) = 11.5$ ,  $p = 0.015$  for Contrast and Dot Displacement;  $F(4, 24) = 19.1$ ,  $p < 0.0001$  for Contrast and Viewing Condition. The three-way interaction between Dot Displacement, Contrast, and Viewing Condition was no longer significant,  $F(4, 24) = 0.28$ ,  $p = 0.888$ .

Figure 4 replots the data from Figure 3 as a difference in threshold SNR between low and high contrasts, as a function of dot displacement.

In Figure 4, panel (a), the effect of contrast is plotted as a function of spatial displacement for each observer separately, collapsing across all viewing conditions. A positive value for the ordinate indicates a benefit at high contrast, and a negative value indicates a benefit at low contrast. Four observers (S3, S4, S6, and S7) had worse performance at high contrast at the smallest displacement but had better performance at high contrast at large displacement, in agreement with Seitz et al. (2008); two observers (S1 and S5) had better performance at high contrast for all displacements; and one observer (S2) had better performance at high contrast, but only at the largest displacement. Figures 3 and 4 show that observers behaved more similarly to

one another at high speed than at low speed, both in terms of absolute threshold (Figure 3) and in terms of the effect of contrast (Figure 4).

Figure 4, panel (b) plots the mean effect of contrast across observers as a function of spatial displacement, with a separate series for each viewing condition. Within every viewing condition the contrast effect went from negative to positive as the spatial displacement was increased from  $1.2^\circ/s$  to  $8.1^\circ/s$ , showing that the cross-over interaction between contrast and spatial displacement is robust to the viewing condition. Yet the slope of the segregated-condition data appears to be larger than the slopes for the other two conditions. This was confirmed by a posthoc  $t$  test: size of Contrast Benefit at Dot Displacement of 1.2 vs. 8.1,  $t(6) = -3.7081$ ,  $p = 0.01$ .

## Experiment 2

### Methods

Experiment 1 showed an advantage for segregating signal and noise to different eyes when contrast was fixed, and percent coherence was the staircase variable. To test the robustness of this finding, in Experiment 2 we measured the effect of viewing condition using a fixed percent coherence (8%) and continuous dot motion, with contrast ratio as the staircase variable. We tested seven observers at four dot speeds ( $4.2^\circ/s$ ,  $8.4^\circ/s$ ,  $12.6^\circ/s$ , and  $16.8^\circ/s$ ) under monocular, segregated, and balanced conditions, and a new “duplicate”

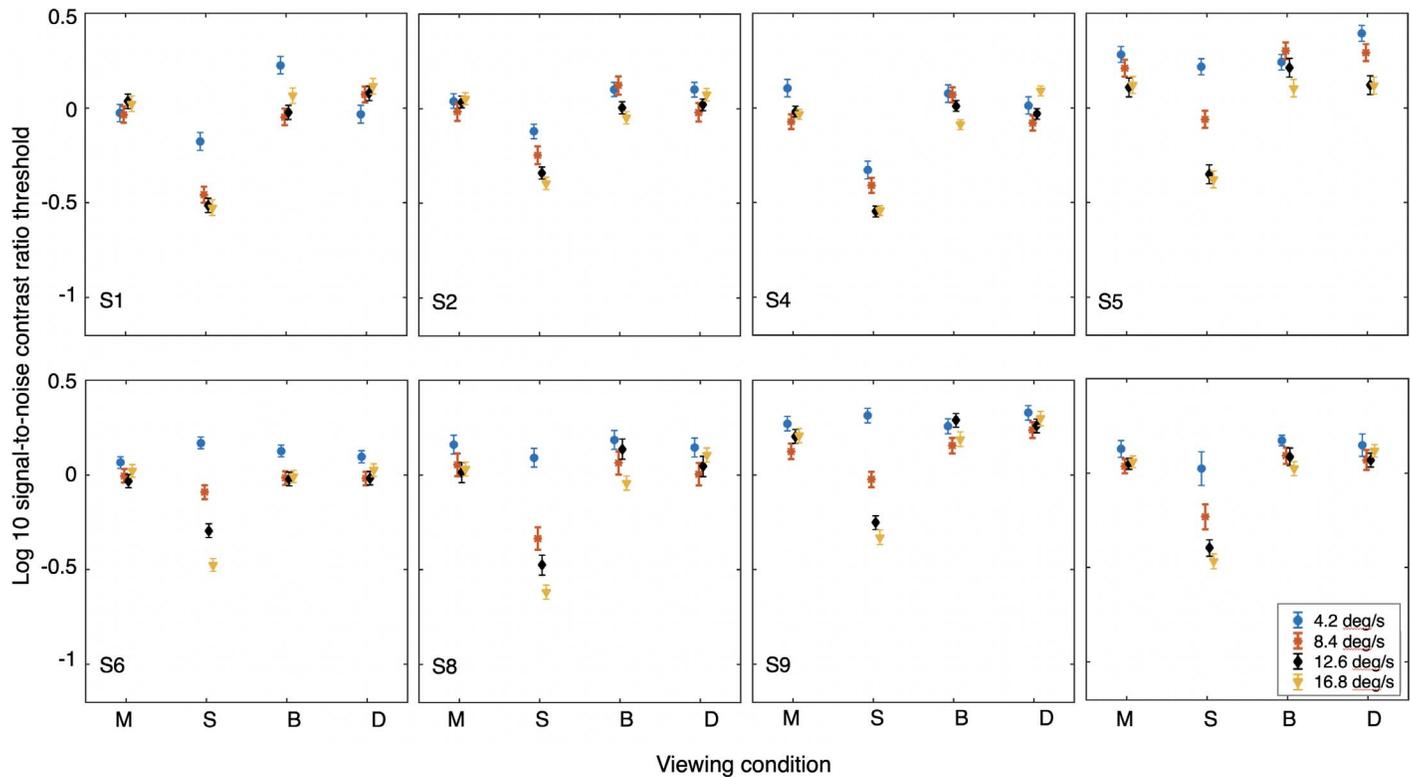


Figure 5. Threshold  $\log_{10}$  signal-to-noise luminance contrast ratio comparison among monocular (M), segregated (S), balanced (B), and duplicate (D) conditions for display of signal and noise dots. Blue circles, red stars, black diamonds, and yellow triangles are data from four spatial displacements ( $4.2^\circ/s$ ,  $8.4^\circ/s$ ,  $12.6^\circ/s$ ,  $16.8^\circ/s$ ), respectively. The first seven panels are individual observers' data, and the last panel is the average performance across the seven observers. Error bars are SEs.

condition in which each eye saw the same set of 100 dots that included both signal and noise.

## Results

Figure 5 plots the 82% correct threshold signal-to-noise luminance contrast ratio in  $\log_{10}$  units for the monocular, segregated, balanced, and duplicate conditions. For the monocular and segregated conditions, data were pooled across signal in the left eye and signal in the right eye. A threshold value of 0 indicates that the luminance of signal and noise were the same. A positive or a negative threshold indicates that the signal dots had greater or lower luminance than the noise dots, respectively. Experiment 2 shows that the segregation benefit was robust to the choice of staircase variable used to assess it.

All observers showed a remarkable segregation benefit that was amplified at higher spatial displacements. We ran a repeated measures ANOVA with Spatial Displacement and Viewing Condition as factors. The threshold value for signal-to-noise contrast ratio was significantly lower (better) with larger spatial displacement,  $F(3, 12) = 13.2$ ,  $p = 0.0004$ , and in the segregated viewing condition,  $F(5, 20) = 130.3$ ,  $p <$

$0.0001$ . The interaction between Spatial Displacement and Viewing Condition was also significant,  $F(15, 60) = 6.375$ ,  $p < 0.0001$ .

By comparison, performance in the duplicate condition was similar to performance in the monocular and binocular conditions. This result replicates Hess et al. (2007), who previously compared performance across monocular, segregated, and duplicate viewing conditions and found no difference except at threshold-level luminance contrast. In the CBBC Model (Figure 1), the effect of the duplicate condition is to place more motion energy into the two eyes' Pinwheels regions, and to increase the motion energy in the binocular Nonpinwheel regions. Evidently, our displays had sufficiently high contrast not to show a benefit of binocular summation; the signal-to-noise ratio in the representation of the stimulus was equal to the signal-to-noise ratio in the display itself. As expected, performance was not different for the duplicate condition.

Results from Experiments 1 and 2 demonstrate a benefit of segregating signal and noise dots into different eyes at high contrast and high spatial displacement. Three ways this benefit could come about are monocular luminance gain control, interocular contrast gain control, and/or monocular bias in global

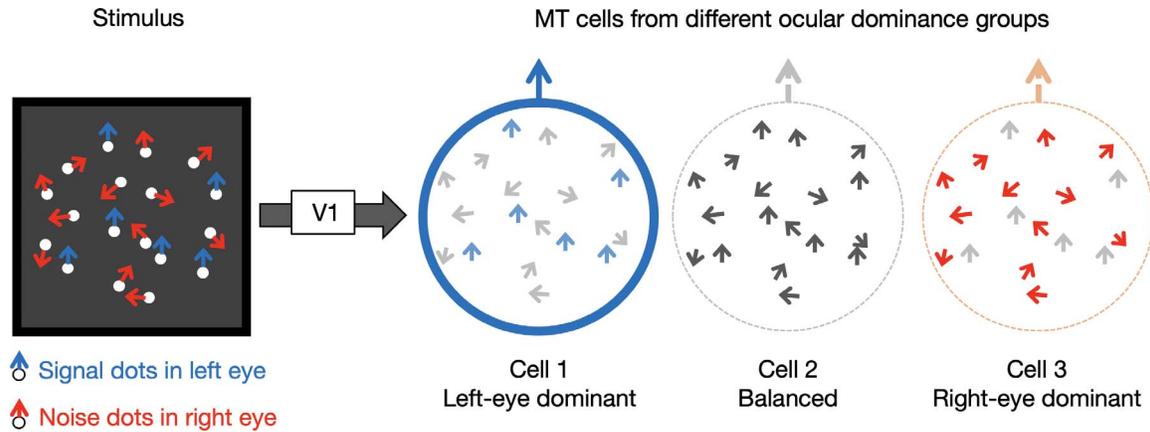


Figure 6. Interocular imbalance of the inputs to regional motion integrators, such as single neurons in area MT, could lead to a benefit for detecting signals that are segregated to a different eye from noise. The square panel on the left shows the stimulus with signal dots (with blue arrows) in the left eye and noise dots (red arrows) in the right eye. Area V1 then extracts monocular local motion signals, which contribute to the receptive fields of three regional motion detectors, presumably in area MT, shown by the three circular panels on the right. These regional detectors are integrators that all respond to upward signal, but they have different ocular dominance. Cell 1 (left) gives greater weight to left-eye local motion signals. Cell 2 (middle) gives them equal weight. Cell 3 (right) gives greater weight to right-eye signals. Thus, Cell 1 is protected from noise in the right eye. Cells 1, 2, and 3 can be thought of as corresponding to cell types 1–3, 4, and 5–7, respectively, in the Hubel et al. (1968) ocular dominance scale, that runs from 1 (response to stimuli in left eye only) through 4 (equal response to either eye) to 7 (response to stimuli in right eye only).

motion integrators. The mechanisms to account for these effects are related to certain stages along the motion processing cascade. Monocular luminance gain control could explain the benefit in the sense that, because there were more noise dots in the noise eye than signal dots in the signal eye, each noise dot could have become relatively less effective as compared to a signal dot before binocular combination. This could happen at retina, LGN, or V1 (Shapley & Enroth-Cugell, 1984; Truchard, Ohzawa, & Freeman, 2000). Interocular contrast gain control is useful to explain the appearance of dichoptic stimuli (Ding & Sperling, 2006; Ding et al., 2013), but if it acts within V1 mechanisms, which are small in size, it is not likely to explain effects in our experiments because a given V1 neuron would typically have responded to at most one signal dot or noise dot, from just one eye, so it would not distinguish between stimuli from one eye or the other based on the number of dots per eye.

The third explanation for the segregation benefit is our new proposal, that it depends on monocular bias in regional motion integrators at a site downstream from V1. This idea is illustrated in Figure 6. In this example, a stimulus in the segregated viewing condition has signal dots displayed to the left eye and noise dots displayed to the right eye. Three hypothetical MT neurons with different ocular dominance receive the afferent input from V1. Cell 1 is categorized in ocular dominance groups 1 through 3 (Hubel & Wiesel, 1968) and has monocular bias towards the left eye; therefore it gives greater weight to local motion from the left eye and does not see noise dots well. Cell 2 is categorized in

ocular dominance group 4 and has no monocular bias, so it gives equal weight to all local motion. Cell 3 is categorized in ocular dominance groups 5 through 7 and has monocular bias towards the right eye, so it gives greater weight to local motion from the right eye and does not see signal dots well. Among all the three regional motion integrators, only Cell 1 would respond well to the signal in the left eye without much interference by noise dots in the right eye.

Downstream, a global motion detector for upward motion, such as a cell in area MST, that collects the responses from regional upward-motion detectors, would respond to the upward motion based on output from Cell 1. Similar imbalance in the regional detectors for downward motion would not cause any of those detectors to respond strongly to this stimulus, giving rise to better performance for segregated stimuli.

## Experiment 3

### Methods

In Experiment 3 we tested whether monocular biases in regional motion detectors and/or monocular gain control, induced by imbalance of dot numbers in the two eyes, contribute to the segregation effect. We compared performance in three conditions: (a) a “balanced” condition in which the number of signal dots was the same in each eye and the number of noise

dots was also the same in each eye; (b) a “segregated-signal balanced” (SSB) condition in which the total number of dots in each eye was the same, but with all signal in one eye; and (c) a “segregated-signal imbalanced” (SSI) condition, in which the noise eye always had 24 noise dots, and the signal eye always had 76 dots that included all signal dots together with the remaining noise dots. In condition SSB the maximum possible percent coherence was 25%, because half of the dots were shown to each eye, but all signal was in one eye. In condition SSI the maximum possible percent coherence was 38%.

If there is monocular bias, then putting all signal in one eye should be beneficial. Otherwise, monocular luminance gain control alone can be presumed responsible for the segregation benefit. On the other hand, if having fewer dots in the signal eye explained the segregation benefit in the S condition, then having fewer dots in the “pure-noise” eye would give benefit to dots in the noise eye, hurting performance in the SSI condition compared to the other conditions.

Monocular and segregated conditions were included as well, for comparison, so there were five conditions (M, S, B, SSB, and SSI). The threshold value for percent coherence was measured at high contrast (Weber contrast 1.0) and at the largest dot displacement from Experiment 1 (8.1°/s). Ten observers participated. Other aspects of the experimental design were the same as in Experiment 1.

## Results

Individual results and the average performance across observers are shown in the left and the right panels of Figure 7, respectively. This experiment replicated the segregation benefit from Experiments 1 and 2 in eight out of 10 observers (not S11 or S14). The SSB condition did not show any significant benefit as compared to the B condition, but for eight of 10 observers, performance in the SSI condition was indeed worse than in the B condition (not S4 or S6).

The right panel of Figure 7 shows observer means for each condition. Repeated measures ANOVA shows a significant effect of Viewing Condition,  $F(4, 36) = 5.631$ ,  $p = 0.0013$ . Then paired-sample  $t$  test was used to look at pairwise Viewing Condition effects. Performance in the balanced condition was similar to that of the monocular condition ( $p = 0.15$ ). Performance in the segregated condition was again better than in the other conditions (paired  $t$  tests comparing S to: M,  $p = 0.006$ ; SSI,  $p < 0.0001$ ; B,  $p = 0.046$ ; SSB,  $p = 0.035$ ). There was no segregation benefit in either the SSB or SSI condition, relative to M or B. Performance in the SSI condition was worse than in the B condition ( $p = 0.002$ ) and marginally worse than in the M condition ( $p =$

0.089). These results show that dots from the eye with fewer dots had a greater effect during binocular combination, no matter whether they were signal dots (S) or noise dots (SSI). Thus, some part of the segregation benefit can be attributed to monocular gain control that reduced the contributions of the dots from the eye that saw more dots

## General discussion

Efficiency of global motion perception after binocular combination can be measured by global percent coherence thresholds or by signal-to-noise contrast ratio thresholds. We made these measurements across a range of values for several stimulus parameters: luminance contrast, spatial displacement (dot speed), dichoptic viewing condition, and percent coherence. As expected, the data ruled out separate global motion estimates for each eye: When dots had low contrast or low speed, noise dots were equally effective to disrupt performance whether they were in the same eye as signal or the other eye; at high contrast and high speed, however, noise was more effective when presented in the same eye as signal. We now consider various effects and interactions between the stimulus parameters.

### Interaction between contrast and dot speed

We found, regardless of viewing condition, an interaction between luminance contrast and dot spatial displacement, in agreement with Seitz et al. (2008). High luminance contrast was helpful when the displacement of dots was large, but harmful when displacement was small. Seitz et al. (2008) gave a possible neurophysiological explanation for this effect, relating it to previous studies of single unit activity in monkey brain area MT. We extended their human psychophysical findings beyond monocular and binocular viewing conditions to a segregated viewing condition (Figure 4).

The main interaction is still of significant interest, quite apart from its further interaction with segregated viewing. In our experiments, the speed at which increased contrast went from being harmful to being helpful was about half of that reported by Seitz et al. (2008); this speed was between 0.13°/frame and 0.27°/frame (60 Hz) in the various conditions their experiments (from their Figures 1 and 2), and between 0.11°/frame and 0.27°/frame (30 Hz) in our experiments (Figure 4). These transition points are within the same order of magnitude as one another, and small differences could be expected given differences between

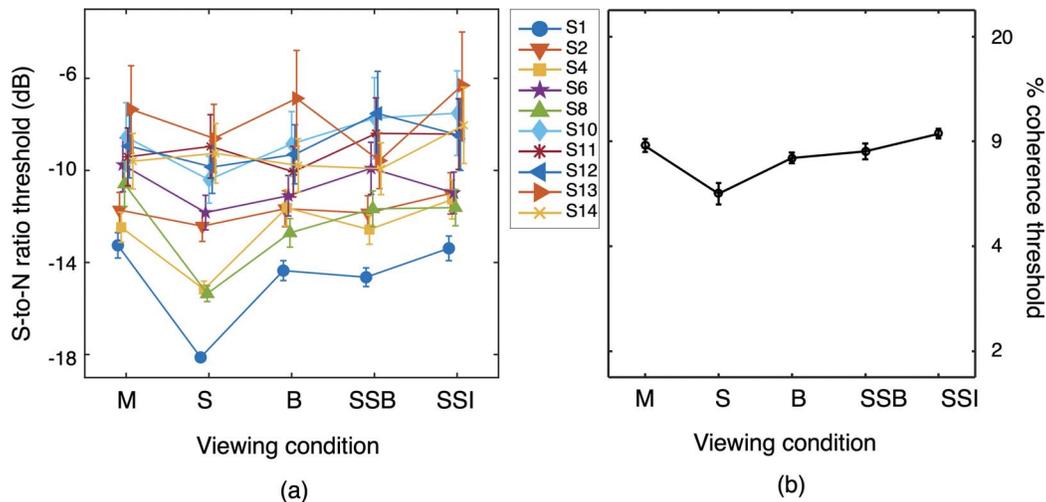


Figure 7. Signal-to-noise ratio and percent coherence threshold comparison between monocular (M), segregated (S), balanced (B), segregated-signal balanced (SSB), and segregated-signal imbalanced (SSI) viewing conditions. The left panel shows individual results; each color and marker shape show data from one observer. Each marker shows the average threshold across the observer's two eyes and across the two repeated measures. The right panel plots average performance across observers. The open black circles with error bars are population mean values within one standard error that was calculated after subtraction of the individual average across viewing conditions for each observer.

the experiments in background luminance and other stimulus attributes.

Seitz et al. explained this interaction as reflecting the operation of mechanisms that maximize information transmission efficiency by suppressing large slow stimuli with redundant information at high contrast, and preserving redundancy for better spatial and temporal pooling at low contrast, along lines similar to the explanation given by Tadin, Lappin, Gilroy, and Blake (2003) for their observation of greater surround-suppression at high contrast. Indeed, low contrast causes expanded spatial summation and sharpens the spatial frequency tuning of neurons in primate area V1 (Sceniak, Ringach, Hawken, & Shapley, 1999; Sceniak, Hawken, & Shapley, 2002). A difficulty for this explanation in the case of RDK stimuli, however, is that both signal dots and noise dots were affected by the changes in contrast. Why should increasing the contrast make it easier to extract the signal when, by definition, the signal-to-noise ratio in the stimulus did not change?

A plausible explanation is that in our stimuli and those of Seitz et al., the signal was narrowband, whereas noise was broadband. We suppose that two different populations of V1 neuron, tuned respectively to high contrast and high speed, and to low contrast and low speed, extracted motion energy from the stimuli. Because noise dots were unpaired (as opposed to having defined motion directions), they did not have a speed, and the speed of the signal dots would not have affected neuronal responses to noise dots. As a result, the signal-to-noise ratio at low contrast would have been better at low speed than at high speed, and

the signal-to-noise ratio at high contrast would have been better at high speed than at low speed.

This explanation could account not only for the interaction between contrast and speed, but also for the triple interaction whereby segregated viewing is additionally beneficial at high contrast and high speed (see Figure 4, above, and Relationship to V1 functional tiling, below). Hess et al. (2007) found no overall benefit of segregation at a dot speed of  $5.9^\circ/\text{s}$ . The benefit of segregation appeared in our experiments most clearly at higher dot speeds ( $8.1^\circ/\text{s}$  in Experiments 1 and 3; and  $8.4^\circ/\text{s}$ ,  $12.6^\circ/\text{s}$  and  $16.8^\circ/\text{s}$  in Experiment 2). Thus, dot speed can account for the difference in segregation benefits between our results and those of Hess et al.

## Monocular bias in motion integrators

In primate vision, most neurons beyond the input layer of primary visual cortex respond to input from either eye, so that the system is mostly binocular (Hubel & Wiesel, 1970; Poggio & Fischer, 1977; Maunsell & Van Essen, 1983c; Burkhalter & Van Essen, 1986; Felleman & Van Essen, 1987; Poggio, Gonzalez, & Krause, 1988; Roy, Komatsu, & Wurtz, 1992; DeAngelis, Cumming, & Newsome, 1998). To explain the segregation benefit in dichoptic RDK, we proposed a role for monocularly biased regional motion integrators (Figure 6). Our Experiment 3 was inconclusive; nevertheless, ocular dominance in area MT would enable a better extraction of motion signal in the

segregated viewing condition than in the monocular viewing condition or the balanced viewing condition, and this idea agrees with the known neurophysiology.

Maunsell and Van Essen (1983c) showed a distribution of ocular dominance in 91 MT neurons from macaque monkeys with normal binocular vision. In their data, about 45% of the neurons were unbiased (group 4: one eye caused within 1.5 times response of the other eye); 21% of the neurons had clear ocular dominance (groups 1, 2, 6, 7: for which one eye caused 3 times or more response than the other eye), although those with stronger ocular dominance had weaker responses in general. Kiorpes, Walton, O'Keefe, Movshon, and Lisberger (1996) and El-Shamayleh, Kiorpes, Kohn, and Movshon (2010) studied the distribution of ocular dominance of 218 neurons in area MT, also in macaque monkeys. Normally sighted and strabismic monkeys both had MT neurons distributed across the ocular dominance groups from left-eye dominance (groups 1 to 3) to right-eye dominance (groups 5 to 7), although the width of ocular dominance distribution of the normal monkeys was much narrower (figure 9 from Kiorpes et al. (1996), and figure 2 from El-Shamayleh et al. (2010)).

Ocular dominance in MT neurons presumably derives from asymmetric input from monocular cells in the lateral geniculate nuclei and cells with ocular dominance in cortical area V1. The absence of complete monocularity in MT, and/or the manner in which MT unit activities are read out, may account for the limited nature of the segregation effect in people with normal binocular vision—in other words, why it is that pure signal in one eye can be masked by pure noise in the other.

## Relationship to V1 functional tiling

While monocular bias in regional motion detectors could account for the benefit of segregation, it remains unexplained why this benefit was observed only when dots had both high contrast and high speed. It is unlikely that this dependence emerges at the level of the regional motion detectors themselves, because that would require the MT neurons to be of two sorts, with some MT neurons being tuned for high contrast and high speed and having monocularly biased inputs, and others being tuned for low contrast or low speed and having binocularly balanced inputs. We know of no evidence for that being the case. However, we speculate that known properties in the joint encoding of direction, speed, and contrast by neurons in V1 may explain this dependence: MT neurons may inherit their blindness to eye-of-origin, for low contrast and low speed stimuli, from upstream neurons in V1.

The argument goes as follows: In mammals, spatial frequency tuning, orientation tuning, and ocular dominance are not independently distributed in V1. Instead, neurons are clustered, so that when maps are drawn for each of these stimulus properties, the tuning patterns vary across the cortical surface. First, cortical iso-orientation domains are laid out geographically such that orientation tuning progresses either around a pinwheel, or linearly (nonpinwheel) (Bonhoeffer & Grinvald, 1991). Second, low spatial frequency regions overlap with low binocularity (strong ocular dominance), and high spatial frequency regions overlap with high binocularity (weak ocular dominance; Nauhaus, Benucci, Carandini, & Ringach, 2008; Nauhaus, Nielsen, & Callaway, 2016). The orientation tuning map is orthogonal to both the spatial frequency tuning map and the ocular dominance map, so orientation tuning is balanced across spatial frequency and ocular dominance. However, the specificity (narrowness) of the orientation tuning is not homogeneous. The pinwheel regions show tuning for low spatial frequency with less cross-orientation suppression and a more linear response to contrast, whereas nonpinwheel regions show tuning for high spatial frequency with more cross-orientation suppression and a response to contrast that saturates earlier Koch et al. (2016). These features of functional structure for the two types of region in V1 are listed in Table 2.

These response characteristics can be related to local motion extraction. For many V1 neurons, local motion direction selectivity arises from orientation selectivity and is orthogonal to it (Hubel & Wiesel, 1968; Adelson & Movshon, 1982; Albright, 1984; Movshon & Newsome, 1996; Gur & Snodderly, 2007), although the exact relationship between orientation maps and direction maps varies across species (reviewed by Nauhaus & Nielsen, 2014). As a result, we can identify orientation selectivity with motion direction selectivity. Furthermore, neurons that are broadly tuned for orientation will be broadly tuned to direction. At a given temporal frequency, low spatial frequency tuning is necessary for high-speed tuning, so in combination with the second property above, fast motions will tend to be detected by V1 neurons that show ocular dominance, while slower motions will be detected by neurons that are more binocularly driven.

Meier and Giaschi (2014, 2017) have recently shown that the time course for development of global motion perception in RDK during childhood is different for large- and small-displacement stimuli. Global motion was detected with adult-like performance in young children for stimuli with a dot displacement of 5 or 30 arcmin per frame (regardless of frame rate), but adult-like performance did not emerge for 1 arcmin displacements until about age 16. This finding provides additional evidence for the existence of two pathways

	Within ocular dominance columns (more monocular)	Between ocular dominance columns (more binocular)
Cortical organization of iso-orientation domains	Pinwheels with neurons tuned to many orientations	Progressive orientation tuning in a linear direction
Spatial frequency tuning	Prefer lower spatial frequency, more of low-pass response profile	Prefer higher spatial frequency, band-pass response profile
Cross-orientation suppression	Less cross-orientation suppression	More cross-orientation suppression
Contrast response linearity	Relatively linear response to contrast	Nonlinear response to contrast
Selectivity of neuronal orientation tuning	Broader orientation tuning	Narrower orientation tuning
Implication for speed and contrast encoding	Better for representing high speed high contrast stimuli	Better for representing low speed low contrast stimuli

Table 2. Functional features comparison between monocular and binocular zones of V1.

that contribute separately to global motion perception. We speculate that binocular zones in V1 take longer to fully develop, or that their outputs require a longer period of development to be utilized efficiently by MT neurons, as compared to monocular zones that respond to larger displacements.

Putting all of these facts together, we see that pinwheel neurons show ocular dominance, narrow tuning for direction, increasing response as contrast increases, and preference for high speed. MT cells that are not binocularly balanced will inherit from these V1 neurons the ability to discriminate upwards from downwards motion with special sensitivity, provided that the signal and noise dots are segregated to different eyes, and provided the dots have high contrast and high speed. On the other hand, non-pinwheel neurons show binocularity, broad tuning for direction, a response that saturates at low contrast, and a preference for low speed. MT cells, whether binocularly balanced or not, must use as their input the output from these neurons when stimuli are slow or have low contrast, and because they are binocular, there can be no exclusion of noise by MT neurons on the basis of segregation.

### Segregation benefit and monocular gain control

We have developed an explanation for the effect of segregation that appeals to ocular imbalance at the site of combination by regional motion detectors. That explanation is in agreement with known neurophysiology. However, from Experiment 2 we were forced to conclude that some part of the benefit of segregation—perhaps a large part of it—resulted from making signal dots more effective by presenting fewer total dots to the eye that sees signal. If monocular gain control reduces the effectiveness of any dot, whether signal or noise, in the dot-majority eye, then relative

performance would be ordered as we have seen:  $S > SSB = M = B > SSI$ .

Is this monocular-gain-control explanation of the segregation effect consistent with the observation that  $S = M = B$  when dots have low contrast or low speed? It seems possible in principle that a gain control mechanism would exhibit an accelerating nonlinearity as a function of motion energy, which increases with dot number, dot contrast, and dot speed. In that case there could indeed be no penalty for dots that share an eye with many other dots, at low contrast or low speed. Only when total motion-energy was high enough would the gain control engage to reduce the dots' effectiveness as compared to those in the eye with fewer dots. We have not added additional dots to just one eye, so we have limited ability to make inferences about the effects of per-eye dot numbers per se.

Given that we used light dots on a black background, luminance could also have had a monocular gain control effect. Shapley and Enroth-Cugell (1984) suggested that perceptual sensitivity is reduced when luminance gain goes down as a consequence of an increase in background luminance, and Ding and Levi (2015) included a monocular luminance gain control stage in their model. In our condition S, having more dots in the noise eye caused luminance to be greater in that eye as compared to the signal eye, which might have given some benefit to signal dots in later stages after binocular combination.

These gain-control explanations are a bit paradoxical: One might predict that the eye with more dots, and thus greater contrast and motion energy, would inhibit the eye with fewer dots. Interocular suppression and rivalry both work this way, opposite to what we observed. Evidently, for the brief displays used in our experiment, monocular gain control played a greater role than did rivalry, or interocular suppression of the lower dot number eye by the higher dot number eye.

## Further comments on binocular combination

### *No benefit of binocular viewing*

Hess et al. (2007) found that percent coherence thresholds were lower for binocular viewing than for monocular viewing, whereas we found them similar. However, their binocular condition differed from ours because it had very low contrast (0.02 Michelson contrast). Under those conditions, internal noise would have limited the process of measuring dot motions, and more so in their monocular condition. Presenting the dots binocularly did not change the signal-to-noise ratio of the stimulus itself, but binocular contrast summation would have caused each dot's motion to be measured more precisely. Thus, internal noise from sensory encoding would have caused a difference between their monocular and binocular conditions that would not be predicted for our higher contrast stimuli, for which performance was limited in all viewing conditions by the signal-to-noise ratio for percent coherence (Experiments 1 and 3) or dot contrast (Experiment 2).

### *Binocular rivalry*

When a visual stimulus supports the perception of two surfaces in the same location, perceptual rivalry between the surfaces can occur. One might ask whether the benefit of segregation resulted from this sort of rivalry. This explanation is unlikely. For two surfaces to rival, the visual system must be able to construct distinct representations of the two surfaces, respectively. However, in our stimuli there was no common property available for the binding of dots in one eye to each other, except for their having been presented to the same eye. Furthermore, binocular rivalry depends on attention to stimulus features (Mitchell, Stoner, & Reynolds, 2004; H.-H. Li, Rankin, Rinzel, Carrasco, & Heeger, 2017). Therefore, our random dot stimuli were designed to minimize rivalry: The duration of the stimulus was short (300 ms), and we used two-frame motion. Short-duration stimuli encourage fusion and discourage rivalry (Wolfe, 1983; Paffen, Alais, & Verstraten, 2006; Robertson, Kravits, Freyberg, Baron-Cohen, & Baker, 2013). Second, dot density was low, so a dot seen by one eye generally had spatial overlap with blank display in the other eye. Finally, observers reported being able to simultaneously see dots presented to both eyes, and they did not report being able to see the signal dots in percepts with relatively fewer dots, as would have been the case for rivalry between the low-density, signal-dot eye and high-density, noise-dot eye.

### *Binocular combination in static stimuli versus RDK*

For the perception of static patterns, the merging of signals, as contrast is varied separately in the two eyes,

can be described by the Ding-Sperling-Klein-Levi model (the DSKL model), in which interocular contrast gain-control plays an important role (Ding & Sperling, 2006; Ding et al., 2013). These models suggest that within a local region of the cyclopean visual field, each eye exerts gain control on the other eye's signal in proportion to the contrast energy of its own input, and in addition exerts gain control on the other eye's gain control. Both of these mechanisms act so that high contrast in one eye reduces contrast gain for the other eye. A similar mechanism might, in principle, operate during binocular local motion extraction (Figure 1, “nonpinwheels” box) when extracting global motion, but the binocular combination of motion signals during global motion perception need not use the same mechanisms as binocular combination during static pattern perception. Thus, the extent to which binocular combination of static stimuli can inform our understanding of binocular combination during global motion perception seems limited.

## Conclusion

Global motion perception is an essential function of midlevel vision. It has recently been exploited to measure interocular balance in amblyopia, using stimuli that put noise and signal dots into different eyes. We asked whether segregating signal and noise dots to different eyes has an effect on performance, *per se*. In normally sighted observers we found that segregation does matter, but not for all stimuli, and that known physiology can explain when the segregation benefit occurs. There is a clear benefit of segregating signal and noise, but only for high contrast, high-speed dots, and this interaction can be explained by known properties of V1 functional segregation into pinwheel and nonpinwheel regions, together with imbalance in the inputs to regional motion detectors. The explanation is consistent with previous reports that high contrast is beneficial only at higher dot speeds (Seitz et al., 2008). Monocular gain control based on the numbers of dots in each eye accounted for some of the benefit of segregation. Our experimental results are consistent with a hybrid model we proposed (Figure 1) that included both monocular gain control and putative imbalance in the inputs of the regional motion detectors. We suggest that some local motions (high contrast, high speed) are transmitted to MT predominantly before binocular combination while others (low contrast, low speed) are transmitted predominantly after combination, in agreement with results from single unit recording.

*Keywords:* binocular combination, global motion, random-dot kinematogram, dichoptic

## Acknowledgments

We thank Erin Koch for comments on Table. 2. This study was funded by National Institutes of Health grants R01-EY013988 to BTB and T35-EY020481 to SUNY College of Optometry.

Commercial relationships: none.

Corresponding authors: Benjamin T. Backus; Lanya T. Cai.

Email: bbackus@sunyopt.edu; lanyatcai@berkeley.edu.

Address: Graduate Center for Vision Research, SUNY College of Optometry, New York, NY, USA.

## Reference

- Adelson, E. H., & Movshon, J. A. (1982, December 9). Phenomenal coherence of moving visual patterns. *Nature*, *300*(5892), 523–525.
- Albright, T. D. (1984). Direction and orientation selectivity of neurons in visual area MT of the macaque. *Journal of Neurophysiology*, *52*(6), 1106–1130.
- Baker, D. H., Meese, T. S., Mansouri, B., & Hess, R. F. (2007). Binocular summation of contrast remains intact in strabismic amblyopia. *Investigative Ophthalmology and Visual Science*, *48*(11), 5332–5338.
- Barbeito, R., Levi, D., Klein, S., Loshin, D., & Ono, H. (1985). Stereo-deficients and stereoblinds cannot make utrocular discriminations. *Vision Research*, *25*(9), 1345–1348.
- Bonhoeffer, T., & Grinvald, A. (1991, October 3). Iso-orientation domains in cat visual cortex are arranged in pinwheel-like patterns. *Nature*, *353*(6343), 429–431.
- Born, R. T., & Bradley, D. C. (2005). Structure and function of visual area MT. *Annual Review of Neuroscience*, *28*, 157–189.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.
- Burkhalter, A., & Van Essen, D. C. (1986). Processing of color, form and disparity information in visual areas VP and V2 of ventral extrastriate cortex in the macaque monkey. *Journal of Neuroscience*, *6*, 2327–2351.
- Carney, T., & Shadlen, M. N. (1992). Binocularity of early motion mechanisms: Comments on Georgeson and Shackleton. *Vision Research*, *32*, 187–191.
- Carney, T., & Shadlen, M. N. (1993). Dichoptic activation of the early motion system. *Vision Research*, *33*(14), 1977–1995.
- DeAngelis, G. C., Cumming, B. G., & Newsome, W. T. (1998, August 13). Cortical area MT and the perception of stereoscopic depth. *Nature*, *394*, 677–680.
- Ding, J., Klein, S. A., & Levi, D. M. (2013). Binocular combination of phase and contrast explained by a gain-control and gain-enhancement model. *Journal of Vision*, *13*(2):13, 1–37, <https://doi.org/10.1167/13.2.13>. [PubMed] [Article]
- Ding, J., & Levi, D. M. (2015). Interocular contrast gain control plus monocular luminance gain control can explain binocular luminance summation. *Journal of Vision*, *15*(12):263, <https://doi.org/10.1167/15.12.263>. [Abstract]
- Ding, J., & Levi, D. M. (2016). Binocular contrast discrimination needs monocular multiplicative noise. *Journal of Vision*, *16*(5):12, 1–21, <https://doi.org/10.1167/16.5.12>. [PubMed] [Article]
- Ding, J., & Levi, D. M. (2017). Binocular combination of luminance profiles. *Journal of Vision*, *17*(13):4, 1–32, <https://doi.org/10.1167/17.13.4>. [PubMed] [Article]
- Ding, J., & Sperling, G. (2006). A gain-control theory of binocular combination. *Proceedings of the National Academy of Sciences, USA*, *103*, 1141–1146.
- El-Shamayleh, Y., Kiorpes, L., Kohn, A., & Movshon, J. A. (2010). Visual motion processing by neurons in area MT of macaque monkeys with experimental amblyopia. *Journal of Neuroscience*, *30*(36), 12198–12209.
- Enoch, J., Goldmann, H., & Sunga, R. (1969). The ability to distinguish which eye was stimulated by light. *Investigative Ophthalmology & Visual Science*, *8*(3), 317–331.
- Felleman, D. J., & Van Essen, D. C. (1987). Receptive field properties of neurons in area V3 of macaque monkey extrastriate cortex. *Journal of Neurophysiology*, *57*, 889–920.
- Georgeson, M. A., Wallis, S. A., Meese, T. S., & Baker, D. H. (2016). Contrast and lustre: A model that accounts for eleven different forms of contrast discrimination in binocular vision. *Vision Research*, *129*, 98–118.
- Gilaie-Dotan, S. (2016). Visual motion serves but is not under the purview of the dorsal pathway. *Neuropsychologia*, *89*, 378–392.

- Gur, M., & Snodderly, D. M. (2007). Direction selectivity in V1 of alert monkeys: Evidence for parallel pathways for motion processing. *Journal of Physiology*, 585(2), 383–400.
- Hayashi, R., Miura, K., Tabata, H., & Kawano, K. (2008). Eye movements in response to dichoptic motion: Evidence for a parallel-hierarchical structure of visual motion processing in primates. *Journal of Neurophysiology*, 95(5), 2329–2346.
- Hayashi, R., Nishida, S., Toliás, A., & Logothetis, N. K. (2007). A method for generating a “purely first-order” dichoptic motion stimulus. *Journal of Vision*, 7(8):7, 1–10, <https://doi.org/10.1167/7.8.7>. [PubMed] [Article]
- Hess, R. F., Hutchinson, C. V., Ledgeway, T., & Mansouri, B. (2007). Binocular influences on global motion processing in the human visual system. *Vision Research*, 47, 1682–1692.
- Hess, R. F., Mansouri, B., & Thompson, B. (2010). A binocular approach to treating amblyopia: Anti-suppression therapy. *Optometry and Vision Science*, 87(9), 697–704.
- Huang, C., Zhou, J., Lu, Z. L., & Zhou, Y. (2011). Deficient binocular combination reveals mechanisms of anisotropic amblyopia: Signal attenuation and interocular inhibition. *Journal of Vision*, 11(6), 1–17. <https://doi.org/10.1167/11.6.4>. [PubMed] [Article]
- Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, 195, 215–243.
- Hubel, D. H., & Wiesel, T. N. (1970, January 3). Stereoscopic vision in macaque monkey cells sensitive to binocular depth in area 18 of the macaque monkey cortex. *Nature*, 225, 41–42.
- Kiorpes, L., Walton, P. J., O’Keefe, L. P., Movshon, J. A., & Lisberger, S. G. (1996). Effects of early-onset artificial strabismus on pursuit eye movements and on neuronal responses in area MT of macaque monkeys. *The Journal of Neuroscience*, 16(20), 6537–6553.
- Kleiner, M., Brainard, D. H., & Pelli, D. G. (2007). What’s new in psychtoolbox-3? *Perception ECVF Abstract Supplement*, 36, 437–442.
- Koch, E., Jin, J., Alonso, J. M., & Zaidi, Q. (2016). Functional implications of orientation maps in primary visual cortex. *Nature Communications*, 7(13529), 1–13.
- Legge, G. E. (1979). Spatial frequency masking in human vision: Binocular interactions. *JOSA*, 69(6), 838–847.
- Li, H.-H., Rankin, J., Rinzel, J., Carrasco, M., & Heeger, D. J. (2017). Attention model of binocular rivalry. *Proceedings of the National Academy of Sciences, USA*, 114(30), E6192–E6201.
- Li, J., Hess, R. F., Chan, L. Y., Deng, D., Yang, X., Chen, X., . . . Thompson, B. (2013). Quantitative measurement of interocular suppression in anisotropic amblyopia: A case-control study. *Ophthalmology*, 120, 1672–1680.
- Li, J., Thompson, B., Lam, C. S., Deng, D., Chan, L. Y., Maehara, G., . . . Hess, R. F. (2011). The role of suppression in amblyopia. *Investigative Ophthalmology and Visual Science*, 52(7), 4169–4176.
- Maunsell, J. H., & Van Essen, D. C. (1983a). The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *Journal of Neuroscience*, 3(12), 2563–2586.
- Maunsell, J. H., & Van Essen, D. C. (1983b). Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. *Journal of Neurophysiology*, 49(5), 1127–1147.
- Maunsell, J. H., & Van Essen, D. C. (1983c). 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(5), 1148–1167.
- Meese, T. S., Georgeson, M. A., & Baker, D. H. (2006). Binocular contrast vision at and above threshold. *Journal of Vision*, 6(11):7, 1224–1243, <https://doi.org/10.1167/6.11.7>. [PubMed] [Article]
- Meier, K., & Giaschi, D. (2014). The maturation of global motion perception depends on the spatial and temporal offsets of the stimulus. *Vision Research*, 95, 61–67.
- Meier, K., & Giaschi, D. (2017). Effect of spatial and temporal stimulus parameters on the maturation of global motion perception. *Vision Research*, 135, 1–9.
- Mitchell, J. F., Stoner, G. R., & Reynolds, J. H. (2004, May 27). Object-based attention determines dominance in binocular rivalry. *Nature*, 429(6990), 410–413.
- Movshon, J. A., & Newsome, W. T. (1996). Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. *Journal of Neuroscience*, 16(23), 7733–7741.
- Nauhaus, I., Benucci, A., Carandini, M., & Ringach, D. L. (2008). Neuronal selectivity and local map structure in visual cortex. *Neuron*, 57(5), 673–679.

- Nauhaus, I., & Nielsen, K. J. (2014). Building maps from maps in primary visual cortex. *Current Opinion in Neurobiology*, *24*, 1–6.
- Nauhaus, I., Nielsen, K. J., & Callaway, E. M. (2016). Efficient receptive field tiling in primate V1. *Neuron*, *91*, 893–904.
- Paffen, C. L. E., Alais, D., & Verstraten, F. A. J. (2006). Attention speeds binocular rivalry. *Psychological Science*, *17*(9), 752–756.
- Pelli, D. G. (1997). The videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.
- Poggio, G. F., & Fischer, B. (1977). Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey. *Journal of Neurophysiology*, *40*, 1392–1405.
- Poggio, G. F., Gonzalez, F., & Krause, F. (1988). Stereoscopic mechanisms in monkey visual cortex: Binocular correlation and disparity selectivity. *Journal of Neuroscience*, *8*, 4531–4550.
- Robertson, C. E., Kravits, D. J., Freyberg, J., Baron-Cohen, S., & Baker, C. I. (2013). Slower rate of binocular rivalry in autism. *The Journal of Neuroscience*, *33*(43), 16983–16991.
- Roy, J. P., Komatsu, H., & Wurtz, R. H. (1992). Disparity sensitivity of neurons in monkey extrastriate area MST. *Journal of Neuroscience*, *12*, 2478–2492.
- Sceniak, M. P., Hawken, M. J., & Shapley, R. (2002). Contrast-dependent changes in spatial frequency tuning of macaque V1 neurons: Effects of a changing receptive field size. *Journal of Neurophysiology*, *88*, 1363–1373.
- Sceniak, M. P., Ringach, D. L., Hawken, M. J., & Shapley, R. (1999). Contrast's effect on spatial summation by macaque V1 neurons. *Nature Neuroscience*, *2*(8), 733–739.
- Seitz, A. R., Pilly, P. K., & Pack, C. C. (2008). Interactions between contrast and spatial displacement in visual motion processing. *Current Biology*, *18*, 904–906.
- Shadlen, M. N., & Carney, T. (1986, April 4). Mechanisms of human motion perception revealed by a new cyclopean illusion. *Science*, *232*(4746), 95–97.
- Shapley, R., & Enroth-Cugell, C. (1984). Visual adaptation and retinal gain controls. *Progress in Retinal Research*, *3*, 263–343.
- Tadin, D., Lappin, J. S., Gilroy, L. A., & Blake, R. (2003, July 17). Perceptual consequences of centre-surround antagonism in visual motion processing. *Nature*, *424*(6946), 312–315.
- Thompson, B., & Liu, Z. (2006). Learning motion discrimination with suppressed and unsuppressed MT. *Vision Research*, *46*(13), 2110–2121.
- Truchard, A. M., Ohzawa, I., & Freeman, R. D. (2000). Contrast gain control in the visual cortex: Monocular versus binocular mechanisms. *Journal of Neuroscience*, *20*(8), 3017–3032.
- Ungerleider, L. G., & Desimone, R. (1986). Cortical connections of visual area MT in the macaque. *Journal of Comparative Neurology*, *248*(2), 190–222.
- Van Essen, D. C., & Gallant, J. L. (1994). Neural mechanisms of form and motion processing in the primate visual system. *Neuron*, *13*, 1–10.
- Wolfe, J. M. (1983). Influence of spatial frequency, luminance, and duration on binocular rivalry and abnormal fusion of briefly presented dichoptic stimuli. *Perception*, *12*(4), 447–456.
- Yehezkel, O., Ding, J., Sterkin, A., Polat, U., & Levi, D. M. (2016). Binocular combination of stimulus orientation. *Open Science*, *3*(11):e160534.