Speed and accuracy of visual motion discrimination depend systematically on motion strength. This behavior is traditionally explained by diffusion models that assume accumulation of sensory evidence over time to a decision bound. However, how does the brain decide when sensory evidence is ambiguous, such as in binocular rivalry? Theories on bistable vision propose that such a conflict is resolved through competitive interactions between adapting units encoding the alternative stimulus interpretations. Thus, distinctly different theoretical frameworks have been proposed for deciding under uncertainty and ambiguity; a discrepancy overlooked so far. Here, we studied motion discrimination at stimulus onset under both conditions. In Experiment 1, speed and accuracy were similar when observers viewed noisy, unambiguous motion patterns in which signal dots were either at identical or at different, uncorrelated locations for the two eyes. This result is compatible with a race between two monocular discrimination processes. However, Experiments 2 and 3 showed that reaction times increase under rivalry conditions and that this increase cannot be explained by motion transparency. The data thus reveal competitive rivalry interactions. We discuss a model that can account for the accuracy and latencies observed under both ambiguous and unambiguous conditions, by combining key elements from diffusion and rivalry models.

Keywords: visual motion discrimination, perceptual decision making, binocular rivalry, neural models


Introduction

The process of decision making has often been studied using simple two-choice visual motion discrimination tasks in which subjects indicate the perceived direction of coherent motion in a noisy random dot pattern (Gold & Shadlen, 2007; Newsome, Britten, & Movshon, 1989; Schall, 2001; Shadlen & Newsome, 1996). Both the accuracy and the reaction times of subjects’ motion judgments depend systematically on stimulus strength, manipulated by changing the fraction of coherently moving dots (Palmer, Huk, & Shadlen, 2005). The resulting psychometric and chronometric functions can be understood from diffusion models that assume additive accumulation of sensory evidence over time to a decision bound (Ditterich, 2006; Palmer et al., 2005; Smith & Ratcliff, 2004). Typically, these experiments are done with unambiguous stimuli, that is, both eyes are watching the same motion pattern and there is only one correct response. However, how does the brain decide when the sensory evidence is ambiguous such as in binocular rivalry?

Binocular rivalry arises when the two eyes are presented with random dots moving in opposite directions (Blake, Zimba, & Williams, 1985; Moutoussis, Keliris, Kourtzi, & Logothetis, 2005; van de Grind, van Hof, van der Smagt, & Verstraten, 2001). Under such conditions, the brain not only needs to resolve the direction of motion, but it also needs to decide which of the competing percepts will reach awareness first. Previous studies have shown that stimulus timing and percept history both influence the choice process at rivalry onset (Klink et al., 2008; Leopold, Wilke, Maier, & Logothetis, 2002; Pearson & Clifford, 2004). The underlying mechanisms and interactions, however, remain poorly understood.

The diffusion model, although accurately describing the outcomes of the decision process in unambiguous motion discrimination tasks, may not readily describe the percept choice probability at the onset of a bistable stimulus. One reason for this is that the diffusion model does not account for stimulus history. More importantly, the diffusion model does not accommodate any state changes in perception. As a result, it cannot, in its present form, account for stochastic percept alternations that occur during prolonged stimulus...
viewing. On the other hand, recent findings suggest that rivalry at the beginning of a trial, so-called onset rivalry, is different from sustained rivalry (Carter & Cavanagh, 2007; Mamassian & Goutcher, 2005). During sustained rivalry, dominance is much more balanced between the two alternative interpretations than at stimulus onset, where large location and subject-specific biases have been found.

Models describing percept alternations in binocular rivalry differ distinctly from the diffusion model. They typically feature cross-inhibition between two neuronal pools coding for the two different stimuli, in combination with an adaptation process that causes the percept to switch after a certain period. Noest, van Ee, Nijs, and van Wezel (2007) proposed a model that can explain history-dependent percept choices at the onset of bistable stimuli as well as percept switches during sustained viewing. This model is based on competition between cell populations that encode the rivaling percepts, a near-threshold interaction between local adaptation mechanisms and a small neuronal bias. It can successfully describe choice probabilities, but it is not clear whether it can also describe choice latencies. In addition, the Noest model, as well as many other rivalry models (Freeman, 2005; Lankheet, 2006; Wilson, 2003) cannot readily cope with unambiguous stimuli to describe the behavior in an unambiguous task.

Thus, it appears that there exist a number of fundamental problems with both diffusion models and rivalry models that have been overlooked so far. In the present study, we therefore investigated response latencies and choice probabilities at stimulus onset under ambiguous and unambiguous conditions.

In a first series of experiments, we showed subjects dynamic random dot stimuli in which the direction and coherence level of the visual motion was the same in the two eyes. Signal dots, however, were presented either at identical or at different, uncorrelated locations on the retina.

By combining the images from the two eyes, the visual system can, in principle, extract more information about the direction of motion when all signal dots appear at different locations in the two eyes. We found, however, that subjects do not use this possible advantage, which is consistent with the idea that sensory evidence from the two eyes may be integrated separately, and that there might be a race between the two resulting signals to reach the decision threshold first.

In a second series of experiments, we compared the 0% correspondence condition from the first experiment with a bistable condition in which the direction of coherent motion was opposite in the two eyes. We found that subjects reacted significantly slower in the rivalry condition, which indicates that the choice process cannot be understood from an independent race between two monocular discrimination processes. Instead, competitive interactions are required.

In a final experiment, we repeated the second experiment using motion signals that were spatially balanced in the two eyes. By using this stimulus, we could exclude the possibility that the increase in reaction time found in Experiment 2 was confounded by motion transparency.

We propose a model, based on the rivalry model by Noest et al. (2007), that can account for percept choices and reaction times of human observers under both ambiguous and unambiguous conditions.

### Methods

#### Subjects and setup

Nine adult human subjects participated in the experiments after giving informed consent. All subjects had normal or corrected-to-normal visual acuity. Subjects S2, S3, and S5 were authors; the other subjects were kept naive of the purpose of the experiments.

The subjects were seated in a darkened room in front of a screen on which stimuli were back-projected (Panasonic PT-AE700E LCD projector). A chin support minimized head movements. Dichoptic stimuli were presented through red/green anaglyph glasses. The intensities of red and green were adjusted before each experiment until the subject judged a red and a green probe stimulus (3 × 1.5°, viewed through the anaglyph glasses) equiluminant.

#### Visual stimulation

The dichoptic stimuli consisted of two superimposed red/green dynamic random dot patterns (frame rate 60 Hz) that were generated on a personal computer using OpenGL rendering software. The dots were red/green squares of approximately 4 × 4 min of arc (3 × 3 pixels) that appeared against a black background. Each frame of the animation typically displayed 400 red and 400 green dots within a 3°-diameter circular area. A small fixation cross always appeared at the center of this circular region.

Motion stimuli were presented to the two eyes for a fixed period of 2 s. Subjects were instructed to fixate the central fixation cross and to judge the direction of coherent motion as fast and accurately as possible by pressing one of two buttons in a two-alternative forced choice (2AFC) task.

We tested motion stimuli that contained either zero-order or first-order flow components: translating and spiraling dot patterns. Both types of stimuli consisted of randomly placed signal dots and noise dots, having limited, asynchronous life cycles. In the translating stimulus, signal dots were moved horizontally to the left or to the right with a velocity of 7.4°/s. In the spiraling stimulus, expansion/contraction at a rate of 2.4 s⁻¹ and rotation with a frequency of 0.38 cycles/s was combined to yield either clockwise-inward motion or counterclockwise-outward motion. Noise dots always appeared at a random location...
in the image and were replaced at a new, random location after their lifetime of three frames (50 ms) expired.

Stimulus strength was manipulated by changing the proportion of signal dots. Positive coherence levels indicate rightward translation or clockwise-inward spirals, whereas negative coherence levels indicate leftward translation or counterclockwise-outward spirals. Spiraling and translating stimuli were presented in separate sessions.

**Experiment 1: 0% versus 100% correspondence**

**Paradigm**

In the first experiment (subjects S1, S2, S3, and S4), the dots presented to the left eye and the dots presented to the right eye moved in the same direction. Coherence levels in the monocular images ranged between 0% and 60%. Signal dots, however, were presented with either 0% or 100% retinal correspondence between the two eyes. In the 100% correspondence condition, the images presented to the right and the left eye were identical, so that all dots fell on the same locations on the retinas (Figures 1A and 1C). In the 0% correspondence condition, signal dots in the left eye fell at different, uncorrelated retinal locations as compared to those in the right eye (Figure 1B). Thus, images in the left and right eyes were different, but there was no motion conflict, because the signal dots moved in the same direction. By combining the images from the two eyes, there is, in principle, more information about the direction of movement in the 0% correspondence condition than in the 100% correspondence condition. If the brain uses this extra information, performance should be better in the 0% correspondence condition than in the 100% correspondence condition. Alternatively, if the brain only uses the information present in two monocular images without extracting the extra information provided in the combined images, responses in the two conditions should be similar.

The coherence levels of the stimuli with 100% correspondence were chosen in such a way that they either matched those of the monocular view (Figures 1A and 1B)

![Figure 1](http://arvojournals.org/)
or those of the combined view (Figures 1B and 1C) in 0% correspondence trials. In the latter case, twice as many signal dots were necessary, which was achieved by increasing the total number of dots. Noise dots were always presented with 100% retinal correspondence. The stimuli were arranged in blocks of 144 trials in which each combination of coherence level, motion direction, and correspondence condition was presented twice. Subjects completed 6–12 of these randomized blocks, resulting in 12–24 trials per test condition.

In an alternative version of the experiment (with subjects S2 and S4, and two new subjects S5 and S6), we tested the effect of 0% correspondence for the noise dots, keeping all other parameters the same. In the 0% correspondence condition of this experiment, the coherence level of the combined images from the two eyes does not change compared with the monocular view. Instead, combining the images from the two eyes would decrease the amount of information in the 100% signal correspondence condition.

Psychometric and chronometric functions

For each subject, choice probabilities and reaction times obtained under the 0% and 100% correspondence conditions were compared using a χ² test and a two-way ANOVA, respectively. Independent variables in the ANOVA analysis were coherence level and correspondence condition.

In addition, we analyzed the response patterns with a quantitative model description of the psychometric and chronometric response functions derived from a diffusion model (Palmer et al., 2005). This diffusion model describes a decision between two mutually exclusive choices by assuming the accumulation of small amounts of evidence over time, until one of two boundaries is reached, triggering one of the possible decisions. Accumulation depends on the coherence level (C), with a constant scaling factor k. Coherence level is a signed measure: positive coherence levels indicate motion to the right; negative levels indicate motion to the left. C = 0 is the condition with pure noise; A is the level of the boundary. The proportion of right decisions (P(R)) as function of coherence level (the psychometric function) can then be described with

\[ P_R(C) = \frac{1}{1 + e^{-2AkC}}. \]  

The chronometric function (the reaction time \( t_r \) as function of stimulus strength) is assumed to be the sum of two variables: the decision time \( t_D \), depending on the coherence level and a residual time \( t_R \), representing processes independent of the decision process, such as sensory delays and motor execution.

The decision time as function of coherence levels was described using the same parameters (A and k) we used for the psychometric function, and the additional parameter \( t_R \) was added, resulting in the following set of equations for the total reaction time:

\[ t_r(C) = \begin{cases} 
\frac{A}{kC} \tanh(AkC) + t_R & \text{for } C \neq 0 \\
A^2 + t_R & \text{for } C = 0
\end{cases} \]  

The psychometric and chronometric functions were fitted simultaneously, using a maximum likelihood method. Likelihoods for the observed proportion were estimated from a binomial probability density function. The likelihoods of the observed mean response times were approximated with a Gaussian, using measured standard deviations as an estimation of the variability.

For each subject and each correspondence condition, we thus obtained 3 fit parameters: the three parameters A, k, and \( t_R \). Standard errors of these parameters were obtained by bootstrapping. Mean values were compared using Student’s t-tests.

Results

By presenting signal dots with 0% and 100% retinal correspondence, we tested if the brain can extract more information about the direction of motion when the signal dots appear at different locations in the two eyes. Figure 2 shows the results from all subjects obtained with the translating stimuli. Data for the spiraling stimuli, for which we obtained very similar results, can be found in Supplementary Figure 1. Top panels in Figure 2 show the proportion of trials in which the subject perceived rightward motion and bottom panels show the corresponding mean reaction times, both as function of coherence level. Filled symbols represent the data. Solid lines show model fits (Methods section).

The two correspondence conditions were compared in two different ways. First, 0% and 100% correspondence trials were matched for monocular coherence levels (see Methods section, Figures 1A and 1B). When the responses are based purely on monocular information from one eye, or if the brain combines the monocular information from the two eyes before fusion, the data from the 0% and 100% conditions would overlap in this matching procedure. Alternatively, when the information from the combined image is used, responses to the 0% correspondence condition should be better and/or faster. Results for this analysis procedure are shown in Figure 2A. Second, we compared the two correspondence conditions by matching trials for the coherence level in the combined image (Figures 1B and 1C). Overlap of the two response curves in this binocular matching procedure would indicate that the decision is based on information from the combined image. The results for this procedure are shown in Figure 2B.

It can be observed in Figure 2A that the curves from the 0% and 100% correspondence conditions were almost
Figure 2. Psychometric and chronometric response functions obtained for translating stimuli. Positive coherence corresponds with motion to the right, negative coherence with motion to the left. (A) Responses are matched according to monocular coherence level. (B) Responses are matched according to combined coherence level (cf., Figure 1). Each column shows data from one subject. (Top) Percentage rightward motion percept. (Bottom) Mean reaction time. Black: 0% retinal correspondence between images in the two eyes. Gray: 100% retinal correspondence between the images in the two eyes. Solid lines show model fits (Methods section). Note that the data for the 0% correspondence condition are the same between (A) and (B), but that these data are shown as function of monocular coherence in (A) and as function of combined coherence in (B). See Supplementary Figure 1 for the results from the spiraling stimulus.
identical when the responses were matched for monocular coherence levels ($\chi^2$ tests on choice data, $p > 0.7$, ANOVAs on reaction times, $p > 0.1$). The data thus demonstrate that the subjects’ performance did not benefit from the extra information that was available in the combined image in the 0% correspondence trials. Accordingly, matching for coherence level in the combined image (Figure 2B) resulted in a significant difference between the 0% and 100% correspondence data; although the information in the combined image is the same, the curves do not overlap. Instead, the higher number of signal dots in the monocular images in the 100% correspondence condition resulted in a higher accuracy (significant for subjects S2 and S3 in the spiraling conditions, $\chi^2$ tests, $p < 0.05$) and shorter reaction times (ANOVAs, $p < 0.001$ for all four subjects) for both stimulus types.

Table 1 compares the fit parameters of the psychometric and chronometric functions that were obtained under the two conditions for all four subjects. No significant differences ($t$-tests, $p > 0.05$) were found between the two correspondence conditions when the responses were matched according to the monocular coherence level. For the matching according to binocular coherence levels, on the other hand, the diffusion rate ($k$) was found to be significantly higher in the 100% condition for all subjects and for both stimulus types ($t$-tests, $p < 0.05$), suggesting that accumulation of evidence follows the higher, monocular coherence levels.

The combined matching procedure resulted in significantly higher $t_R$ values (representing fixed response delays) for the 100% correspondence condition compared with the 0% correspondence condition. In most cases, however, this was due to a somewhat unrealistically low value of $t_R$ in the 0% condition, even though the fit provides a good description of the response patterns. This further corroborates the notion that the combined matching procedure provides an inadequate characterization of the behavior.

Since we found that subjects do not seem to combine information from the two eyes, at least not beyond the benefit from having two independent monocular views, it is better to look at the parameters found for the monocular matching procedure. The fits obtained with this latter matching procedure indeed resulted in realistic $t_R$ values that were not significantly different between the two correspondence conditions. In addition, no significant differences were found for parameter $A$ for any of the subjects or stimulus types in this analysis.

In an alternative version of this experiment (Figure 3), we tested the effect of presenting noise dots with 0% retinal correspondence in subjects S2, S4, S5, and S6. Consistent with the results presented above, we found that when we matched for monocular coherence level, the choice probabilities did not show any significant differences between the 0% and 100% signal correspondence conditions ($\chi^2$ tests, $p > 0.2$). For the translating stimulus, differences in reaction time were also not statistically significant (ANOVAs, $p > 0.08$), except in subject S6 ($p < 0.05$). For the spiraling stimulus, there were some differences (ANOVAs, $p < 0.05$), but these differences were small, and sometimes opposite. Accordingly, the parameters of the response curves fitted to the data were not significantly different in any of the subjects ($t$-tests, $p > 0.05$).

Conversely, when trials were matched for combined coherence levels, differences in reaction times between the 0% and 100% correspondence conditions were highly

| Stimulus type | Matching procedure | Parameter | Subject 1 | | Subject 2 | | Subject 3 | | Subject 4 |
|--------------|--------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
|              |                    |           | Correspondence | | Correspondence | | Correspondence | | Correspondence |
|              |                    |           | 0% | 100% | 0% | 100% | 0% | 100% | 0% | 100% |
| Translation | Monocular | $A$ | 1.00 | 1.00 | 1.02 | 1.01 | 1.05 | 1.01 | 0.61 | 0.56 |
|              |                    | $k$ | 5.27 | 5.48 | 9.30 | 7.95 | 4.87 | 6.30 | 11.86 | 11.34 |
|              |                    | $t_R$ | 0.35 | 0.38 | 0.40 | 0.35 | 0.31 | 0.41 | 0.60 | 0.61 |
|              | Combined | $A$ | 1.14 | 0.98 | 1.08 | 0.95 | 1.22 | 0.99 | 0.66 | 0.57 |
|              |                    | $k$ | 2.46 | 5.87* | 4.64 | 10.08* | 2.17 | 6.67* | 6.12 | 14.23* |
|              |                    | $t_R$ | 0.07 | 0.42 | 0.27 | 0.44* | -0.06 | 0.45 | 0.54 | 0.61 |
| Spiral      | Monocular | $A$ | 0.99 | 0.99 | 0.97 | 0.98 | 0.97 | 0.91 | 0.81 | 0.83 |
|              |                    | $k$ | 5.42 | 5.19 | 7.50 | 6.55 | 7.73 | 7.96 | 7.92 | 6.17 |
|              |                    | $t_R$ | 0.45 | 0.42 | 0.41 | 0.37 | 0.51 | 0.56 | 0.49 | 0.41 |
|              | Combined | $A$ | 1.12 | 0.96 | 1.06 | 0.88 | 1.06 | 0.89 | 0.88 | 0.73 |
|              |                    | $k$ | 2.57 | 5.69* | 3.74 | 9.99* | 3.76 | 8.98* | 4.05 | 8.99* |
|              |                    | $t_R$ | 0.19 | 0.50 | 0.25 | 0.55* | 0.34 | 0.59* | 0.38 | 0.54 |

Table 1. Fit parameters $A$, $k$, and $t_R$ of the psychometric and chronometric response curves obtained in Experiment 1 for all four subjects for translating (transl.) and spiraling stimuli; 0% correspondence and 100% correspondence conditions were compared in two different ways, indicated here with “monocular” and “combined,” respectively. See text for details on the analysis. Asterisks indicate significant differences between the 0% and 100% correspondence conditions ($t$-test). Note: *$p < 0.05$, **$p < 0.01$. 

Subjects responded faster in the 100% correspondence condition, where the monocular coherence levels were higher. The difference in reaction times between the two conditions could be captured by a significant increase in diffusion rate ($k$) for the 100% correspondence condition, as compared to the 0% correspondence condition. This excludes the possibility that the difference observed in Figure 2B results merely...
from the higher number of dots in the 100% correspondence condition.

Discussion

Taken together, the results from Experiment 1 indicate that (in the absence of depth cues) the brain only uses the information provided by the monocular images when making a categorical decision about the direction of movement. Clearly, this result does not exclude the possibility that the brain still uses the information from both eyes because it is possible that the brain combines the available information before fusion. Whether or not this would lead to enhanced performance under the 0% correspondence condition tested in Experiment 1 depends on how the information is combined. In the context of the diffusion model, there are (at least) two distinctly different possibilities.

First, the brain could pool the sensory evidence from the two eyes to achieve a better signal-to-noise ratio (Figure 4A). Note, however, that the benefit provided by pooling is substantially limited by the amount of covariation in the neuronal responses (Zohary, Shadlen, & Newsome, 1994). Thus, the benefit that would result from pooling is expected to be larger in the 0% correspondence condition compared with in the 100% correspondence condition, because in the former condition signal dots in

A. Pooled sensory evidence from the two eyes

![Diagram A](https://example.com/diagram_a.png)

B. Race between independent decision units for each eye

![Diagram B](https://example.com/diagram_b.png)

Figure 4. Schematic illustration of the two ways in which the information from the right and left eyes can be combined in a diffusion-type model. (A) Sensory evidence from the left and right eyes is pooled and integrated ($\int dt$). The decision is made when the pooled evidence reaches one of the decision bounds. (B) Sensory evidence from the two eyes is integrated separately and the two accumulated evidence signals race for their thresholds. The decision is made when one of the two signals reaches one of the decision bounds (dashed vertical line).
the two eyes appeared asynchronously at different locations. Consequently, this scheme would predict enhanced accuracy in the 0% condition. Our data do not support this interpretation.

Second, the sensory evidence from the two eyes could be integrated separately, and there might be a race between the two resulting signals to reach the decision threshold first (Figure 4B). Such a scheme provides enhanced performance by virtue of statistical facilitation in reaction times. However, in contrast with the pooling scheme, it would predict equal performance under the 0% and 100% correspondence conditions, as was observed in the experiments.

This raises the interesting question of what will happen if the information is conflicting between the eyes. Is it possible to describe motion discrimination under such conditions as a race between two independent decision units, each relying on temporal integration of monocular motion information? If so, we would expect similar decision speeds for ambiguous versus unambiguous conditions. Alternatively, the two motion percepts might compete for dominance through subtractive cross-inhibition (as assumed in rivalry models), which would predict a drop in decision speed. In the second experiment, we therefore compared the responses to unambiguous and ambiguous motion stimuli.

### Experiment 2: Reaction times for ambiguous versus non-ambiguous stimuli

#### Paradigm

In the second experiment (subjects S1, S2, S3, and S4), we compared our subjects’ performance on judging the direction of motion of ambiguous (rivaling) and non-ambiguous stimuli. In both conditions, the monocular images gave the same amount of information.

The two stimulus conditions are illustrated in Figure 5. In the unambiguous condition, the direction of the motion in the images presented to the left and right eyes was the same, but signal dots were presented at different retinal locations in the left and right eye images (this condition was identical to the 0% correspondence condition in Experiment 1). In the ambiguous condition, the dots moved to the left (or counterclockwise-outward) in the left eye and to the right (or clockwise-inward) in the right eye or vice versa, resulting in binocular rivalry. Retinal locations of signal dots in the left and right eyes were again uncorrelated.

Stimuli were arranged in blocks of 96 trials in which rivaling and non-rivaling trials as well as trials with different coherence levels and motion directions were randomly interleaved. Each test condition was presented twice per block. Subjects completed 6–9 blocks, resulting in 12–18 trials per test condition.

Subjects were instructed to indicate their first motion percept as fast and accurately as possible (2AFC task). Observers reported afterward that they rarely saw transitions in motion direction (<3% of the trials), and that those transitions, if they occurred, happened only after the button press.

### Psychometric and chronometric functions

Following the same procedure as described for Experiment 1, choice probabilities and reaction times obtained in ambiguous and unambiguous conditions were compared using $\chi^2$ tests and two-way ANOVAs, respectively.

We also fitted psychometric and chronometric curves to the data from the unambiguous trials using Equations 1 and 2, respectively (see Experiment 1). To fit the chronometric curve from the rivalry condition, the parameters $A$ and $t_R$ were assumed not to be affected by the rivalry condition, so these parameters were kept fixed at the level fitted for the unambiguous condition. For the rivalry condition, we thus only fitted $k$ (Equation 2). Psychometric curves were not calculated for the rivalry condition because there was no correct answer in this condition.

### Results

Figure 6 shows the percentage of rightward motion percepts and the corresponding mean reaction times as a function of coherence level in the translating stimulus.

For the unambiguous condition, positive coherence levels indicate rightward motion and negative coherence levels refer to leftward motion in either eye. For the ambiguous condition, the direction of motion was opposite in the two eyes. The responses for this condition are shown as function of motion coherence in the right eye (as schematically depicted below the top axes in Figure 6). So, for this condition, positive coherence means rightward motion in the right eye (and to the left in the left eye), whereas negative coherence means leftward motion in the right eye (and to the right in the left eye). In this way, any preference for perceiving motion in one eye over the other becomes manifest as a response bias for motion direction.

In the unambiguous condition, subjects correctly indicated the motion direction in trials with high coherence levels, but the percentage of correct responses dropped to chance level when the coherence level decreased to zero. Rightward responses in ambiguous trials did not show such a pattern. Instead, the data scattered roughly around 50%. There were, however, subject-specific biases, suggesting an idiosyncratic eye preference. For example, at high coherence levels, subject 1 tended to respond rightward when the
right eye was watching rightward motion (data in top right corner), but leftward when the right eye was watching leftward motion (data in bottom left corner). This bias, suggesting a preference for the right eye, was strongest in trials with the highest coherence levels.

Reaction time decreased with increasing coherence level in both conditions. Interestingly, however, reaction times were significantly shorter for unambiguous stimuli than for rivaling stimuli for all but the lowest and highest coherence levels (t-tests, \( p < 0.05 \)). This systematic difference in reaction time was present for all subjects. Moreover, the same pattern of results was observed for spiral motion (Supplementary Figure 3).

To further quantify this effect, we fitted chronometric response functions to the data (see Methods section). The resulting fit parameters are listed in Table 2. The difference in reaction times between the two conditions could be captured by a significant decrease in diffusion rate (\( k \)) for the ambiguous condition, as compared to the unambiguous condition.

**Discussion**

Our data thus show that subjects reacted slower on ambiguous trials compared with unambiguous trials. Apparently, the conflicting information present in the suppressed eye in ambiguous trials slows down the decision process. Clearly, this behavior cannot be understood from a simple race between two independent decision units each relying on temporal integration of monocular motion information. Instead, our observations suggest competitive interactions.

Unfortunately, it is not immediately clear from these experiments whether the slower responses are caused by rivalry alone. Motion detection thresholds also drop under opposite transparent motion conditions, i.e., when the two patterns of motion are presented to both eyes simultaneously (Mather & Moulden, 1983; Snowden & Braddick, 1989), so discrimination speed might have dropped because of that. Interestingly, however, the percept of transparent motion only exists if the two patterns of motion contain locally unbalanced motion signals (which is the case, e.g., when two random dot motion patterns are uncorrelated), but it is destroyed if the two opposite components of motion are spatially balanced (Qian, Andersen, & Adelson, 1994). In fact, such stimuli look like flickering noise.

We thus decided to conduct control experiments in which the motion patterns in the left and right eyes were either uncorrelated (as in Experiment 2) or spatially balanced. If the observed increase in reaction times is due to rivalry alone, no difference between these two types of rivaling stimuli is expected. However, if transparency was indeed a major confound, the increase in reaction times should be attenuated considerably for the spatially balanced motion patterns compared with unbalanced motion patterns.

**Experiment 3: Rivalry of paired dot patterns**

As explained above, we could not exclude from Experiment 2 that the drop in discrimination speed under the rivalry condition was partly due to motion transparency.
Figure 6. Psychometric and chronometric response functions for ambiguous and unambiguous translating stimuli. Each column shows results of one subject. (Top) Percentage of rightward response as function of motion coherence in both eyes (unambiguous stimuli) or the right eye (ambiguous stimuli). (Bottom) Mean reaction times in the same trials. Black triangles: Unambiguous stimuli. Gray squares: Ambiguous stimuli. Arrows marked L and R underneath the horizontal axis (top) show the direction of motion in the left and right eyes, respectively, in the ambiguous condition. Solid lines show the model fit. See Supplementary Figure 3 for the results from the spiraling stimulus.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Subject 1</th>
<th>Subject 2</th>
<th>Subject 3</th>
<th>Subject 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unambiguous</td>
<td>Ambiguous</td>
<td>Unambiguous</td>
<td>Ambiguous</td>
<td>Unambiguous</td>
</tr>
<tr>
<td>Translation</td>
<td>0.97</td>
<td>0.97</td>
<td>0.85</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>7.30</td>
<td>4.23**</td>
<td>9.53</td>
<td>4.75**</td>
</tr>
<tr>
<td></td>
<td>0.36</td>
<td>0.36</td>
<td>0.47</td>
<td>0.47</td>
</tr>
<tr>
<td>Spiral</td>
<td>0.99</td>
<td>0.99</td>
<td>0.95</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>5.01</td>
<td>3.17</td>
<td>7.38</td>
<td>4.33*</td>
</tr>
<tr>
<td></td>
<td>0.40</td>
<td>0.40</td>
<td>0.60</td>
<td>0.60</td>
</tr>
</tbody>
</table>

Table 2. Fit parameters $A$, $k$, and $t_R$ from Experiment 2 for all four subjects. For fitting the ambiguous condition, the parameters $A$ and $t_R$ were kept fixed at the level found in the unambiguous condition. Asterisks indicate significant differences between the unambiguous and ambiguous conditions. Note: *$p < 0.05$, **$p < 0.01$. 

Kalisvaart, Klaver, & Goossens 11

Downloaded From: http://arvojournals.org/ on 10/29/2015
Another limitation of the experimental setup in Experiments 1 and 2 was that the anaglyph glasses (Methods section) might have produced an insufficient separation between the left and right eye images. In the third experiment, we therefore used a setup that allowed us to present the stimuli through a front-mirror stereoscope. We repeated the test conditions of Experiments 1 and 2 for translating motion patterns but with slightly different parameters. We also included a new rivalry condition in which motion signals in the two eyes were spatially balanced to control for the possible influence of motion transparency.

Subjects and setup

Four subjects participated in this experiment. Subject S2 also participated in the first two experiments. The three new subjects S7, S8, and S9 were naive with regard to the purpose of the study. The subjects were seated in front of an LCD monitor (Dell 2007WFPb) on which the visual stimuli were displayed at 60 Hz. Subjects watched the screen through a front-mirror stereoscope (HyperView, Berzin, USA) at an effective viewing distance of 67 cm.

Paradigm

Figure 7 illustrates the four stimulus conditions used in Experiment 3. In all four conditions, the motion patterns in each eye consisted of 200 dots (4 × 4 min of arc) in a circular aperture of 3°. Signal dots moved horizontally at a speed of 2.4°/s. The signal dots and noise dots had asynchronous, limited lifetimes of 100 ms (6 frames) and 17 ms (1 frame), respectively.

There were two ambiguous conditions, one with paired dots and one with unpaired dots. In the paired condition, each signal dot in the left eye was paired with a signal dot in the right eye, and the two dots moved in opposite direction without vertical offset, crossing each other at 50% of their lifetime. The unpaired condition was similar to the rivalry condition from Experiment 2, except that the number of dots, the speed of motion, and the lifetimes of the signal and noise dots were different. These latter parameters were modified in such a way that the paired motion patterns, when presented transparently to both eyes simultaneously, were perceived as flickering noise while the unpaired motion patterns produced a clear transparent motion percept (Supplementary Figure 4 shows the results of a separate 2AFC experiment, which confirmed this for all four subjects). There were also two unambiguous conditions with paired and unpaired dots. These conditions were similar to the 100% and 0% correspondence conditions of Experiment 1, respectively, except for density, speed, and lifetimes of the dots.

As in the previous experiments, subjects were instructed to indicate the direction of motion by pressing one of two mouse buttons as fast and accurately as possible (2AFC task). Stimuli were presented until the subject responded, or until the maximum presentation time of 2 s was exceeded. All four test conditions and coherence levels were randomly interleaved and presented in blocks of 160 trials in which each condition was shown twice. Each subject completed 10–15 blocks, resulting in 20–30 trials per condition.

Results

Figure 8 shows the percentage of rightward motion percepts and the corresponding mean reaction times as a
function of coherence level for each of the four test conditions. Data are presented in the same way as in Figure 6. Solid and dashed lines show fit results obtained with the diffusion model, and the corresponding fit parameters are listed in Table 3. Note that the overall pattern of results is very similar to the results from Experiment 2.

In the unambiguous condition, subjects were fast and accurate in discriminating the motion direction in trials with high coherence levels, but the percentage of correct responses dropped to chance level and reactions got slower when the coherence level decreased to zero. No difference was found between the unambiguous paired and unpaired conditions ($\chi^2$ tests on choice probabilities, $p > 0.2$; two-way ANOVAs on reaction times, $p > 0.05$, except for S7), thereby replicating the results from Experiment 1.

In the ambiguous conditions, rightward motion percepts either scattered around 50% (S8) or showed strong individual eye biases (left eye preference in subjects S2 and S9 and right eye preference in subject S7). As in Experiment 2, reaction times were shorter for unambiguous stimuli than for rivaling stimuli for all subjects. This difference was statistically significant in all subjects except S9 (ANOVAs, $p < 0.001$) and also reflected in significant decreases in the diffusion rate ($k$) for the ambiguous conditions as compared to the unambiguous condition (Table 3).

Importantly, the decrease in reaction time under the ambiguous conditions was not significantly different between the paired and the unpaired conditions (ANOVA, $p > 0.05$). If anything, the paired stimuli produced the largest drop in discrimination speed, while the opposite is expected if this decrease were (partly) due to transparency. In addition, the fit parameters (Table 4, $t$-test, $p > 0.05$) obtained for the chronometric response functions were not significantly different (although it should be noted that the fit for the ambiguous conditions did not provide a good

Figure 8. Psychometric and chronometric response functions for ambiguous and unambiguous conditions with paired and unpaired signal dots. Same layout as Figure 6. Black triangles: Unambiguous, paired motion. White triangles: Unambiguous, unpaired motion. Gray squares: Ambiguous, paired motion. White squares: Ambiguous, unpaired motion. Solid and dashed lines show fit results. Note the different scale on the vertical axis for subject S9.
description of the data from subject S2). The results thus indicate that the increases in reaction time are primarily due to rivalry.

### General discussion

In this study, we presented subjects with a motion discrimination task in which the left and right eyes were shown dissimilar stimuli. The 0% correspondence condition that we used in Experiment 1 could be regarded as a condition in between binocular rivalry and unambiguous motion. On the one hand, the images in the left and right eyes differ, while on the other hand, the information in the two images, i.e., the direction of movement, is the same. This way the stimulus, although different for the left and right eyes, does not give rise to a binocular conflict nor to a depth percept. To use all the information from the two eyes when presented with such a stimulus, images should be combined, not by using stereovision but by superimposing the images of the two eyes, or by simply averaging the sensory evidence provided by each eye. None of our subjects were able to use these possible benefits to improve performance, neither for zero-order (translation) nor for first-order (spiral) motion patterns.

Alternatively, subjects could take advantage of the binocular correlation at zero disparity between signal dots to perform better in the 100% correspondence condition. The strong overlap of the data from the two correspondence conditions, however, showed that subjects did not use this possible advantage either. In all subjects, the parameters of the diffusion model (drift rate, k, and bound levels, A), as derived from the observed psychometric and chronometric functions, were identical. These results corroborate and extend the work by Muller, Lankheet, and van de Grind (2004). Using random pixel arrays, they found that binocular correlation in translating stimuli did not significantly improve nor decrease the thresholds for motion detection. In the present study, we confirmed these findings for translating stimuli by showing that binocular correlation of dynamic random dot patterns has no effect on accuracy in a motion discrimination task. In addition, we found that the chronometric functions remained unchanged too. Muller et al. (2004) only studied linear motion. Linear motion is processed at a relatively low level in the visual system (in area MT, Albright, Desimone, & Gross, 1984; Baker, Petersen, Newsome, & Allman, 1981; Maunsell & Van Essen, 1983). We also included more complex spiral motion patterns, which are processed at higher stages of the visual system (in area MST, Duffy & Wurtz, 1991; Graziano, Andersen, & Snowden, 1994; Orban et al., 1992). For spiraling stimuli, however, we found the same results as for simple linear translation: neither speed nor accuracy changed.

The results from our first experiment thus seem to suggest that subjects base their decision on motion direction in noisy random dot patterns purely on the information provided by the monocular images. Clearly, this result does not exclude the possibility that the brain still uses the information from both eyes. A simple race between two independent decision

<table>
<thead>
<tr>
<th>Stimulus type</th>
<th>Parameter</th>
<th>Subject 2</th>
<th>Subject 7</th>
<th>Subject 8</th>
<th>Subject 9</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>Paired</td>
<td>Unpaired</td>
<td>Paired</td>
<td>Unpaired</td>
</tr>
<tr>
<td>Unambiguous</td>
<td>0.76</td>
<td>0.74</td>
<td>1.03</td>
<td>1.03</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>k</td>
<td>33.8</td>
<td>23.1</td>
<td>7.93</td>
<td>4.69</td>
</tr>
<tr>
<td></td>
<td>t_r</td>
<td>0.46</td>
<td>0.44</td>
<td>0.42</td>
<td>0.33</td>
</tr>
<tr>
<td>Ambiguous</td>
<td>0.76</td>
<td>0.74</td>
<td>1.03</td>
<td>1.03</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>k</td>
<td>8.38</td>
<td>6.30</td>
<td>3.22</td>
<td>2.92</td>
</tr>
<tr>
<td></td>
<td>t_r</td>
<td>0.46</td>
<td>0.44</td>
<td>0.42</td>
<td>0.33</td>
</tr>
</tbody>
</table>

Table 4. Fit parameters A, k, and t_r from Experiment 3 for all four subjects, calculated separately for the paired and unpaired conditions. For fitting the ambiguous condition, the parameters A and t_r were kept fixed at the level found in the unambiguous condition. No significant differences between the paired and unpaired conditions were found for any of the parameters in any of the subjects (t-tests, p > 0.05).
units each relying on temporal integration of monocular motion information could account for the psychometric and chronometric curves found in Experiment 1. However, it could not explain the increase in reaction time observed in Experiment 2, where we presented stimuli that moved in opposite directions in the left and right eyes. In Experiment 3, we showed that this decrease in reaction time could not be explained by a transparent motion percept in the ambiguous conditions. The slowing down of the decision process therefore should be accounted for by the subconscious presence of the rivaling stimulus.

Andrews and Blakemore (2002) have reported that motion signals from perceptually suppressed grating indeed continue to influence the perception of motion. Subconscious processing of visual images has also been demonstrated by investigating the build up of afterimages by stimuli that were suppressed part of the time by binocular rivalry. Linear motion patterns and spirals are still effective stimuli that were suppressed part of the time by binocular rivalry. It should be noted, however, that the latency changes we found also depended strongly on stimulus uncertainty (i.e., coherence level). Takei and Nishida (2010) used stimuli with a fixed, very low uncertainty and manipulated the amount of ambiguity using a different visual feature. The rotating cylinder, for example, was defined by motion, but they controlled the amount of ambiguity through binocular disparity. Indeed, when they used a four-dot apparent motion stimulus, in which both the signal strength and the degree of ambiguity were defined by a rotation angle, the increase in reaction time was about two times larger (i.e., nearly 10%) than for the rotating cylinders.

Model simulations

The diffusion-to-bound model (Methods section) could describe reaction times and psychometric curves in situations without binocular motion conflict (Experiment 1) but not in ambiguous, rivalry situations (Experiment 2). In order to explain both the psychometric and chronometric curves under both conditions, a different type of model is required. A parsimonious extension to the diffusion model would be to assume that the motion discrimination process relies on an independent race between two monocular discrimination processes, each relying on temporal integration of monocular motion information (Figure 4B). However, such an idea cannot account for the observed increases in latencies in the rivalry conditions. These latter observations instead suggest competitive interactions.

Several quantitative models have been proposed for describing competitive interactions in binocular rivalry (e.g., Klink et al., 2008; Laing & Chow, 2002; Lankheet, 2006; Lehky, 1988; Noest et al., 2007; Shpiro, Curtu, Rinzel, & Rubin, 2007; Wilson, 2003). These models feature two main characteristics: mutual inhibition between the two eyes to ensure that only one image is visible at the same time and self-adaptation that decreases the dominance of one eye over time so that the other eye can take over. The model by Noest et al. (2007) can describe bistable percept choices at stimulus onset during intermittent stimulus presentation as well as percept switches during sustained stimulus viewing (Klink et al., 2008; Noest et al., 2007). However, its inputs and subsequent rivalry-resolving stage do not incorporate the existence of opponent sensory channels from each eye, which means that it cannot accommodate different stimulus strengths and different motion directions simultaneously. The latter is needed to cope with both rivalrous and non-rivalrous stimuli in different directions. In addition, it lacks a representation of sensory noise. In an attempt to resolve this problem, we extended the rivalry model of Noest et al. to accommodate opponent motion channels in each eye and combined it with elements of the diffusion model for motion discrimination to accommodate sensory noise.

Figure 9 illustrates the modifications we implemented. For each eye, the pooled activity of two populations of motion sensitive neurons with opposite preferred directions was simulated (Figure 9A). For each population, we assumed that the mean response for motion in its preferred direction was larger than for motion in the opposite (null) direction, and that the response variance is proportional to the mean. Such responses to random dot motion have been found, for example, in macaque areas V1 and MT (Snowden, Treue, & Andersen, 1992). We also assumed that the sensory responses depend systematically on motion coherence. As in previous diffusion models, we modeled this relationship according to responses measured in area MT (Britten, Newsome, Shadlen, Celebrini, & Movshon, 1996; Shadlen, Britten, Newsome, & Movshon, 1996; Shadlen, Hanks, Churchland, Kiani, & Yang, 2007). More specifically, the output of each population was described as a constant depending on motion coherence with additive white noise (see Appendix A, for details). The two signals from each eye then undergo a subtraction, which results in two monocular evidence signals. One is signaling evidence for rightward and against leftward motion. The other is signaling evidence for leftward and against rightward motion in the same eye.
These monocular signals are then integrated over time by four non-linear units, which adapt and compete for dominance through subtractive cross-inhibition. As in Noest et al.’s (2007) model, each unit integrates its visual input with a “shunting”-type gain control, which implements the adaptation, and the level of adaptation interacts with a constant neural baseline (see Appendix A for details). We assumed that each unit is only driven by monocular inputs because the results of Experiment 1 suggested that a binocular input stage may not be needed to account for our results. The rivalry stage of the model is therefore limited to two units for each eye that have opposite “on-directions.” Note that, although driven by monocular input, the rivalry solving units in the model are in fact binocular (because of the cross-inhibition). In our scheme, these binocular units represent leftward and rightward directions of motion, but they may of course also represent other, more complex features.

Figure 9. Schematic representation of our percept-choice model for visual motion discrimination under uncertainty and rivalry. The model consisted of an input stage, which calculates the sensory evidence for rightward and against leftward motion, and vice versa, from each of the two eyes, and a rivalry-resolving/decision stage. A choice is made when the difference between activation levels for leftward and rightward motions in this latter stage reaches the upper or lower bound, respectively. (A) The input stage consisted of motion sensitive visual units. Their activity as a function of motion coherence and direction was modeled according to previously measured responses in macaque area MT (Britten et al., 1993; Shadlen et al., 1996). For motion in the preferred direction, the mean activity increased linearly (slope $a_p$) with increasing coherence level with respect to the baseline ($b$) at 0% motion coherence. For motion in the null direction, the mean activity decreased with increasing coherence but less steeply (slope $a_n$). The response variance, which was simulated by adding Gaussian noise, was assumed to be proportional to the mean activation level. (B, C) The rivalry-resolving stage consisted of four nonlinear units that adapt and compete for dominance. Each unit received momentary evidence for rightward and against leftward motion, or vice versa, from one of the two eyes (black) as well as subtractive cross-inhibition (blue). The two units for each eye had opposite on-directions (indicated by the arrows) due to their opponent visual inputs and the threshold imposed by the sigmoid “firing-rate” function (which converts the “local field”, $H$, of each unit into momentary spike rates). In one version of the model (B), the cross-inhibition was derived from the pooled outputs of its opposing units in both eyes. In the second version (C), each unit received only cross-inhibition from the opposing unit in the other eye. In both cases, the cross-inhibition and visual inputs were integrated with a “shunting”-type gain control that implements the adaptation, and the level of adaptation interacted with a constant neural baseline as in Noest et al.’s (2007) model. The dynamics of the model was given by a set of differential equations that specified the dynamics of the “local field” variable $H_{ij}$ and adaptation state $A_{ij}$ of each unit (see Appendix A for details).

In one version of the model (Figure 9B), we assumed that the cross-inhibition signals are binocular signals produced by pooling the output from the left- and right-eye units with the same on-direction, thus supporting competition between percepts. In a second version (Figure 9C), we assumed that the cross-inhibition signals remained monocular, thus supporting competition between eyes. In both cases, a response is made when the mean activity of the units competing for one motion direction exceeds the mean activity of the others by a criterion amount. Parameters of the rivalry stage and the level of the decision bound were adjusted manually to roughly fit the experimental data.

Interestingly, both version of the model produced very similar response patterns and could adequately capture the key changes in percept choices and mean reaction times that we found between the rivalrous and non-rivalrous conditions. As shown in Figure 10, both models produced a sigmoid psychometric curve in the non-rivalry condition (triangles), which adequately describes the increasing accuracy with increasing coherence level (cf., Figure 6). Simulations of the rivalry condition (squares) with the same model parameters produced a flat line at 50% rightward choices. In addition, we obtained realistic bell-shaped pattern of reaction times. Reaction times decreased with increasing coherence level but less steeply in the ambiguous than in the unambiguous condition, which is consistent with the results from Experiment 2. Increasing/decreasing the scaling factor $g$ of the monocular inputs from one or the other eye could readily simulate the subject-specific eye preferences that were observed in previous onset rivalry experiments (Carter & Cavanagh, 2007) and in the rivalry condition of Experiments 2 and 3. Note that the simulated curves are smoother than the curves obtained by the experimental data because the simulations were run for 1000 trials per condition, as compared to ~20 trials per condition in the experimental data.

It is clearly oversimplified to assume that the rivalry-resolving stage only receives purely monocular inputs. For example, the two schemes can, in their present form, not accommodate the results by Meng, Chen, and Qian (2004), who showed that subjects still perceive motion when presented with stimuli that contain binocular motion signals but no monocular motion signals.

We believe, however, that this shortcoming could be “repaired” by assuming that the two rightward and the two leftward motion units in the rivalry-resolving stages are partially driven by inputs from both eyes, because the net function of these units is that they resolve the conflict in motion direction, not eyes.

### Conclusions

Our experiments demonstrate that binocular correlations between unambiguous dichoptic motion stimuli do not facilitate visual motion discrimination. One would be
tempted to conclude therefore that visual motion is primarily detected based on the monocular information from the two eyes. However, conflicting motion information in the two eyes does slow down the motion discrimination process, which means that the process cannot be understood from an independent race between two monocular discrimination processes. Instead, competitive interactions are required. We investigated two parsimonious models that implement such competitions. One assumed competition between the opponent motion channels from the two eyes, and the other assumed competition between the opponent motion percepts. Both models could readily simulate the observed response patterns under non-rivalry and rivalry conditions. The model we propose combines key elements from existing diffusion and rivalry models into one, unified theoretical framework.

Appendix A

The model in Figure 9 was implemented in Simulink (Version 7.4; The MathWorks). It consisted of an input stage, which modeled monocular sensory responses, and a rivalry-resolving/decision stage.

The input stage simulated for each eye the activity of two populations of motion sensitive neurons with opposite preferred directions. As in previous diffusion models, we assumed that each population (1) responds vigorously for motion in the preferred direction and only weakly for motion in the opposite direction, (2) the mean response increases (decreases) systematically as a function of motion coherence in the preferred (null) direction, and (3) the response variance is proportional to the mean. More specifically, for motion in the preferred direction we assumed that the mean response, $R$, of each unit increased linearly (slope $a_p = 0.256$) with increasing coherence level with respect to the baseline ($b = 23.32$) at 0% motion coherence. For motion in the null direction, we assumed that the mean responses decreased with increasing coherence but less steeply (slope $a_n = -0.072$). The response variance within each population was simulated by adding Gaussian noise with mean $\mu = 0$ and variance $\sigma^2dt$ to the mean, where $\sigma^2 = r \cdot \phi_1 R(C)$ with $\phi_1 \approx 1.5$ spk$^2$ as the variability of an individual neuron and $r \approx 0.15$ as the correlation between neurons in the same population. The values for response parameters $a_p$, $a_n$, $b$, $r$, and $\phi_1$ were adopted from Britten, Shadlen, Newsome, and Movson (1993) and Shadlen et al. (1996).

The rivalry-resolving stage consisted of four non-linear units that adapt and compete for dominance. As shown in Figure 10. Simulation results with two versions of the model. Same layout as Figure 6. The simulations consisted of 1000 trials per coherence level per condition with bounds at ±0.15. (Left) Each adapting unit received cross-inhibition from the pooled outputs of its opposing units in both eyes. (Right) Each adaptive unit received only cross-inhibition from the opposing unit in the other eye. Note the similarity between the two types of cross-inhibition.
Figure 9, each unit was driven by monocular inputs and had either a leftward or rightward on-direction as determined by its input connections. The dynamics of each unit in this stage was given by a set of differential equations that specified the “local field” dynamics and the “shunting-type” adaptation component of each unit (after Noest et al., 2007):

\[
\tau \partial H_{i,j} = X_{i,j} - (1 + A_{i,j})H_{i,j} + \beta A_{i,j} - \gamma I_{i,j}, \quad (A1)
\]

\[
\tau_{ad} \partial A_{i,j} = -A_{i,j} + aS[H_{i,j}]; i, j \in \{1, 2\}, \quad (A2)
\]

where the indices \(i\) and \(j\) refer to eye of origin and motion on-direction, respectively. The local field activity of each unit \((H)\) was converted into a spike-rate output via a sigmoid function \((S)\) and depended on the visual input \((X)\), the adaptation dynamics \((A)\), the amount of cross-inhibition \((I)\), and an adaptation-dependent bias (for details, see Noest et al., 2007).

The inputs to the rivalry-resolving stage were calculated from the responses of the sensory units by subtracting and scaling their outputs:

\[
X_{i,j} = g(R_{i,p} - R_{i,q}); p, q \in \{1, 2\}, p = j, q \neq p, \quad (A3)
\]

where \(g = 0.1\) was the fixed gain factor that we adopted for scaling the inputs.

In one version of the model (Figure 9B), each adapting unit received cross-inhibition from the pooled outputs of its opposing units in both eyes:

\[
I_{i,j} = 0.5 \Sigma_k S[H_{k,l}]; k, l \in \{1, 2\}; l \neq j. \quad (A4)
\]

In the second version (Figure 9C), each adaptive unit received only cross-inhibition from the opposing unit in the other eye:

\[
I_{i,j} = S[H_{k,l}]; k, l \in \{1, 2\}; k \neq i, l \neq j. \quad (A5)
\]

In both versions of the model, the gain \((\gamma)\) of the cross-inhibition feedback was kept the same.

To account for the low temporal resolution and long temporal integration times in motion coherence detection, the integration time constant was set to \(\tau = 0.5\) s. The time constant for the adaptation was set to \(\tau_{ad} = 1.0\) s. Values of the remaining parameters were set to: \(\alpha = 3; \beta = 0.27; \gamma = 3.3\). The sigmoid “firing rate” function that converted the “local fields” \(H_{i,j}\) into momentary spike-rate outputs consisted of a non-linear, Naka–Rushton function:

\[
S[z > 0] = z^2/(z^2 + 1); S[z \leq 0] = 0. \quad (A6)
\]


