Chromatic input to motion processing in the absence of attention

Alexander Thiele a,1, Amy Rezec b, Karen R. Dobkins b,*

a Salk Institute for Biological Studies, La Jolla, CA 92037, USA
b Department of Psychology, University of California, 0109, San Diego, La Jolla, CA 92093, USA

Abstract

While several previous psychophysical and neurophysiological studies have demonstrated chromatic (red/green) input to motion processing, the nature of this input is still a matter of debate. In particular, recent controversy has developed regarding whether chromatic motion processing relies on lower-level processes [J. Neurosci. 14 (1994) 4854; 19 (1999) 6571] versus higher-level, attention- or salience-based mechanisms [Science 257 (1992) 1563; Proc. Natl. Acad. Sci. 96 (1999a) 8289; 96 (1999b) 15374]. In this study, we investigated the degree to which chromatic motion is influenced by attentional mechanisms by employing a dual-task paradigm, which allowed us to compare the strength of chromatic motion under conditions of poor versus full attention. Here, we found that for equiluminant red/green gratings, chromatic motion processing is as robust in poor, as in full, attention conditions. This lack of an attentional effect suggests that chromatic motion processing must rely, at least in part, on lower-level (i.e., pre-attentive) motion mechanisms. For non-equiluminant (e.g., red brighter than green) gratings, attention significantly modulates chromatic motion strength. Possible explanations for this latter result are discussed in the context of inherent salience differences between the bright-red and dim-green stripes of the heterochromatic grating. © 2002 Elsevier Science Ltd. All rights reserved.

Keywords: Attention; Chromatic; Motion; Salience

1. Introduction

Numerous psychophysical studies have shown that the primate motion system can use chromatic (red/green) contrast to discern direction, although the motion percept is typically less robust than that observed for achromatic (light/dark) stimuli (see Gegenfurtner and Hawken (1996) or Dobkins and Albright (1998) for reviews). Attempts to explore the neural basis of chromatic motion processing have focused on the middle temporal (MT) area of monkey visual cortex, which is thought to play a key role in motion perception. Mirroring psychophysical findings, several neurophysiological studies have shown that MT neurons are able to signal the direction of moving red/green gratings, although responses are significantly weaker than for moving achromatic gratings (e.g., Dobkins & Albright, 1994; Gegenfurtner et al., 1994; Thiele, Dobkins, & Albright, 1999; Thiele, Dobkins, & Albright, 2001). In the past, we (Dobkins & Albright, 1994; Thiele et al., 1999) have argued that the chromatic responses observed perceptually and in MT can be explained by signals arising within low-level, early stages of the magnocellular division of the visual system, which provides the bulk of the input to area MT (Maunsell, Nealey, & DePriest, 1990). Hence, we proposed that chromatic motion processing is mediated, at least in part, by lower-level mechanisms.

In contrast to the evidence for lower-level contribution, other studies have suggested that chromatic motion processing relies on higher-level, attentionally-based mechanisms. Of particular relevance is a study by Cavanagh (1992), which employed an “opposed motion stimulus” consisting of two sinusoidal gratings—one heterochromatic (equiluminant, red/green), the other achromatic (light yellow/dark yellow)—superimposed and moving in opposite directions. When subjects were required to attentively track the motion of the heterochromatic grating, they could do so with ease, even under conditions when the global percept of motion was in the opposite direction (i.e., in the direction of the achromatic grating). By contrast, subjects were poor at...
tracking the motion of the achromatic component that dominated the global percept. These results led Cavanagh to conclude that motion of an achromatic grating is mediated by lower-level mechanisms, while motion of a heterochromatic grating is mediated by higher-level, attentional tracking of select local features.

More recently, Lu, Lesmes, and Sperling (1999a,b) have suggested that chromatic motion at equiluminance is mediated entirely by a “third-order” motion mechanism (Lu & Sperling, 1996). This third-order system is believed to compute motion based on the outputs of a “feature-salience” map, which is influenced by both “bottom-up” cues (i.e., inherent salience differences between the features that make up a stimulus) and “top-down”, selective attention (that serves to increase the effective salience of an attended feature). Based on their model, these authors propose that motion of an equiluminant red/green grating can be discerned only when either (1) bottom-up differences in salience exist between the red and green stripes of the grating (for example, when it is presented on a green background, thereby increasing the inherent salience of the red stripes (Lu et al., 1999a)), or (2) top-down attention increases the salience of either the red or green stripes (for example, by asking subjects to attentively track the red stripes (Cavanagh, 1992), or simply attend to “red” (Blaser, Sperling, & Lu, 1999)).

In the present study, we investigated the influence of lower- versus higher-level mechanisms on chromatic motion processing by comparing the strength of chromatic motion under conditions of poor versus full attention. The results of these studies reveal robust chromatic input to motion processing even in the near-absence of attention, suggesting that chromatic motion processing relies, at least in part, on lower-level (i.e., pre-attentive) mechanisms.

2. Methods

Subjects: Eight subjects (including two of the authors) participated in these experiments. All subjects had normal or corrected-to-normal vision, and normal red–green color vision (as assessed by the Ishihara test for color deficiency).

Apparatus: Visual stimuli were generated using a SGT Pepper Graphics board (Number Nine Computer Corp.: 640 × 480 pixel resolution, analog RGB output, 8 bits/gun) residing in a Pentium-based PC. Stimuli were displayed on a 20° analog RGB monitor (Sony GDM 2000TC, 60 Hz, non-interlaced).

Stimuli: We employed an “opposed motion” stimulus (after Cavanagh and Anstis (1991), and see Thiele et al. (1999) for details), which consists of two sinusoidal gratings—one heterochromatic (red/green), the other achromatic (light yellow/dark yellow)—superimposed and moving in opposite directions (upwards vs. downwards in these experiments). The mean luminance (28 cd/m²) and chromaticity (CIE: x = 0.492, y = 0.446) of the stimulus was equated with that of the background. (Stimulus size = 4.7° × 4.7°, spatial frequency = 0.4 c/deg). Using this stimulus, an estimate of the strength of chromatic input to motion—referred to as the “equivalent luminance contrast” (or “EqLC”)—is obtained by determining the amount of luminance contrast in the achromatic grating required to null the motion of the heterochromatic grating. To this end, nine different luminance contrasts in the achromatic grating were employed (ranging from 1% to 50%). Note that these luminance contrast values refer to “effective” contrast, i.e., the contrast of the achromatic grating when superimposed upon the heterochromatic grating, which is half the contrast of the grating if it were presented alone. For the heterochromatic grating, the effective chromatic contrast was 50% the maximum available on our monitor, which produced 14.4% root mean squared contrast in L- and M-cones at V½ equiluminance (see Dobkins, Thiele, and Albright (2000) for details on calculations).

In order to investigate whether chromatic input to motion processing, as well as attentional effects, vary as a function of temporal frequency, estimates of EqLCs were obtained at both 2 and 8 Hz. In addition, EqLCs were obtained for three different luminance contrasts in the heterochromatic grating, which allowed us to measure the strength of chromatic input to motion processing when luminance contrast is added to a red/green grating: (1) 0%, i.e., equiluminance, which was determined separately for each subject using a minimally-distinct motion technique (Dobkins et al., 2000), (2) +25% (green more luminous than red, G > R) and (3) −25% (red more luminous than green, R > G). Note that these ±25% values were set relative to each subject’s equiluminance point, and that they refer to the “effective” luminance contrast in the grating (see above).

Paradigm: Subjects were tested in a dark room and viewed the video display binocularly from a chin rest situated 57 cm away. They were instructed to maintain fixation on a small central square, and provide perceptual reports via key-presses on a keyboard. No feedback was provided. Data were obtained in two separate attention conditions. In both, a stream of five letters (each lasting 66.6 ms with 100 ms blank in between) was presented inside the fixation square at the same time that the opposed motion stimulus was presented parafoveally, centered 3.7° to the right of fixation. The total stimulus duration was 800 ms. ²

²The stimulus parameters employed in these experiments, including stimulus eccentricity and duration, were chosen so that we could make comparisons with our neurophysiological data from macaques (Thiele et al., 1999, 2001).
In the first condition, which we refer to as full-attention, subjects performed a single task, i.e., they reported the perceived global direction (“upward” versus “downward”, 2-AFC) in the opposed motion stimulus using two digits on their right hand, ignoring the (irrelevant) stream of letters at the center of gaze. In the second condition, which we refer to as poor-attention, subjects performed a dual task. Here, they were required to count the number of vowels appearing inside the fixation square while concurrently discerning the direction of the opposed motion stimulus. Because the vowel-task was extremely demanding, subjects paid substantially less attention to the motion stimulus in this condition. For the vowel-task, as few as one, and as many as five, of the letters were vowels. Subjects used the five digits on their left hand to first report the number of vowels (5-AFC), and then two digits on their right hand to report the perceived direction of the opposed motion stimulus.

Before beginning the dual task, subjects received ample practice on the vowel-task alone (and were instructed to ignore the simultaneously-presented opposed motion stimulus). Testing on the dual-task paradigm commenced once the subject’s performance on the vowel-task asymptoted. Averaged across subjects, performance on the vowel-task alone was $75.5 \pm 6.8\%$ (where chance performance = $20\%$). In order to ensure that, in the dual-task condition, subjects devoted their attention almost exclusively to the vowel-task, we required that their vowel-counting performance in the dual-task be nearly identical to that obtained when tested on the vowel-task alone. Accordingly, the average performance on the vowel-task was only $3.3 \pm 4.4\%$ points lower in the dual-task condition, and this difference was not significant ($t = 2.08, p = 0.08$).

**Determining EqLC:** For each subject, the percentage of trials for which motion was reported in the achromatic direction was calculated as a function of luminance contrast in the achromatic grating. Weibull functions were fit to the data and the luminance contrast yielding 50% decisions in favor of the achromatic grating direction provided the “motion null point”. EqLC was then calculated as the difference between the motion null value and the luminance contrast in the heterochromatic grating. Thus, in the non-equiluminant case, EqLC reflects the strength of the chromatic portion of the heterochromatic grating.

In total, EqLC was obtained for 12 conditions, i.e., two attention conditions (full-attention and poor-attention), three heterochromatic luminance contrasts ($-25\%$, $0\%$ and $+25\%$) and two temporal frequencies (2 and $8\,$Hz). One hundred and forty trials went into each EqLC estimate, for a total of 1680 trials per subject. Testing was divided into 12 blocks, with heterochromatic luminance contrast varied within, and temporal frequency varied across, blocks. The first six blocks consisted of the full-attention condition, and were followed by six blocks of the poor-attention condition.

### 3. Results

**Example EqLC data:** Example psychometric functions from one subject tested with the opposed motion stimulus are presented in Fig. 1. The percentage of trials for which the subject reported motion in the achromatic direction is plotted as a function of luminance contrast in the achromatic grating, separately for the “full-attention” (open circles) and “poor-attention” (filled circles) conditions. In this example, the heterochromatic component contained +25% luminance contrast ($G > R$). Although the slopes of the functions were nearly identical in the two conditions, EqLC values were substantially higher in the full-attention condition. Specifically, the full-attention condition yielded a motion null at 32% contrast, and thus an EqLC value of 7% (i.e., 32–25%), whereas the poor-attention condition yielded an EqLC value of 2% (i.e., 27–25%).

**Group mean data:** Group mean EqLC values are plotted as a function of heterochromatic luminance contrast in Fig. 2(A), separately for the 2 Hz (left panel) and 8 Hz
In the full-attention condition, EqLC values were found to be significantly greater than zero across all conditions ($t > 3.82, p < 0.003, 1$-tail $t$-test). In addition, EqLCs were roughly constant with varying levels of heterochromatic luminance contrast (2 Hz: $F(2, 7) = 1.27, p = 0.31$, 8 Hz: $F(2, 7) = 1.20, p = 0.33$), a result that corroborates findings from previous studies (Cavanagh & Anstis, 1991; Thiele et al., 1999). Averaged across the three heterochromatic luminance contrasts, mean EqLC values were $6.8 \pm 2.7\%$ and $4.7 \pm 1.8\%$ for 2 and 8 Hz stimuli, respectively, and the effect of temporal frequency was significant ($F(1, 7) = 23.4, p < 0.02$).

The results from the poor-attention condition provide the first measure of chromatic input to motion in the near-absence of attention. Here, mean EqLC values were $4.7 \pm 1.8\%$ and $4.0 \pm 1.5\%$ for 2 and 8 Hz stimuli, respectively, which were not significantly different from one another ($F(1, 7) = 1.89, p = 0.21$). As for the full-attention condition, all EqLC values in the poor-attention condition were significantly above zero ($t > 3.00, p < 0.01, 1$-tail $t$-test). However, in contrast to the full-attention condition, EqLC values for the poor-attention condition were largest at equiluminance, declining away from equiluminance (although this effect was only significant for 2 Hz data: $F(2, 7) = 5.19, p < 0.02$; 8 Hz data: $F(2, 7) = 1.82, p = 0.20$). This decline in the effectiveness of chromatic input to motion with increasing luminance contrast in the heterochromatic grating can be accounted for by responses of magnocellular neurons at early stages of visual processing, an issue we return to in the Discussion.

In order to observe more directly the effects of attention on chromatic motion processing, difference scores, i.e., $\text{EqLC}_{\text{full-attention}} - \text{EqLC}_{\text{poor-attention}}$, were calculated for each subject. Group averaged data are plotted in Fig. 2(B). The results of these analyses reveal...
two main findings. First, attention significantly influenced the strength of chromatic input to motion at 2 Hz ($F(1, 7) = 23.4, p < 0.002$), but not at 8 Hz ($F(1, 7) = 2.55, p = 0.15$). Note that this differential effect of attention at 2 versus 8 Hz is unlikely to be explained by a lesser effectiveness of the concurrent vowel-task in reducing attention in the 8 Hz condition, since percent correct performance on the vowel-task was essentially identical for 2 Hz (72.3 ± 6.8%) and 8 Hz (72.1 ± 4.5%) conditions. Second, the effects of attention at 2 Hz were significant only for non-equiluminant stimuli (−25% (R > G): $t = 6.75, p < 0.001$, +25% (G > R): $t = 3.54, p < 0.01$), but not for equiluminant stimuli ($t = 0.44, p = 0.67$). Under non-equiluminant conditions, EqLC values were, on average, 3.2 ± 1.3% points lower (i.e., a decrease of 47%) when attention was largely withdrawn from the motion stimulus.

4. Discussion

The results of these experiments provide the first demonstration of robust chromatic input to motion processing in the near absence of attention, as evidenced by EqLC values that remain significantly above zero under poor attention conditions. This finding suggests that the perception of chromatic motion does not require the use of higher-level, attentionally-based mechanisms. We should point out, however, that our EqLC measures in the poor-attention condition could be slightly overestimated if, on some trials, subjects mistakenly directed their attention away from the central vowel-task and towards the opposed motion stimulus. However, we believe such occurrences (had they existed) were extremely infrequent since subjects’ performance on the concurrent vowel-task was as good as their performance on the vowel-task alone (see Methods), suggesting that subjects were, in fact, adequately maintaining their attention on the central task.

In sum, our results suggest that chromatic motion can be processed without attention, and thus support the notion that chromatic motion can be processed by solely lower-level (i.e., pre-attentive) mechanisms. This, of course, does not preclude the possibility that chromatic motion may also employ higher-level, attentionally-based mechanisms under other circumstances. As reviewed in the Introduction, previous studies have demonstrated the ability to attentively track a red stripe in a moving equiluminant (red/green) grating (Cavanagh, 1992), and that simply attending to “red” strengthens the motion percept of an equiluminant grating (Blaser et al., 1999). Both of these previous studies invoked top-down selective attention, which presumably acted to increase the effective salience of the red stripe(s). The resulting salience difference between the red and green stripes, in turn, should have boosted the overall effectiveness of the red/green grating. Thus, while these previous studies clearly demonstrate top-down influences on chromatic motion processing, the results of the present experiment demonstrate that lower-level mechanisms also contribute significantly.

Also relevant to the present study, Lu et al. (1999a,b) have recently suggested that chromatic motion at equiluminance can be discerned only when either top-down selective attention increases the effective salience of either the red or green stripes (as discussed above) or when bottom-up differences in salience exist between the red and green stripes of the grating (for example, when it is presented on a green background, thereby increasing the inherent salience of the red stripes (Lu et al., 1999a)). While we do not dispute the possibility that chromatic motion stimuli can activate a salience-based mechanism, we believe our finding of significant chromatic input to motion in the near-absence of attention makes a strong case against the notion of complete reliance on a salience-based system, for two reasons. First, the red and green stripes of our equiluminant gratings appeared isosalient with respect to the yellow background (but see Lu et al. (1999b)), and thus this stimulus should not have activated bottom-up salience-based mechanisms. Second, the subjects in our experiment were instructed to simply report the global direction of the stimulus, and thus—especially in the poor-attention condition—were extremely unlikely to have attentively tracked or attended to a particular feature of the red/green grating. Devoid of such top-down processing, the effective salience of the grating should not have been altered. For these reasons, we conclude that neither top-down nor bottom-up salience-based mechanisms contributed to the strong chromatic input to motion observed in the present study.

In further support of our proposal that chromatic motion processing need not rely on salience-based mechanisms, but instead can be explained by lower-level mechanisms, previous studies have suggested that the origins of chromatic signals to motion processing can be accounted for by the responses of magnocellular (M) cells early in visual processing (Dobbins & Albright, 1994, 1998; Thiele et al., 1999). These M cells, which provide the bulk of the input to cortical motion area MT (Maunsell et al., 1990), can provide chromatic signals at equiluminance in the form of responses to borders defined by red/green contrast (e.g., Vilberg, Lee, Kaiser, & Kremers, 1992) as well as residual responses resulting from differential sensitivity to red versus green across the M cell population (e.g., Logothetis, Schiller, Charles, & Hurlbert, 1990). Moreover, this latter type of M cell signal directly predicts decreasing EqLC values with increasing luminance contrast (see Cavanagh and Anstis (1991) and Thiele et al. (1999)). Based on these early M cell responses, it is not surprising that chromatic motion is robust in the near-absence of attention, and that the
effect of adding luminance contrast to the heterochromatic grating follows the M cell prediction (see Fig. 2(A), filled squares).

On a final note, we address our finding of significant attentional effects for only non-equiluminant stimuli moving at 2 Hz (see Fig. 2(B)). One possible explanation for this result is that, despite being instructed to report the global direction of motion, subjects tended to attentively track the direction of the red/green grating under the full-attention condition, which served to elevate their EqLCs. Such a phenomenon could account for our results if the ability to track moving heterochromatic gratings were restricted to (1) slowly moving (i.e., 2 Hz) gratings, and (2) non-equiluminant gratings. With regard to the latter, non-equiluminant stimuli may be particularly trackable if, for example, the bright-red stripes are perceived as more salient than the dim-green stripes. This salience difference may allow subjects to “latch on” to the salient red stripes and track their direction. Although we do not have an independent measure for salience, previous studies have shown that perceived differences between red and green stimuli are more salient when luminance contrast is added to the two colors (Hilz, Huppmann, & Cavonius, 1974; Boynton, Hayhoe, & MacLeod, 1977; Eskew, Stromeyer, Picotte, & Kronauer, 1991). Based on these findings, we expect that the luminance contrast in our non-equiluminant gratings served to produce an inherent salience difference between the bright-red and dim-green stripes.

To test this tracking hypothesis, in six of our subjects we employed a paradigm that measured the trackability of our stimuli (see Cavanagh (1992) for details). Tracking performance was found to be near ceiling for all luminance contrasts at 2 Hz ($R > G$: 95.7 ± 5.6%, equiluminance: 95.0 ± 3.2%, $G > R$: 97.7 ± 2.3%), yet near chance at 8 Hz ($R > G$: 64.2 ± 12.9%, equiluminance: 45.2 ± 7.1%, $G > R$: 55.0 ± 13.9%). Thus, in line with the tracking hypothesis, 2 Hz stimuli were significantly more trackable than 8 Hz stimuli. However, contradictory to this hypothesis, equiluminant 2 Hz gratings were found to be perfectly trackable. It is, of course, still possible that equiluminant gratings are less trackable than non-equiluminant gratings, however, our results cannot address this possibility since performance was at ceiling for all 2 Hz stimuli. Future experiments will be required to more precisely quantify differences in trackability across different stimuli.

An alternative, yet related, possibility for why significant effects of attention were observed only for non-equiluminant stimuli is based on a notion originated by Lee, Itti, Koch, and Braun (1999). Results from their study, which similar to the present study measured visual performance under poor and full attention conditions, suggest that global attention may serve to accentuate inherent salience differences between the features that make up an object. By this account, the attentional effects observed for non-equiluminant stimuli might be explained by global attention serving to accentuate the inherent salience difference between the bright-red and dim-green stripes, thus boosting the effectiveness of the red/green grating. Conversely, the lack of an attentional effect for equiluminant stimuli might be explained by the fact that the red and green stripes were roughly isosali ent, and thus contained no salience difference to be accentuated. If this type of salience-based mechanism can, in fact, account for the differential attentional effects for equiluminant versus non-equiluminant stimuli, the lack of attentional effects observed for our 8 Hz stimuli would suggest that salience-based mechanisms are relatively insensitive at higher temporal frequencies (and see Lu & Sperling (1995a,b, 1996)). Clearly, future experiments will be required in order to elucidate the mechanism underlying the differential attentional effects observed across stimulus conditions. In any event, in addition to our finding of robust chromatic input to motion processing under poor attention conditions, which suggests contribution from lower-level mechanisms, our finding of attentional modulation under some stimulus conditions suggests (in accordance with previous studies) that chromatic motion processing is also subject to higher-level influences.

Acknowledgements

This work was supported by a NIH grant EY12153 (KRD). We are grateful to Thomas Albright for the use of his equipment, and to Jochen Braun, John Reynolds, Patrick Cavanagh and Zhong-Lin Lu for helpful feedback on the manuscript.

References


