PERGAMON

Available online at www.sciencedirect.com
SCIENCE DIRECT®

Vision Research 43 (2003) 1323–1328

Vision Research
www.elsevier.com/locate/visres

Rapid Communication

Attentional selection of superimposed surfaces cannot be explained by modulation of the gain of color channels

Jude F. Mitchell, Gene R. Stoner, Mazyar Fallah, John H. Reynolds *

Systems Neurobiology Laboratory, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA 92037-1099, USA

Received 7 November 2002; received in revised form 4 February 2003

Abstract

When two differently colored, superimposed patterns of dots rotate in opposite directions, this yields the percept of two superimposed transparent surfaces. If observers are cued to attend to one set of dots, they are impaired in making judgments about the other set. Since the two sets of dots are overlapping, the cueing effect cannot be explained by spatial attention. This has led to the interpretation that the impairment reflects surface-based attentional selection. However, recent single-unit recording studies in monkeys have found that attention can modulate the gain of neurons tuned for features such as color. Thus, rather than reflecting the selection of a surface, the behavioral effects might simply reflect a reduction in the gain of color channels selective for the color of the uncued set of dots (feature-based attention), as if viewing the surfaces through a colored filter. If so, then the impairment should be eliminated when the two surfaces are made the same color. Instead, we find that the impairment persists with no reduction in strength. Our findings thus rule out the color gain explanation.

1. Introduction

Valdes-Sosa, Cobo, and Pinilla (2000) recently provided one of the most compelling examples yet of non-spatial attentional selection. Observers viewed two superimposed sets of dots, one red and one green, which rotated rigidly in opposite directions, yielding the percept of two superimposed transparent surfaces. Fixation point color directed attention to one or the other surface. Following a 750-ms period of rotation, the cued surface underwent a brief (150 ms) translation in one of eight directions. Following this first translation, both sets of dots resumed rotating and, after a variable delay, either one or the other surface translated. Observers were able to report the first translation accurately, and could also accurately report the second translation of the same surface, even if the two translations were separated by delays as short as 150 ms. In stark contrast, observers were very poor at judging the direction of the uncued surface and this impairment lasted for approximately 600 ms. These findings were taken as strong evidence for an object- or surface-based attention model (Duncan, 1984). According to this model, objects or surfaces are selected as integrated wholes and judgments about the selected object (in this case, translation direction) can be made more accurately than judgments about unattended objects.

An alternative, and arguably simpler, model is that the observed effects result from changes in the gain of color channels selective for the two sets of dots. If the gain of the color channel responding to the non-cued surface were reduced, this would be like viewing the uncued dots through a color filter. The resulting reduction in the salience of the uncued dots could easily account for impairments in judging changes in their motion. Similarly, if the color-gain of the cued dots increased, the highlighted dots would act as a source of noise that could also impair judgments of the uncued dots. A color gain explanation is consistent with single-unit recording studies in the monkey, which have found evidence for attention-dependent changes in the gain of individual features, including color (Haenny, Maunsell, & Schiller, 1988; Haenny & Schiller, 1988; Maunsell, Sclar, Nealey, & DePriest, 1991; McAdams & Maunsell, 2000; Motter, 1994a, 1994b; Treue & Martinez-Trujillo, 1999). This suggests that the behavioral impairments described above may be the result of *feature*-, not *surface*-based attention.

* Corresponding author.
E-mail address: reynolds@salk.edu (J.H. Reynolds).

0042-6989/03/$ - see front matter
doi:10.1016/S0042-6989(03)00123-8

Valdes-Sosa et al. (2000) partially addressed this issue by removing the initial rotational difference between the two sets of colored dots so that they appeared to move as a single surface. The logic of this control was that the impairment, if truly surface-based, should be abolished if both sets of colored dots are seen as a single surface due to their common motion. Indeed, under these conditions, the impairment did vanish. However, even if we accept the conclusion that the only relevant effect of removing the motion differences is the creation of a single perceptual surface, the failure to find an impairment when dots move together is a negative result. As such, from it we can only conclude that the perception of two surfaces (or, more precisely, motion differences between the two dot fields) is necessary to generate the impairment. We cannot conclude that a perception of two surfaces is sufficient to cause the impairment. Nor can we conclude that color differences are not necessary. The controls advanced by Valdes-Sosa et al. cannot, therefore, rule out the color gain model. By comparing performance with and without a color difference between the two surfaces, we performed a direct test of the color gain model. If the impairment survived this removal of the color difference, this would constitute a positive result, which would rule out the color gain model.

One obstacle to testing this hypothesis is that in the paradigm of Valdes-Sosa et al. (2000), attention was endogenously cued to the surface that matched the color of the fixation point. Recently, however, Reynolds, Alborzian, and Stoner (2003) found that the impairment reported by Valdes-Sosa et al. (2000) persisted with no reduction in strength after removal of the endogenous cue. This showed that the first translation acted as an exogenous cue that caused the surface that translated first to be selected, and impaired judgments of the other surface. In the present study we were therefore able to dispense with the endogenous cue and test whether the impairment persisted when the two surfaces were equated in color.

2. Methods

2.1. Stimuli and task

All experiments were conducted in a dark, quiet room. Equiluminance between red and green guns was established for each subject using heterochromatic flicker fusion (Ives, 1912), with a flicker rate of 60 Hz. The red gun was held constant at 255 and the green gun adjusted until minimal flicker was reported. This procedure was repeated eight times and the results averaged. The resulting gun values were used throughout the remainder of the experiment for each subject.

For the first session, subjects were given verbal instructions and practiced the task with translations of 150-ms duration. After the subject was comfortable that he or she understood the instructions, the experimenter varied the duration of the translation until subjects could achieve 70% accuracy in judging the first translation. Data from this practice/calibration session were discarded, and all analysis was performed on data collected in subsequent sessions.

The experimenter sat with the subject throughout every session to ensure that eye fixation monitoring was accurate. Subjects were allowed to pause and stretch any time they felt fatigued. Except during these pauses, they sat comfortably with head resting in a chin and forehead rest, to stabilize the head for eye position monitoring. Eye position was continuously monitored using an ISCAN Model ETL-400 infrared eye tracking system, operating at a 60 Hz sampling rate (ISCAN, Inc., Burlington, MA).

At the beginning of each trial, a gray fixation point (0.25×0.25 deg of visual arc) appeared at the center of a computer monitor (Trinitron Multiscan TC, operating at 60 Hz). After achieving fixation within a 1 degree square window observers initiated trials by key-press. Pressing the key caused two circular patterns of dots to appear, centered over the fixation point. One pattern rotated clockwise around the fixation point and the other rotated counter-clockwise around the fixation point. This yielded the percept of two superimposed transparent surfaces (see Fig. 1).

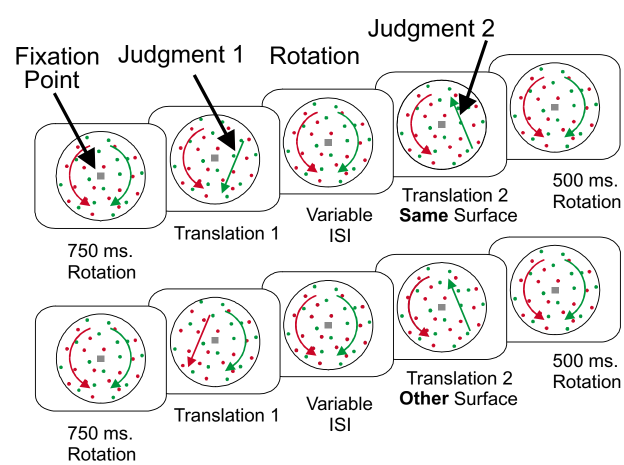

Fig. 1. Task: panels are arranged from left to right according to the sequence of events in each trial. The observer began each trial with a key press, resulting in a period of 750 ms during which the two surfaces rotated around the fixation point, in opposite directions. One surface then translated for 150 ms in one of the eight cardinal directions, while the other surface continued to rotate. Following this first translation, the two surfaces continued to rotate for a variable delay of 150–1050 ms, at which point one of the two surfaces, with equal probability, shifted for 150 ms. After this second shift, both surfaces continued to rotate for an additional 500 ms. Observers had to maintain fixation throughout the trial, and report the direction of each shift. On half of the trials, selected at random, one surface was red and the other was green. On the remaining trials, both surfaces were the same color, either both red or both green with equal probability.

On half of the trials, selected at random, one pattern of dots was red and the other was green. On the remaining half of the trials, the two sets of dots were identical in color (both red or both green, with equal probability). On trials in which both colors were present, the two sets of dots could rotate clockwise or counterclockwise, with equal probability. The average dot density of each dot field was 5 dots per square degree of visual arc. Each pattern of dots was circular and 2.75 deg in diameter. Each dot subtended 0.03 deg of visual arc. Both patterns rotated 50 deg around the center of rotation per second.

Every trial began with a 750-ms period during which both populations of dots continuously rotated. After this period of rotation, one of the sets of dots (selected at random, with equal probability) underwent a brief shift in one of eight directions while the uncued dots continued to rotate. As in the original study of Valdes-Sosa and colleagues, 60% of the dots translated coherently, while the remaining 40% of dots moved in the remaining 7 directions. This discouraged subjects from solving the task by attending to individual dots. All dots translated at a speed of 0.75 deg of visual arc per second.

At the end of this translation, both sets of dots rotated for a variable period of time, selected randomly with equal probability from five possible inter-stimulus intervals (ISIs: 150, 300, 450, 800, or 1050 ms). Following this rotation, one or the other sets of dots, selected at random with equal probability, translated. This was followed by a period of 500 ms during which both sets of dots resumed rotation, thereby masking the second translation.

On each trial, observers reported the directions of the two shifts, by pressing the key in the corresponding position around a numeric keypad. Observers were allowed to report the direction of each shift as soon as it occurred, but were required to maintain fixation within a 1 deg fixation window throughout the trial. Breaks of fixation, incorrect responses, and correct responses were signaled immediately by three different computer generated sounds.

2.2. Observers

All eight observers were paid to participate in the experiment. All had normal or corrected to normal vision. Six were women and two were men. All subjects were naïve as to the purpose of the experiment. Ages ranged from 17 to 21 years.

3. Results

Subjects ran between 960 and 1280 trials (mean, 1120 trials), yielding a mean of 56 repetitions (standard deviation, 5.2) in each of the 20 experimental conditions (five inter-stimulus intervals; two cueing conditions: translations on same surface, different surfaces; two color conditions: different colored surfaces or same colored surfaces).

In agreement with the findings of Valdes-Sosa et al. (2000) and Reynolds et al. (2003), subjects were able to report successive translations of one surface accurately, even when they occurred within 150 ms of one another, but were severely impaired in judging translations of first one surface, then the other. Fig. 2 shows average performance across subjects when the two surfaces were defined by dots of different colors. By convention, negative ISI's correspond to judgments of the first translation, and positive ISI's correspond to judgments of the second translation. Line color indicates whether the first and second translations occurred on the same surface (black) or different surfaces (grey). Error bars indicate

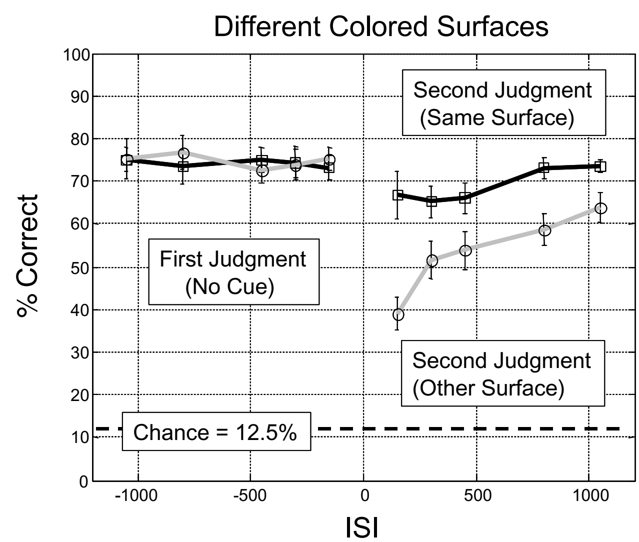

Fig. 2. Mean accuracy across eight subjects in reporting the direction of two successive translations, averaged across trials in which the dots defining the two surfaces were of different colors. Chance performance, indicated by dashed horizontal line, was 12.5%. ISI indicates the duration of the interval between the offset of the first translation and the onset of the second translation. By convention, negative ISI's correspond to judgments of the first translation, and positive ISI's correspond to judgments of the second translation. Thus, points at −1050 correspond to accuracy in judging the first translation, averaged across trials when the two translations were separated by an ISI of 1050 ms. Points at +1050 correspond to the second judgment, averaged across the same trials. Line color indicates whether the first and second translations occurred on the same surface (black) or different surfaces (gray). Error bars indicate standard errors of mean (SEM) performance across subjects. Observers accurately reported the direction of the first translation, regardless of whether the second translation also occurred on the same surface (black line) or occurred on the other surface (gray), and regardless of how soon after the first translation the second translation occurred. Subjects also reported the second translation accurately if it occurred on the surface that translated first. However, subjects were severely impaired in making judgments about the second translation when it occurred on the other surface. This impairment was greatest at the shortest ISI tested (150 ms) and gradually diminished over time.

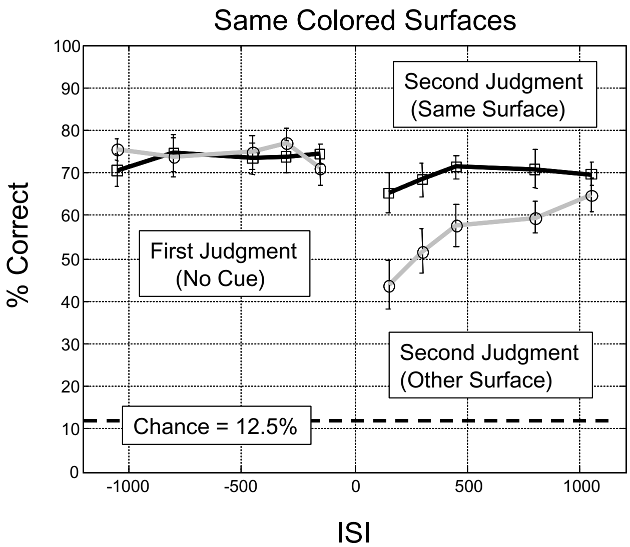

Fig. 3. Mean accuracy across eight subjects in reporting the direction of two successive translations, averaged across trials in which the dots defining the two surfaces were of the same color (either both red or both green). All conventions are identical to those in Fig. 2.

standard errors of the mean (SEM) performance across subjects (Fig. 3). Subjects accurately judged the first translation (left side of graph), and were also able to report the second translation of the same surface (solid black line, right side of graph). The performance for second judgments was slightly, but significantly, impaired (relative to the first translation judgments) when the translation occurred on the same surface, ($p < 0.01$) according to a three way ANOVA with (1) ISI, (2) subject, and (3) first/second judgment as factors (mean accuracy on first judgment = 73.8%, on second judgment = 69.1%). In contrast, subjects were severely impaired in judging second translations if they were of the other surface (grey line, right side of graph).

Removal of the color difference between surfaces did not change the impairment (Fig. 3). Performance on the first judgment was the same in both color conditions (mean accuracy 74.5% with color, 74% without color). The pattern of impairment observed for second judgments was remarkably similar in both conditions. On trials without a color difference, observers were still severely impaired when first one, and then the other surface translated, beginning at the shortest ISI tested (150 ms). Judgments of the second translation of the same surface were comparable across the two conditions (mean accuracy 69.0% with color, 69.2% without color). The presence or absence of a color difference had no significant effect on second translation judgment accuracy, according to a three-way ANOVA, with (1) ISI, (2) the presence or absence of a color difference and (3) surface (same versus different) as factors. There was also no significant interaction between the presence of a color difference and the other two variables ($p > 0.05$).

4. Discussion

4.1. Summary

Consistent with the findings of Valdes-Sosa et al. (2000) and Reynolds et al. (2003), we find that when one set of dots translates, this momentarily impairs the observer's ability to discriminate brief translations of a superimposed set of dots. Reynolds et al. (2003) found that this impairment is due to the first translation acting as an exogenous attentional cue. The present finding, that this impairment persists at full strength even when the two sets of dots are identical in color, rules out the possibility that the impairment is the result of a change in the gain of color selective neurons following the exogenous cue. As Valdes-Sosa and colleagues have already noted, the eight different directions of translation that were discriminated are identical across the two surfaces. The impairment cannot, therefore, be attributed to modulation of the gain of motion channels such as have been reported in a single-unit recording study of feature-based attention in area MT (Treue & Martinez-Trujillo, 1999). Thus, the observed pattern of results can be explained neither by a change in motion or color gain.

4.2. The importance of the rapid serial object transformation paradigm

The rapid serial object transformation (RSOT) paradigm that Valdes-Sosa and colleagues introduced has proven influential because it has enabled researchers to measure the time course of object-based selection, and it has done so while avoiding confounds that have complicated the interpretation of other object-based attention paradigms. Valdes-Sosa and colleagues have provided strong evidence that following attentional selection of an object, judgments of another, spatially superimposed object are severely impaired for a remarkably long period of time (~600 ms). This finding has led to additional psychophysical and ERP studies that have begun to relate the time course of the behavioral impairment to underlying neuronal mechanisms. For example, Pinilla, Cobo, Torres, and Valdes-Sosa (2001) found that translations of the unattended surface elicit than do translations of the attended surface, a smaller N1 ERP component, implicating a change in early sensory processing. The similarity of the time course of the impairment and the time course of the attentional blink has led to experiments that have investigated whether or not the two phenomena depend on the same underlying neural mechanisms.

Despite the obvious importance of the RSOT paradigm, all studies that have employed it have confounded object-based and color-based attention. The feature-based attention interpretation seems quite plausible considering that single-unit and fMRI studies (McAdams & Maunsell, 2000; Saenz, Buracas, & Boynton, 2002; Treue & Martinez-Trujillo, 1999) have demonstrated feature-based attentional modulation at the neuronal level in monkeys and humans. Given the importance of the RSOT paradigm and its prominent role in supporting object-based theories of attention, it was essential to resolve this confound. The present results rule out the feature-based interpretation, thus clarifying the meaning of the RSOT paradigm, and reinforce its importance in understanding the neural mechanisms of attention.

4.3. Possible remaining roles for color

These results do not rule out the possibility that a color difference might, under some circumstances, influence motion discrimination performance. Indeed, Croner and Albright (1997) found that when discriminating the motion of a set of coherently moving dots appearing among randomly moving dots, color differences between the two sets of dots allowed signal dots to be segmented from task-irrelevant noise dots thereby enhancing motion discrimination, relative to same-color conditions. Under the present conditions, however, segregation by common motion was apparently sufficient to completely segment the two sets of dots, as there were no observable differences in performance with and without a color difference.

4.4. Ruling out spatial attention and differences in frequency

Kramer and Jacobson (1991) have observed that some earlier studies of attention to superimposed objects or surfaces have used stimuli that were not entirely overlapping, raising the possibility that improved performance in discriminating features of cued stimuli could result from different distributions of resources in space. In an important study which addressed this concern, Blaser, Pylyshyn, and Holcombe (2000) superimposed two Gabor patches and had subjects track these "objects" as they changed independently in their orientation, spatial frequency, and color. They found that observers could reliably report which of the two Gabors had been cued at the beginning of a trial, indicating that they could attend to that Gabor, despite the presence of the second, spatially superimposed stimulus. In addition, observers were better at tracking two features of one Gabor than they were at tracking two features of different Gabors. This is consistent with a model in which attention selected one of the stimuli for processing, thereby inhibiting processing of the other stimulus. The individual stripes of the two Gabors are, however, salient features that occupy different locations in space. Because the color, spatial frequency, and orientation attributes of each Gabor are locally available from each stripe observers might have tracked the cued stimulus by attending to the location of one Gabor stripe. According to this explanation, subjects' impairments in reporting changes in the features of two Gabors could reflect difficulty in dividing spatial attention across the locations occupied by individual stripes of the two Gabors. Blaser and colleagues argue against this on the grounds that the spatial frequencies of their Gabors (1–8 cycles per degree, corresponding to stripe widths ranging from 30′ to 3.75′) were beyond the spatial resolution limit of attention. However, attentional resolution at the fovea has been estimated to be approximately 6′ (Intriligator & Cavanagh, 2001). Therefore, over most of the spatial frequency range used by Blaser and colleagues, attention to the individual bars of the superimposed Gabors cannot be ruled out. In contrast, in the Valdes-Sosa paradigm adapted in this study, the dots composing each rotating surface are homogenously distributed. Moreover, the dots making up the translating surface moved with partial coherence and this eliminates attention to individual dots as a viable strategy.

Another concern was raised by Watt (1988), who pointed out that the stimuli in most object-based attention studies differed in spatial frequency content. The dots defining the two surfaces in the present study were drawn from the same probability distribution, so any differences in the spatial frequency content of the two surfaces are minimal, arguing against the possibility that the observed attention effects reflect modulation of frequency filters.

4.5. Neural correlates of surface-based attention

Pinilla et al. (2001) have recently found direct evidence that motion selective neurons in human cortex are modulated in the RSOT paradigm. They recorded event-related potentials as subjects performed the standard RSOT task, and found task-dependent changes in the N1 component, which is elicited by motion-onset or changes in motion direction (Bach & Ullrich, 1994; Kuba & Kubová, 1992) and is thought to reflect activity in motion-sensitive cortical areas. The N1 component that was elicited by the second translation was diminished in magnitude if the translation occurred on the unattended surface.

This modulation in the magnitude of the N1 component was observed despite the fact that the translations in the two attention conditions were identical. A general increase in the gain of motion-selective neurons cannot, therefore, account for the object-specific modulation of N1. This effect could, however, be explained by changes in color gain. A change in the gain of color selective neurons would, in effect, be like viewing the stimuli through a color attenuating filter. For example, the N1 component elicited by the translation of the red surface would be diminished if subjects wore red-attenuating glasses.

Thus, the Pinilla et al. (2001) study does not, by itself, show that the modulation of the N1 component reflects the operation of surface-based attention. By demonstrating that the behavioral impairment persists at full strength when the color difference is eliminated, the present study rules out this feature-based explanation. Thus, the present study lends support to models in which motion-selective neuronal responses are modulated, but this modulation is guided by surface segmentation (Duncan, Humphreys, & Ward, 1997; Fallah & Reynolds, in press).

5. Conclusion

The RSOT paradigm introduced by Valdes-Sosa et al. (2000) to study surface-based attention convincingly ruled out contributions from spatial attention as well as feature-gain mechanisms based on direction of motion and spatial frequency. However, because the two stimuli in the original paradigm were differentiated by color, feature-gain based on color was still a plausible mechanism. We removed the color difference and found the same pattern of results. Our findings thus rule out feature-based attention and provide additional support for the conclusion that the observed impairment is surface-dependent and reflects surface-based attention.

Acknowledgements

We thank Tom Albright for discussions that helped motivate these experiments. We thank John Curtis, Catherine Williams and Laura Abavare for testing subjects. Funding provided by grants from the George Hoag Family Foundation and NEI grants 1R01EY13802-01 and RO1EY13802-01.

References

Bach, M., & Ullrich, D. (1994). Motion adaptation governs the shape of motion-evoked cortical potentials. *Vision Research, 34*, 1541–1547.

Blaser, E., Pylyshyn, Z. W., & Holcombe, A. O. (2000). Tracking an object through feature space. *Nature, 408*(6809), 196–199.

Croner, L. J., & Albright, T. D. (1997). Image segmentation enhances discrimination of motion in visual noise. *Vision Research, 37*(11), 1415–1427.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology. General, 113*(4), 501–517.

Duncan, J., Humphreys, G., & Ward, R. (1997). Competitive brain activity in visual attention. *Current Opinion in Neurobiology, 7*(2), 255–261.

Fallah, M., & Reynolds, J. H. (in press). The role of competitive circuits in macaque extrastriate cortex in selecting spatially superimposed stimuli. In N. Kanwisher & J. Duncan (Eds). *Attention and performance XX: Functional brain imaging of visual cognition*. Oxford: Oxford University Press.

Haenny, P. E., Maunsell, J. H., & Schiller, P. H. (1988). State dependent activity in monkey visual cortex. II. Retinal and extraretinal factors in V4. *Experimental Brain Research, 69*(2), 245–259.

Haenny, P. E., & Schiller, P. H. (1988). State dependent activity in monkey visual cortex. I. Single cell activity in V1 and V4 on visual tasks. *Experimental Brain Research, 69*(2), 225–244.

Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of attention. *Cognitive Psychology, 43*, 171–216.

Ives, H. E. (1912). On heterochromatic photometry. *Philosophical Magazine, 24*, 845–853.

Kramer, A. F., & Jacobson, A. (1991). Perceptual organization and focused attention: the role of objects and proximity in visual processing. *Perception and Psychophysics* (3), 267–284.

Kuba, M., & Kubová, Z. (1992). Visual evoked potentials specific for motion onset. *Documenta Ophthalmologica, 80*, 83–90.

Maunsell, J. H. R., Sclar, G., Nealey, T. A., & DePriest, D. D. (1991). Extraretinal representations in area V4 in the macaque monkey. *Visual Neuroscience, 7*, 561–573.

McAdams, C. J., & Maunsell, J. H. (2000). Attention to both space and feature modulates neuronal responses in macaque area V4. *Journal of Neurophysiology, 83*(3), 1751–1755.

Motter, B. C. (1994a). Neural correlates of attentive selection for color or luminance in extrastriate area V4. *Journal of Neuroscience, 14*(4), 2178–2189.

Motter, B. C. (1994b). Neural correlates of feature selective memory and pop-out in extrastriate area V4. *Journal of Neuroscience, 14*(4), 2190–2199.

Pinilla, T., Cobo, A., Torres, K., & Valdes-Sosa, M. (2001). Attentional shifts between surfaces: effects on detection and early brain potentials. *Vision Research, 41*, 1619–1630.

Reynolds, J. H., Alborzian, S., & Stoner, G. R. (2003). Surface-based exogenous cueing triggers automatic competitive selection. *Vision Research, 43*(1), 59–66.

Saenz, M., Buracas, G. T., & Boynton, G. M. (2002). Global effects of feature-based attention in human visual cortex. *Nature Neuroscience, 5*(7), 631–632.

Treue, S., & Martinez-Trujillo, J. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature, 399*(6736), 575–579.

Valdes-Sosa, M., Cobo, A., & Pinilla, T. (2000). Attention to object files defined by transparent motion. *Journal of Experimental Psychology. Human Perception and Performance, 26*(2), 488–505.

Watt, R. J. (1988). *Visual processing: computational, psychophysical and cognitive research*. Hillsdale, NJ: Erlbaum.