

SHORT COMMUNICATION

Binding of motion and colour is early and automatic

Erik Blaser,¹ Thomas Papathomas² and Zoltán Vidnyánszky^{3,4}

¹University of Massachusetts Boston, Department of Psychology, 100 Morrissey Blvd., Boston, Massachusetts 02125 USA

²Rutgers University, Laboratory of Vision Research and Department of Biomedical Engineering, 152 Frelinghuysen Road, Piscataway, New Jersey 08854 USA

³Neurobiology Research Group, Hungarian Academy of Sciences, Semmelweis University, 1094 Budapest, Hungary

⁴Faculty of Information Technology, Péter Pázmány Catholic University, Budapest, H-1083, Hungary

Keywords: after-effect, contingent adaptation, human, primary visual cortex, psychophysics

Abstract

At what stages of the human visual hierarchy different features are bound together, and whether this binding requires attention, is still highly debated. We used a colour-contingent motion after-effect (CCMAE) to study the binding of colour and motion signals. The logic of our approach was as follows: if CCMAEs can be evoked by targeted adaptation of early motion processing stages, without allowing for feedback from higher motion integration stages, then this would support our hypothesis that colour and motion are bound automatically on the basis of spatiotemporally local information. Our results show for the first time that CCMAEs can be evoked by adaptation to a locally paired opposite-motion dot display, a stimulus that, importantly, is known to trigger direction-specific responses in the primary visual cortex yet results in strong inhibition of the directional responses in area MT of macaques as well as in area MT+ in humans and, indeed, is perceived only as motionless flicker. The magnitude of the CCMAE in the locally paired condition was not significantly different from control conditions where the different directions were spatiotemporally separated (i.e. not locally paired) and therefore perceived as two moving fields. These findings provide evidence that adaptation at an early, local motion stage, and only adaptation at this stage, underlies this CCMAE, which in turn implies that spatiotemporally coincident colour and motion signals are bound automatically, most probably as early as cortical area V1, even when the association between colour and motion is perceptually inaccessible.

Introduction

Our visual experience is based on coherent objects; we do not perceive scenes of disembodied colours, orientations and spatial scales. However, processing is often specialized by such visual features, and indeed to some extent isolated in functionally distinct neural pathways and brain areas (Livingston & Hubel, 1988; Felleman & Van Essen, 1991). To solve this puzzle, binding mechanisms have been proposed that link visual features that are associated with the same surface or object. Recently, it was proposed that binding could take place at two different stages of visual processing (Treisman, 1999; Wolfe & Cave, 1999; Holcombe & Cavanagh, 2001): an early stage of processing where an initial, automatic binding of spatiotemporally colocalized visual features takes place, and/or a higher, object-level binding mechanism that requires attention and links all the features that belong to the same object.

Contingent after-effects: evidence for early feature binding

The conjecture of an automatic feature binding mechanism at the early stages of visual processing was based to a substantial extent on results from studies of visual contingent after-effects (Treisman, 1999; Wolfe

& Cave, 1999). There is converging evidence that the neural substrate of contingent after-effects, e.g. motion after-effects contingent upon colour (Favreau *et al.*, 1972) and colour after-effects contingent upon orientation (McCollough, 1965), is located in the early stages of visual cortical processing. It has been shown that contingent after-effects are specific for the retinal position of induction (Gibson, 1933; Gibson & Radner, 1937; Stromeyer, 1972), lack interocular transfer (Favreau *et al.*, 1972; Mayhew & Anstis, 1972; Murch, 1972), and are selective for low-level properties of stimuli such as orientation, colour and spatial frequency (for review see Humphrey & Goodale, 1998). These results have been interpreted as evidence for early, automatic binding mechanisms based on spatiotemporal proximity of visual features.

However, there is an important shortcoming of the experimental designs used in these previous studies, which questions the ability to draw conclusions regarding the locus of the feature binding mechanisms based on their results. In order for contingent after-effects to provide evidence for the existence of automatic local binding mechanisms, both feature pairs (say, red-vertical tilt vs. green-horizontal) must be presented simultaneously and must overlap spatially during adaptation. If, instead, contingent pairs of features are presented in temporal alternation (for instance, only red-vertical or green-horizontal at any given time) as they were in the above-mentioned studies on the McCollough effect, then a contingent after-effect could evolve due to the existence of an adaptation process that is contingent for all the features that are present simultaneously in the

Correspondence: Drs Erik Blaser & Zoltán Vidnyánszky, as above.

E-mail: erik.blaser@umb.edu; vidnyanszky@ana.sote.hu

Received 20 October 2004, revised 27 January 2005, accepted 2 February 2005

image, whether or not the features that are associated with the same object are bound together during adaptation. In other words, it is only when both feature pairs overlap in space and time that we can be assured that the contingency of after-effects reflects the existence of binding between the contingent features during adaptation. The present study used an experimental design that met this requirement.

Isolating early automatic binding of colour and motion

The aim of the present study was to investigate whether colour and motion, the prototypical example of features thought to be processed in distinct channels (Livingstone & Hubel, 1988; Felleman & Van Essen, 1991), are actually bound automatically and early. To this end, we tested whether colour-contingent motion after-effects (CCMAE) could be induced by a stimulus that targets adaptation to the early, local stage of motion processing, such as that corresponding to cortical area V1. Critically, too, this stimulus was designed to guard against any contribution from higher motion stages such as that corresponding to cortical area MT (the hallmark of such an area is that it computes 'global' motion by performing a motion integration process over a relatively large spatiotemporal scale; in contrast, the earlier, 'local' stage, such as cortical area V1, operates at what is probably the smallest meaningful spatiotemporal scale for motion processing). The adapting stimulus consisted of an aperture containing two overlapping sets of bi-coloured, oppositely moving dots, where the opposing motion fields were either transparent or locally paired. While transparent motion stimuli are perceived as two families of dots sliding over one another, locally paired stimuli are perceived as motionless flicker (Qian *et al.*, 1994a). At the higher stage, there is a strong mutual inhibition between the mechanisms responsive to the opposite, locally paired vectors, thereby suppressing activity, as confirmed by physiological studies in area MT of the macaque (Qian *et al.*, 1994b) and in the human MT+ complex (Heeger *et al.*, 1999). Thus in the case of the locally paired display, the two colours are associated with two different motion directions only locally; any broader spatiotemporal window cannot discern the correspondence and, indeed, perceptually there is no global motion and both colours are associated with a static, flickering surface. Area V1, in fact, is thought to be the area that works at a sufficiently spatiotemporally local scale to resolve locally paired dot displays (Qian *et al.*, 1994b). Therefore, if CCMAE can be evoked by adaptation with locally paired stimuli, this would provide evidence that colour and motion are bound early. Moreover, given the 'cancellation' at the higher stage with locally paired stimuli, this and higher stages can not provide any feedback that could contribute to the CCMAE. If an after-effect can nonetheless be induced then this early binding of colour and motion signals must happen automatically, that is, in the absence of feedback from higher stages.

To identify the locus of contingent adaptation mechanisms even more reliably, we set the spatiotemporal relationship between the different motion signals in our adaptor such that only neural mechanisms of very high resolution, requiring receptive fields as small as those of area V1 neurons, could associate the specific direction and colour signals in the stimulus.

Materials and methods

Subjects

Five naïve and two expert subjects were run in these experiments; all subjects were adults and had normal or corrected-to-normal vision. These experiments were undertaken with the understanding and

written consent of each subject, the review of the Federally mandated Institutional Review Board of the University of Massachusetts Boston, and in conformation with The Code of Ethics of the World Medical Association (Declaration of Helsinki).

Stimuli and apparatus

All of our stimuli consisted of 128 red and 128 green limited-lifetime dots, distributed randomly in a circular aperture with a diameter of 4.25° of visual angle. Red and green dots were calibrated for isoluminance for each of our subjects. Each dot subtended 0.05° and had a lifetime of 80 ms. When moving, dots drifted at a rate of 2.5° per second, rendering four 20-ms motion 'steps' before a dot disappeared and was redrawn in a random location in the aperture. Dots were presented on an otherwise dark screen. Subjects were seated in a dark room, 57 cm away from the display, and were instructed to maintain fixation on a central fixation point. Displays for particular experimental conditions are described in more detail below.

Procedure and conditions

We used a 'top-up' procedure for adaptation: trials were composed of an 'adaptation' phase (8 s in general, 45 s for the first trial) and a 500-ms 'test' phase. There were three conditions, run in blocks, which were differentiated by the type of adaptation that was used.

During transparent adaptation, subjects viewed a motion stimulus that consisted of one population of coloured dots drifting in one lateral direction, say red rightward, while the other family of coloured dots drifted in the opposite direction, in this case green leftward. Such a stimulus appears as two transparent surfaces sliding over one another.

Locally paired adaptation differs from transparent adaptation in that each red dot is paired with an oppositely moving green partner at close proximity and on a collision course (Qian *et al.*, 1994a). Given this close proximity and short lifetime, the two dots approach, cross, and pass each other in a very small spatiotemporal window (Fig. 1). Such a stimulus exhibits no global motion at all, and instead just appears as directionless flicker (Qian *et al.*, 1994a). (All subjects reported that locally paired displays appeared subjectively as directionless flicker but, to further ensure that local pairing does indeed eliminate global motion, we had subjects perform a control experiment where they attempted to make judgements about the direction of the dots. Subjects viewed a brief, 500 ms, display and were asked to report the motion direction of either the red or green dots. This control showed that subjects were not able to determine the direction when the dots were locally paired; average performance was at chance level, 50.3% correct.) We used a similar stimulus to calibrate each subject for red–green isoluminance. Using a minimum motion procedure, subjects adjusted green luminance, in a red–green locally paired stimulus, until global motion was obliterated.

For alternating adaptation, red and green dot populations alternated temporally, every second (1 s of red dots alone drifting rightward, for instance, then 1 s of green leftward, then 1 s of red rightward, and so on). To keep the total exposure to the motion of a particular set of colour dots constant across all these conditions, the adaptation time for the alternating condition was twice that of transparent and locally paired conditions.

We used a novel test stimulus to measure the magnitude of the resulting after-effect for all of these conditions: during test, both red and green dots are present in the display, with one family of dots, say green, completely static (with infinite lifetime) and with the other family, in this case red, drifting upward. This upward drift was also

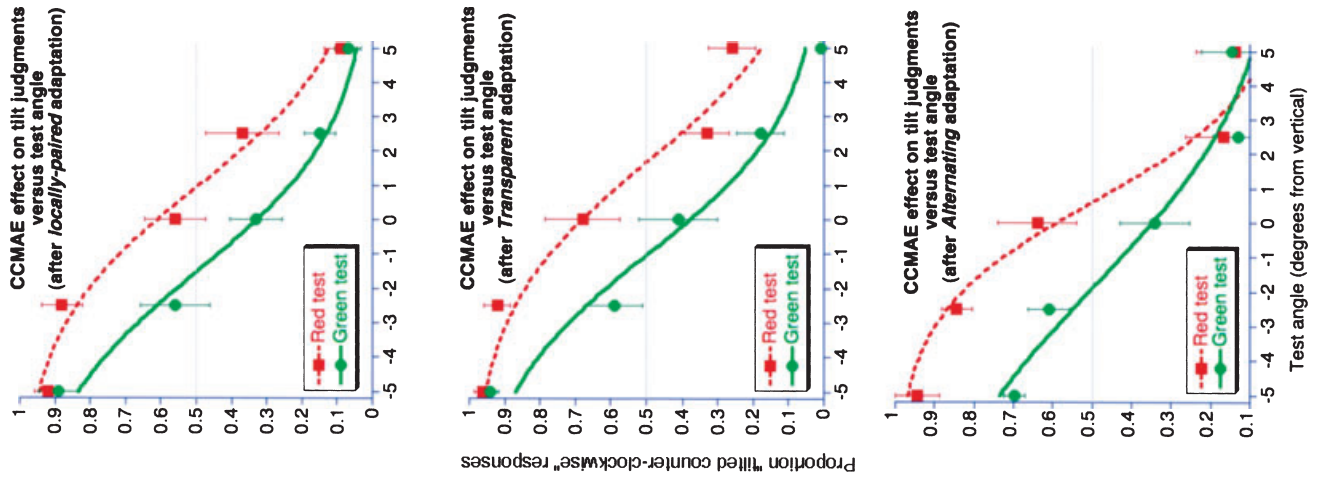


FIG. 2.

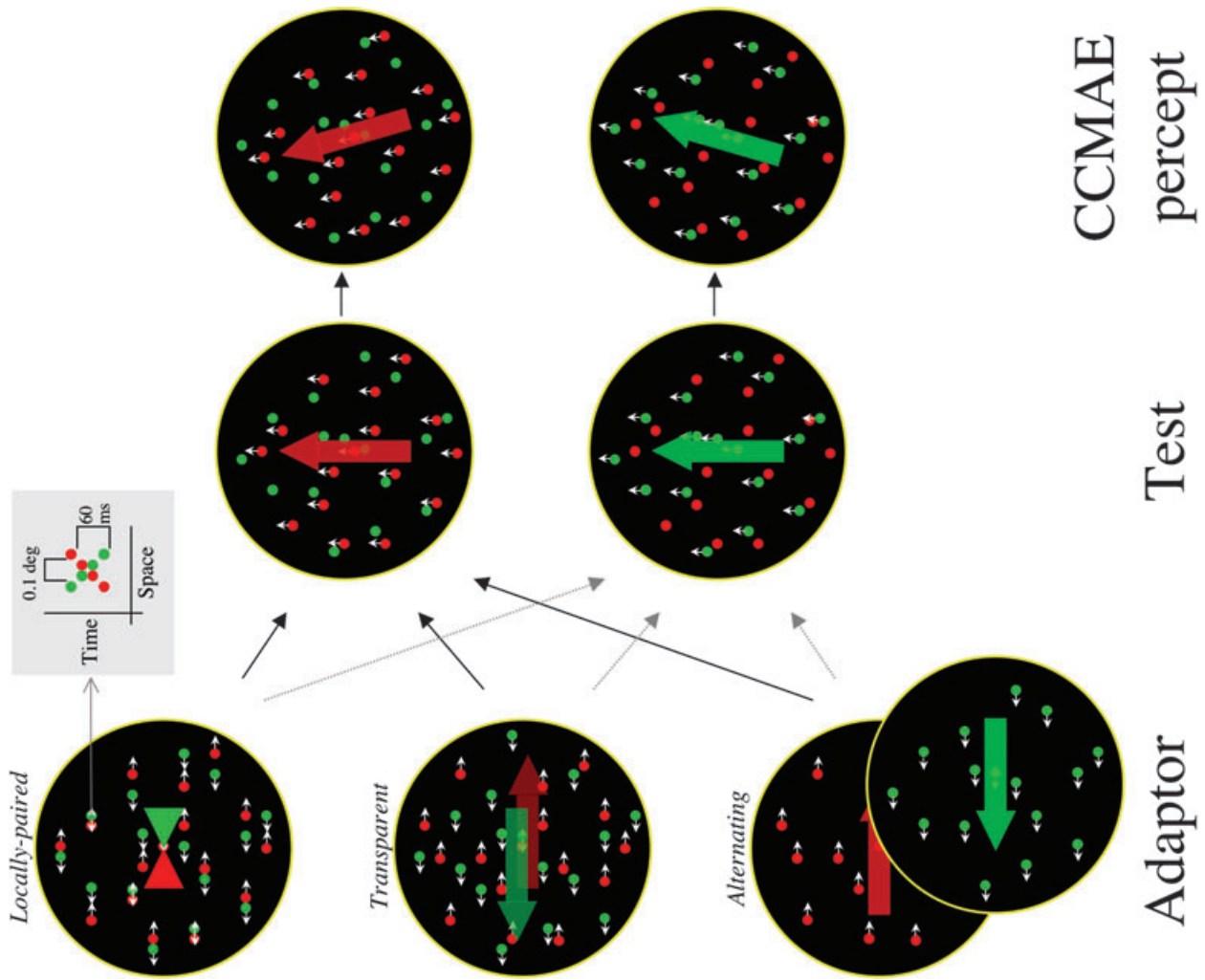


FIG. 1.

2.5 deg/s, the same as the adapting motion. The exact angle of this upward motion was determined by a method of constant stimuli, and took on a particular value (-5° , -2.5° , 0° vertical, 2.5° or 5°) randomly from trial to trial. In a two-alternative forced-choice paradigm, the subject's task was to indicate whether the axis of this upward vector was tilted to the right or the left. If there is a colour-contingent motion after-effect it will introduce a tilt in the motion of this upward moving field; in keeping with the above example, red dots would be deflected counter-clockwise and green clockwise. The magnitude of this angular deflection is our measure of the strength of the after-effect (Fig. 1).

Results

Significant colour-contingent motion after-effects were found for the transparent condition and, critically, for the locally paired condition as well. After adaptation with locally paired displays, which appear as directionless flicker, all five subjects exhibited CCMAE's, with an average induced deflection of 1.25° . Importantly, the magnitude of these effects is not significantly different from the 1.35° deflection obtained in the transparent adaptation condition (Fig. 2).

However, it has been shown that opposite directions inhibit each other at the higher stage responsible for global motion processing not only in the locally paired condition but also, although to a lesser degree, in the transparent motion condition. Therefore, one might argue that the reason why the magnitude of CCMAE was similar in the transparent and locally paired conditions is that adaptation at the higher cortical stage is substantially diminished already in the transparent condition and thus local pairing cannot have an additional inhibitory effect. The alternating control condition was designed to test this possibility. Because only one of the motion directions was present at a given time in this case, opposite motion directions could not inhibit each other. The average deflection in this experiment was 1.24° , which is very similar to that found in the transparent and locally paired conditions. This result strongly suggests that the magnitude of CCMAE's depends only on the strength of motion adaptation at the early stage.

Discussion

Our finding that CCMAE's can be evoked following adaptation to locally paired bivectorial motion provides evidence that, during adaptation, the specific colour and motion signals of the locally paired oppositely moving dots were bound early, at a stage where the two directions are still represented separately by the local direction-tuned filters (Qian *et al.*, 1994b). As was pointed out earlier, having the

different combinations of colours and motion directions present simultaneously in very close proximity during adaptation guards against the concern that CCMAE's actually result from adaptation mechanisms that form contingencies between all co-occurring visual features. Moreover, because local pairing diminished the representation of the opposite motion signals at the higher stage, it is highly unlikely that attention or feedback from this, or any other, higher cortical stage (including MT+) could contribute to the binding of these colour and motion signals. Thus, the results of the present study provide evidence for the existence of an early and automatic binding mechanism.

The neural locus of early binding

Our results also constrain the possible neural locus of binding mechanisms reflected in CCMAE: the size of the receptive fields of the neuronal populations involved must be small enough to resolve the contingencies between the closely paired colour and motion signals. Because the maximum spatial separation between the dots defining the oppositely moving vectors was approximately a tenth of a degree in the present study, it suggests that the locus of the feature binding that is reflected in contingent after-effects should be at the earliest cortical stage of visual processing where directional selectivity exists, i.e. in the primary visual cortex where the sizes of receptive fields provide the necessary resolution (Heeger *et al.*, 1999; Angelucci *et al.*, 2002).

There are two plausible explanations that could account for such early association between the processing of colour and motion: the first is based on the existence of double-duty neurons that are selective simultaneously both for colour and direction of motion. In fact, such colour-motion double-duty neurons were found in early visual cortical areas, including V1 (for review see Lennie, 1998; Gegenfurtner & Kiper, 2003). Another possible explanation is that active binding mechanisms exist to link the processing of neurons that are tuned to a single feature dimension (i.e. either colour or motion direction). Although our results cannot discriminate between these two possible neural substrates, they do provide a testable prediction regarding the locus of these processes: the neural correlates of feature binding, as well as contingent adaptation, should already be present at a very early stage of visual cortical processing, most probably as early as V1.

Contribution of additional binding stages

Our locally paired stimuli were specifically designed to eliminate the contribution, whether direct or via feedback, of any potential higher level binding mechanisms. It seems this approach was successful, as

FIG. 1. Schematic of experimental stimuli. The left three apertures contain a representation of the stimuli used in locally paired conditions (which appear as directionless flicker), transparent conditions (which appear as two fields of dots sliding over one another), and alternating conditions (where red and green dot families are interleaved in time), respectively. The inset panel shows spatial and temporal parameters for locally paired motion. This was 'four-step' limited-lifetime motion and after a particular pair completed this sequence it disappeared and reappeared in a random location. (transparent and alternating stimuli have identical parameters, save for the forced collisions). Tests were the same, independent of the adapting conditions: green static/red upward or red static/green upward (middle two apertures, respectively). The rightmost two apertures show the subjects' percept, given the presence of a CCMAE induced by the adaptor. The illusory deflection of the upward moving test vector occurs because of vector averaging between the upward vector of the test and the horizontal vector of the motion after-effect. The angular magnitude of this deflection is our measure of CCMAE strength.

FIG. 2. Results. The top panel shows psychometric functions reflecting subjects' judged tilt of the upward test vector as a function of the actual tilt, after locally paired adaptation. The presence of a CCMAE deflects this vector, introducing a bias into the subjects' responses. The strength of the motion after-effect can be measured in degrees by taking half of the distance between the psychometric functions garnered from red-upward test vectors and green-upward test vectors. The middle and bottom panels similarly show results after transparent and alternating adaptation, respectively. All psychometric functions are averages over four (one expert and three naive) subjects, and SEMs are shown. In all three conditions (locally paired, transparent and alternating) CCMAE's are significant, yielding 1.25, 1.35 and 1.24 degrees of illusory deflection, respectively, but are not significantly different from one another.

the size of the CCMAE's that resulted from locally paired conditions were not significantly different from those yielded by the transparent or alternating conditions, both of which had no such restrictions against the contribution of higher visual cortical areas. In spite of all this, it is still logically possible that there are additional binding stages besides the early, automatic stage that we have identified here; these stages could, for instance, link the features that belong to the same object. Several lines of evidence suggest that such object-level binding exists and requires focal attention (for review, see Robertson, 2003). The most convincing support is provided by the existence of illusory feature conjunctions (misbinding) in peripheral vision (Wu *et al.*, 2004) or when attentional resources are limited, and also in clinical cases with damage to attentional functions (Robertson, 2003).

Even if higher-level binding stages do exist in addition to the earlier stage our results point to it seems that these stages operate independently, as adaptation of the bound visual features at the higher cortical stages of visual processing does not contribute significantly to these contingent after-effects. This latter conclusion is based on our surprising findings that the magnitude of CCMAE in the locally paired condition was similar to that in the transparent and alternating conditions. It is well documented that there is mutual inhibition between the opposite motion directions at area MT of the macaque both in the transparent and locally paired condition, and that it is much stronger in the latter case than in the former (Qian *et al.*, 1994b). Thus, adaptation at the higher, global motion stage should be strongest in the alternating condition, weakest in the locally paired condition, and attain an intermediate value for the transparent condition. Our finding that the magnitude of CCMAE was similar in all three conditions implies that it was not affected by adaptation at this higher stage, this in turn further supports our conclusion that CCMAE is based solely on the adaptation of the earlier stage.

In conclusion, our findings imply that clusters of spatiotemporally co-occurring visual features, which are bound automatically and much earlier than previously thought, represent fundamental units of perceptual organization and the building blocks of our unified visual experience.

Acknowledgements

This work was supported by grants from the Hungarian National Research and Development Program (2/046/2004) to Z.V. and from the U.S. National Institutes of Health/National Eye Institute (R01 EY 013758-01) to T.V.P.

Abbreviation

CCMAE, color-contingent motion after-effect.

References

- Angelucci, A., Levitt, J.B., Walton, E.J., Hupe, J.M., Bullier, J. & Lund, J.S. (2002) Circuits for local and global signal integration in primary visual cortex. *J. Neurosci.*, **22**, 8633–8646.
- Favreau, O.E., Emerson, V.F. & Corballis, M.C. (1972) Motion perception: a color-contingent aftereffect. *Science*, **176**, 78–79.
- Felleman, D.J. & Van Essen, D.C. (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex*, **1**, 1–47.
- Gegenfurtner, K.R. & Kiper, D.C. (2003) Color vision. *Annu. Rev. Neurosci.*, **26**, 181–206.
- Gibson, J.J. (1933) Adaptation, after-effect and contrast in the perception of curved lines. *J. Exp. Psychol.*, **16**, 1–31.
- Gibson, J.J. & Radner, M. (1937) Adaptation, after-effect and contrast in the perception of tilted lines. *J. Exp. Psychol.*, **20**, 453–467.
- Heeger, D.J., Boynton, G.M., Demb, J.B., Seidemann, E. & Newsome, W.T. (1999) Motion opponency in visual cortex. *J. Neurosci.*, **19**, 7162–7174.
- Holcombe, A.O. & Cavanagh, P. (2001) Early binding of feature pairs for visual perception. *Nat. Neurosci.*, **4**, 127–128.
- Humphrey, G.K. & Goodale, M.A. (1998) Probing unconscious visual processing with the McCollough effect. *Conscious Cogn.*, **7**, 494–519.
- Lennie, P. (1998) Single units and visual cortical organization. *Perception*, **27**, 889–935.
- Livingstone, M. & Hubel, D. (1988) Segregation of form, color, movement, and depth. *Anat. Physiol. Percept. Sci.*, **240**, 740–749.
- Mayhew, J.E.W. & Anstis, S.M. (1972) Movement after-effects contingent on color, intensity and pattern. *Perception Psychophysics*, **12**, 77–85.
- McCollough, C. (1965) Color adaptation of edge-detectors in the human visual system. *Science*, **149**, 1115–1116.
- Murch, G.M. (1972) Binocular relationships in a size and color orientation specific after-effect. *J. Exp. Psychol.*, **93**, 30–34.
- Qian, N., Andersen, R.A. & Adelson, E.H. (1994a) Transparent motion perception as detection of unbalanced motion signals I. Psychophysics. *J. Neurosci.*, **14**, 7357–7366.
- Qian, N., Andersen, R.A. & Adelson, E.H. (1994b) Transparent motion perception as detection of unbalanced motion signals III. Modeling. *J. Neurosci.*, **14**, 7367–7380.
- Robertson, L.C. (2003) Binding, spatial attention and perceptual awareness. *Nat. Rev. Neurosci.*, **4**, 93–102.
- Stromeyer, C.F. (1972) Contour-contingent color aftereffects: retinal area specificity. *Am. J. Psychol.*, **85**, 227–235.
- Treisman, A. (1999) Solutions to the binding problem: Progress through controversy and convergence. *Neuron*, **24**, 105–110.
- Wolfe, J.M. & Cave, K.R. (1999) The psychophysical evidence for a binding problem in human vision. *Neuron*, **24**, 11–17.
- Wu, D.A., Kanai, R. & Shimojo, S. (2004) Steady-state misbinding of colour and motion. *Nature*, **429**, 262.