ELSEVIER

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Vision Research 46 (2006) 2615–2624

Vision Research

www.elsevier.com/locate/visres

Pushing the limits of transparent-motion detection with binocular disparity

John A. Greenwood *, Mark Edwards

School of Psychology, The Australian National University, Canberra 0200, Australia

Received 25 October 2005; received in revised form 13 January 2006

Abstract

When transparent motion is defined purely by direction differences, observers fail to detect more than two signal directions simultaneously [Edwards, M., & Greenwood, J.A. (2005). The perception of motion transparency: A signal-to-noise limit. *Vision Research, 45*, 1877–1884]. This limit is strongly related to signal-detection thresholds for transparent motion, which are several times higher than uni-directional thresholds. When the effective signal intensities are elevated by speed differences that drive independent global-motion systems, the transparent-motion limit can be extended to allow detection of three signals [Greenwood, J.A., & Edwards, M. (2006). An extension of transparent-motion detection limit using speed-tuned global-motion systems. *Vision Research, 46*, 1440–1449]. Because there are independent disparity-tuned global-motion systems, distributing transparent-motion signals across distinct depth planes also allows an increase in their effective signal intensity. In the present study, the addition of depth differences enabled the simultaneous detection of three signals. However, as with the addition of speed differences, observers were not able to detect four signals, which would be predicted if signal intensity were the sole constraint on transparent-motion detection. The combination of depth and speed produced similar results, suggesting that there is a strict higher-order limit, possibly related to attention, restricting the maximum number of signals that can be detected simultaneously to three.

Keywords: Motion perception; Transparency; Global motion; Depth perception; Stereopsis

1. Introduction

To detect moving objects, the visual system must integrate motion signals arising from the same object, whilst also discriminating between the motion signals arising from distinct objects. These processes are complicated in the case of transparent motion, which occurs when two or more objects pass through the same region of space at the same time without occlusion. By examining the detection of multiple transparent-motion signals, we can therefore gain insight into the limits of the segmentation and integration abilities of the motion processing system.

When direction is the sole basis for differentiating transparent-motion signals, no more than two global-motion directions can be detected simultaneously (Edwards & Greenwood, 2005; Mulligan, 1992). This limitation occurs because the signal intensities (as defined below) required to detect transparent-motion signals are many times higher than those required to detect uni-directional motion. When signal intensities are elevated through the addition of speed differences, the transparent-motion limit can be extended (Greenwood & Edwards, 2006). In the present study, we sought to extend the transparent-motion limit by introducing differences in binocular disparity to transparent-motion stimuli.

1.1. The transparent-motion limit and its extension

One of the most versatile stimuli used to examine transparent-motion detection is the random dot kinematogram, where two or more spatially intermingled groups of dots move in different directions within the same aperture (e.g.,

---

* Corresponding author. Fax: +612 6125 0499.
E-mail address: john.greenwood@anu.edu.au (J.A. Greenwood).

0042-6989/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2006.01.022

Clarke, 1977). A major advantage of these stimuli is that signal intensity can be clearly defined. Here, we define signal intensity as the proportion of dots moving in one signal direction. Dots moving in other directions, whether as noise (i.e., randomly moving dots) or within other transparent-motion signal directions, will act as noise for the detection of this signal. Thus, with direction differences as the sole basis for transparency, two signals can at most be presented at intensities of 50% each. The addition of a third signal reduces the intensity of all three signals to 33%.

To consider the detection of multiple transparent-motion signals, it is important to distinguish between simultaneous and sequential detection. Previous experiments have ensured simultaneous detection through tasks requiring the detection of all transparent-motion signals present within brief presentation times. In contrast, each signal could be detected sequentially, with the unattended signals treated as noise (Braddick, Wishart, & Curran, 2002). This latter mode of detection may resemble something closer to uni-directional detection for each signal (Edwards & Greenwood, 2005; Edwards & Nishida, 1999).

To determine the maximum number of signals that can be detected simultaneously, Mulligan (1992) required observers to report which of two intervals, presented for 250 ms, contained more signal directions. Comparisons were always between n and $n+1$ signals (e.g., 2 vs. 3), with direction as the sole basis for transparency. Results indicated that no more than two transparent-motion signals could be detected simultaneously. This limit appears to result from high signal-detection thresholds for transparency, as determined in a similar task where observers required intensities of 40% for each of two signals to be detected within transparent-motion stimuli (Edwards & Greenwood, 2005). When this is coupled with the reduction in intensity that results from increasing the number of signals, the maximum number of signals that can be presented at suprathreshold intensities is two. The dependence of this limit on signal intensity is consistent with the operation of the global-motion stage (Britten, Shadlen, Newsome, & Movshon, 1993), where transparent-motion detection is thought to occur (e.g., Qian, Andersen, & Adelson, 1994; Snowden, Treue, Erickson, & Andersen, 1991).

If the transparent-motion limit is the result of high signal-detection thresholds, it follows that an increase in signal intensities within transparent-motion stimuli should allow the detection of more than two signals. One way to increase the effective intensity of global-motion signals is to selectively drive independent global-motion systems. For instance, there appear to be at least two speed-tuned global-motion systems, which independently detect global-motion signals within distinct speed ranges (Edwards, Badcock, & Smith, 1998)[1]. Thus, transparent-motion signals selectively processed by one of these systems will not reduce the intensity of signals detected within other speed-tuned systems. When the appropriate speed differences were added to transparent-motion stimuli, the resulting increase in effective signal intensity allowed the simultaneous detection of three signals—an extension of the transparent-motion limit (Greenwood & Edwards, 2006).

However, if transparent-motion detection were limited solely by signal intensity, dividing the signals between two speed-tuned systems should have elevated four signals above threshold. The fact that the limit could be extended to three, but not four, suggests that there may be an additional higher-order limit on transparent-motion detection. If this were the case, any extension of the transparent-motion limit should be restricted to three. We sought to investigate this by examining the influence of binocular disparity differences on the transparent-motion limit.

1.2. Binocular disparity and transparent-motion detection

As well as selectivity for direction and speed, cells in cortical area V5/MT, the apparent neural site of global-motion processing, show selectivity for binocular disparity (Maunsell & Van Essen, 1983). These preferences appear to be organised into at least two independent global-motion systems. That is, for the detection of motion signals at crossed (near) disparities, noise dots presented at uncrossed (far) disparities have no effect on thresholds. The same is true of uncrossed global-motion signals with crossed noise dots (Snowden & Rossiter, 1999). Distinct planes of depth are required for this independent processing, as when signal dots are distributed across multiple depth planes, there is no evidence of disparity tuning (Hibbard, Bradshaw, & De Bruyn, 1999).

Furthermore, when transparent-motion signals are presented on distinct crossed and uncrossed depth planes, detection thresholds are lowered (Hibbard & Bradshaw, 1999). Distributing transparent-motion signals across these two disparity-tuned systems should therefore allow an extension of the transparent-motion limit by increasing the effective signal intensities. This improvement in performance should be of similar magnitude to that obtained using speed differences (Greenwood & Edwards, 2006), allowing the simultaneous detection of at least three transparent-motion signals.

Two experiments were conducted in order to assess this hypothesis. In Experiment 1, we established binocular disparity values that are sufficient to allow independent global-motion detection. These disparities were then used within transparent-motion stimuli in Experiment 2 in order to extend the transparent-motion detection limit.

2. Experiment 1: Sensitivity of the disparity-tuned systems

We first sought to replicate the results of Snowden and Rossiter (1999) with a uni-directional global-motion task designed to be as similar as possible to that used in our transparent-motion experiments.

---

[1] These systems may represent two non-overlapping points within a continuum of speed-tuned systems (van Boxtel & Erkelens, 2005). This may also be the case for the disparity-tuned systems examined in the present study.

2.1. Method

2.1.1. Observers

Three observers were used in both experiments: one of the authors (J.A.G.) and two naïve observers (E.A.E. and L.E.G.). All had normal or corrected-to-normal visual acuity, with no history of visual disorders.

2.1.2. Apparatus

Stimuli were generated with a Cambridge Research Systems VSG 2/5 in a host PC, and presented on a Clinton Monoray monitor with a refresh rate of 120 Hz and a 1024 × 769 pixel resolution. Binocular disparity was generated by Cambridge Research Systems FE-1 shutter goggles, which allow the presentation of alternating frames to each of the two eyes separately. The alternation of the shutters was synchronised with the monitor, with the image to each eye refreshed at 60 Hz (though stimuli were updated less frequently, as described below). Stimuli were viewed from 1 m, with head movements restricted by a chin rest. Observers responded to trials via the mouse buttons. The same apparatus was used for both experiments.

2.1.3. Stimuli

Global-motion stimuli were presented within a circular aperture of 11.5° diameter. Either 60 or 120 circular dots were presented, each with a diameter of 0.14°. This gave a dot density of 0.6 dots/deg^2 with 60 dots and 1.2 dots/deg^2 with 120 dots, both of which minimise the occurrence of motion correspondence errors (Williams & Sekuler, 1984). The luminance of the background was 82 cd/m^2. Dots were defined by a luminance increment, with a Michelson contrast of 30%.

Crossed disparities (near depth) were generated by shifting dots presented to the left eye by 0.05° (2 pixels) rightwards and dots to the right eye by 0.05° (2 pixels) leftwards, giving a binocular disparity of 0.1°. Uncrossed disparities required the inverse shift. A 0.3 × 0.3° fixation cross was provided to minimise eye movements. To allow observers to monitor their vergence state, nonius lines of 0.2° in length were presented 0.1° above and below the fixation cross to the left eye and right eye, respectively.

Dots were assigned as signal or noise at the beginning of each stimulus interval and moved in a continuous trajectory for the entire duration. Dots that moved beyond the aperture boundary were re-plotted in the opposite half of the aperture, based on the direction of motion.

2.1.4. Procedure

A temporal two-alternative forced-choice (2AFC) procedure was used. One interval contained a global-motion signal of varying intensity, with the remaining dots set as noise. The other consisted solely of noise dots. Observers were required to indicate which interval contained the global-motion signal, analogous to a 0 vs. 1 comparison in the n vs. $n + 1$ scheme used in our transparent-motion tasks.

The signal direction was chosen randomly from a rectangular distribution from 0 to 360°, with noise directions selected from the same distribution without replacement. Dots moved with a step size of 0.14° (6 pixels) per motion frame, giving a speed of 4.1°/s. This is close to the median speed preference for V5 cells (Lagae, Raiguel, & Orban, 1993), without exceeding transparent-motion $d_{\max}$ thresholds (Snowden, 1989).

Prior to each trial, the fixation cross and nonius lines were presented. When the nonius lines were perceived as aligned with the fixation cross, observers initiated the trial via the mouse buttons. The two stimulus intervals were then presented, each consisting of six frames in total. Because two transparent depth planes are more difficult to resolve than single planes in isolation (Akerstrom & Todd, 1988; Wallace & Mamassian, 2004), we wanted to ensure that performance in our mixed-disparity tasks was not hindered by inadequate resolution of the depth planes. Thus, for both intervals (in all conditions) the initial frame of dots was presented for 200 ms. However, the task could not be performed based on this static frame alone.

Subsequent motion frames within each interval were each presented for 33 ms. This refresh rate gave a smooth percept of motion whilst also allowing accurate displacement of the dots in a large number of directions. Dots within our stimuli thus made five displacements, with an effective stimulus duration of approximately 167 ms. Stimulus intervals were separated by a 1 s blank interval to minimise the effects of hysteresis (Williams, Phillips, & Sekuler, 1986).

Signal-to-noise detection thresholds were assessed within six conditions. For each of the crossed and uncrossed disparity planes, thresholds were measured with both 60 and 120 dots. Two mixed-disparity conditions were also conducted. The first contained 60 dots with crossed disparity from which the signal was drawn, with 60 uncrossed noise dots (*near mix*). The other contained the inverse arrangement (*far mix*). If global-motion signals are detected independently on each of these depth planes, thresholds in mixed-disparity conditions should be equivalent to thresholds with 60 same-disparity dots, both of which should be lower than thresholds with 120 same-disparity dots.

Thresholds were assessed with a modified staircase procedure (Levitt, 1971), with each of the six conditions tested within distinct staircases. Signal intensity began at 60 dots (out of 60 or 120 dots total, depending on condition) and was varied with a 3 down/1 up staircase converging on 79% correct performance. Eight reversal points were collected, with thresholds taken as the mean of the last six. Ten staircases were completed for each condition in random order. Stimuli were viewed in a dark room, following 5 min of dark adaptation. No feedback was given regarding performance during trials.

2.2. Results and discussion

Scores in both experiments were screened for outliers, with additional trials run when scores exceeded 2.5

standard deviations from the mean. For each of the six conditions, the resulting mean and standard error are displayed in Fig. 1. Conditions with the signal dots at crossed disparities (near) are presented left of the dividing line, with uncrossed (far) signal conditions on the right.

Each observer displays the expected pattern. Around 10–15 signal dots were required to reach threshold in conditions with 60 dots (*near 60*, *far 60*). The addition of 60 noise dots at the same disparity (*near 120*, *far 120*) raised thresholds to between 20 and 25 dots. In contrast, the addition of noise dots on a different depth plane (*near mix*, *far mix*) had no effect on thresholds, which were the same as with 60 dots in isolation. This suggests that global-motion signals presented at these binocular disparity values were processed independently by distinct global-motion systems, replicating the results of Snowden and Rossiter (1999) with our stimulus parameters. As in our previous study (Greenwood & Edwards, 2006), thresholds are slightly higher than those typically seen in global-motion tasks (Braddick, 1995), which is most likely due to direction uncertainty arising from the randomised signal direction (Ball & Sekuler, 1980).

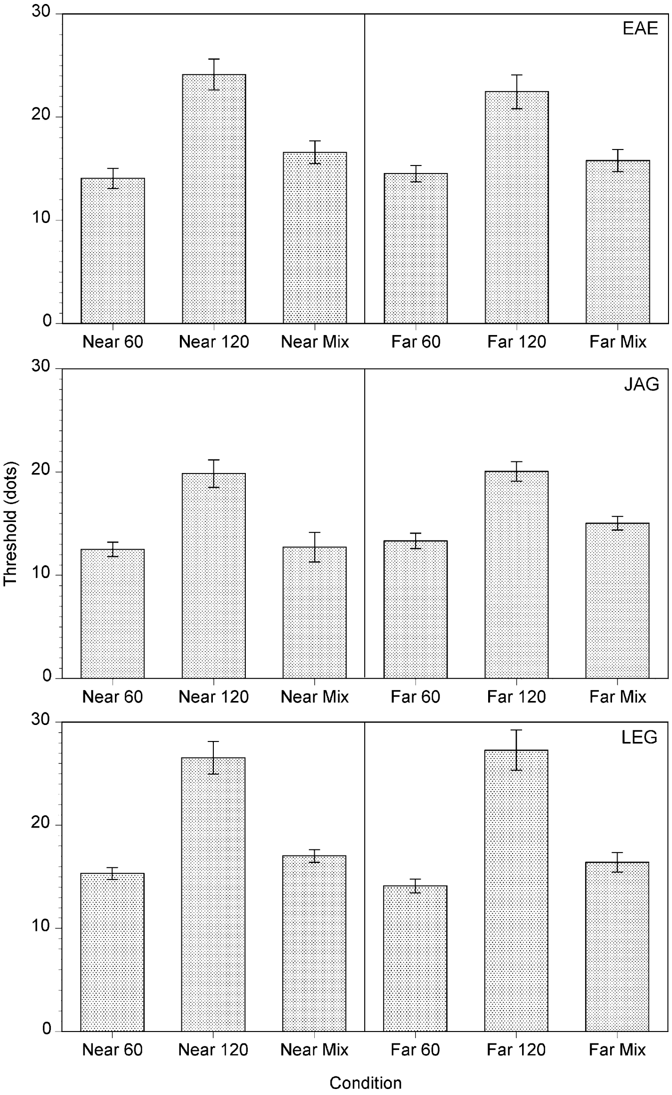

Fig. 1. Mean thresholds for global-motion signal detection, expressed as the number of signal dots required. Conditions with crossed (near) binocular disparity are presented left of the dividing line, with uncrossed (far) conditions on the left. For each of the two disparity planes, thresholds were obtained with 60 and 120 dots. Mixed conditions contained both crossed and uncrossed disparities, with target dots positioned on one plane only. Each data point represents the average of ten staircase estimates; error bars represent 1 *SEM*.

3. Experiment 2: Transparent-motion detection with disparity differences

The disparity values selected in Experiment 1 were subsequently added to transparent-motion stimuli in order to elevate the effective signal intensities. Because the transparent-motion limit appears to be the result of high global-motion signal-detection thresholds, this elevation in signal intensity should allow an extension of the limit. That is, observers should be able to detect more than two transparent-motion signals simultaneously, as in our previous study using speed differences (Greenwood & Edwards, 2006). If the limit of two is applicable to each disparity-tuned system, it may even be possible for observers to detect up to four transparent-motion signals. However, a higher-order limit of three, as suggested by our previous results, would restrict any extension of the limit.

3.1. Method

3.1.1. Procedure

The aperture configuration was the same as in Experiment 1, with 120 dots present. A temporal 2AFC procedure was used, as in our previous studies (Edwards & Greenwood, 2005; Greenwood & Edwards, 2006). A static frame again preceded each interval for 200 ms, with stimuli then moving for approximately 167 ms. Stimulus intervals were separated by a 1 s blank interval.

In each trial, one interval contained n transparent-motion signals, with $n+1$ signals in the other. Between one and five signals were presented according to the method of constant stimuli, making four comparisons: 1 vs. 2, 2 vs. 3, 3 vs. 4, or 4 vs. 5. Presentation order was randomised, with observers required to indicate which interval contained the greater number of signals. This task requires that all signals within an interval be detected in order to perform the required comparisons. Thus, when paired with brief presentation times, simultaneous detection of the signals is ensured (Braddick et al., 2002; Edwards & Greenwood, 2005).

The direction of each signal group was determined randomly, with the sole constraint that directions differed by at least 45°. This was implemented because the detection of transparency with angular separations lower than 45° is difficult when direction is the sole basis for transparency (Edwards & Nishida, 1999; Smith, Curran, & Braddick, 1999). For consistency, the 45° minimum separation was also maintained in the mixed-disparity conditions.

The use of randomised directions ensured that observers had to detect each signal within the interval to perform the task, rather than simply responding to the presence or absence of a single direction. This also minimises the occurrence of any patterns of motion such as motion parallax, as well as any direction-specific adaptation (e.g., Raymond, 1993). *Fixed-walk* dot motion was used to avoid the interference that occurs with rapid changes in direction (Watamaniuk, Flinn, & Stohr, 2003).

Four conditions were conducted: two with all dots at the same disparity (*all near*, *all far*) and two mixed-disparity conditions where dots were divided between crossed and uncrossed disparities. No noise dots were present within stimuli in this experiment. The same-disparity conditions gave the baseline performance at each of the near and far depth planes. For these stimuli, each signal consisted of an equal proportion of the total dots.

The two mixed-disparity conditions differed in the proportion of signal directions on each plane, with either the majority of signals on the near plane (*majority near*) or the far plane (*majority far*). In both cases, each depth plane always contained half the total dots. This meant that the proportion of dots at each depth plane could not be used as a cue to the number of signals present. Given the transparent-motion limit of two when direction is the sole basis for transparency, no more than two signals were assigned to the same depth plane where possible. With five signals, two signals were assigned to one depth plane, and three to the other. Dots with the same signal direction were always on the same depth plane. The number of signals presented within each of the two depth planes for the mixed-disparity conditions, and the resulting signal intensities, are presented in Table 1.

For the same-disparity conditions, a block of trials contained ten of each n vs. $n+1$ signal comparison, making 40 trials per block. The 1 vs. 2 comparisons were excluded from the mixed-disparity conditions, which were interleaved within the same block to make 60 trials. Thus, observers did not know which of the depth planes would contain the greater number of signals. Each block was presented separately, with 10 blocks for each condition completed in random order. Responses were converted into percent-correct scores.

Table 1
Signal composition of the mixed-disparity transparent-motion stimuli for Experiment 2

Total no. of signals	No. of near signals (% of total dots)	No. of far signals (% of total dots)
2	1 (50)	1 (50)
3 (*majority near*)	2 (25)	1 (50)
3 (*majority far*)	1 (50)	2 (25)
4	2 (25)	2 (25)
5 (*majority near*)	3 (16.7)	2 (25)
5 (*majority far*)	2 (25)	3 (16.7)

Numbers in parentheses indicate the percentage of total dots assigned to each of the signals. Note that because the two disparity-tuned global-motion systems are independent, the effective signal intensities would approximately double.

3.2. Results and discussion

The mean percent-correct scores and standard error within the four conditions are plotted in Fig. 2, where chance-level performance corresponds to 50% correct. The same-disparity conditions (*all near*, *all far*) demonstrate the standard pattern of results when direction is the sole basis for transparency: 1 vs. 2 and 2 vs. 3 comparisons were both performed with a high level of accuracy, while performance drops significantly with higher signal numbers. The 3 vs. 4 comparisons were close to chance levels, with a slight decrease continuing for the 4 vs. 5 comparisons. This suggests that performance in the 2 vs. 3 comparison was based solely on the detection of transparent motion in the two-signal interval. These signals could be differentiated from the three-signal interval, though the poor performance in the 3 vs. 4 comparison demonstrates that three signals were not detected. The mechanisms of this limitation will be considered further in Section 4.1.

When the transparent-motion signals were distributed across two depth planes in the mixed-disparity conditions, performance on the 3 vs. 4 comparisons rose from near-chance levels to between 85 and 90% correct. There was a slight tendency for *majority-far* conditions to be more difficult than the *majority-near* conditions, though differences were small. This may relate to the *front effect* (Lehmkuhle & Fox, 1980), though no such asymmetry is evident in the coherence thresholds measured in Experiment 1. Despite this minor asymmetry, our results indicate that observers were able to detect three signals and discriminate them from four. However, they were unable to detect four signals, as seen in the chance-level performance in 4 vs. 5 comparisons. Nonetheless, the addition of depth differences allowed the detection of up to three transparent-motion signals—an extension of the transparent-motion limit.

To assess the magnitude of this extension, difference scores were calculated for the 3 vs. 4 comparisons. For each observer, the same-disparity 3 vs. 4 scores were averaged to give a baseline level of performance. Mixed-disparity scores had this baseline subtracted and were then divided by the baseline to give percent-improvement scores. Averaged across the three observers, performance on the 3 vs. 4 comparisons was improved by 48% with the *majority-near* signal configuration, and by 41% with the *majority-far* configuration. This is of the same magnitude as the improvements produced by speed differences, which improved performance by 40–45% in our previous study (Greenwood & Edwards, 2006). Both of these stimulus features produced improvements that are substantially larger than those elicited by contrast-polarity differences (18–23%), which do not have independent global-motion systems (Edwards & Badcock, 1994).

4. General discussion

This study provides further evidence that the transparent-motion limit of two is not a fixed numerical restriction.

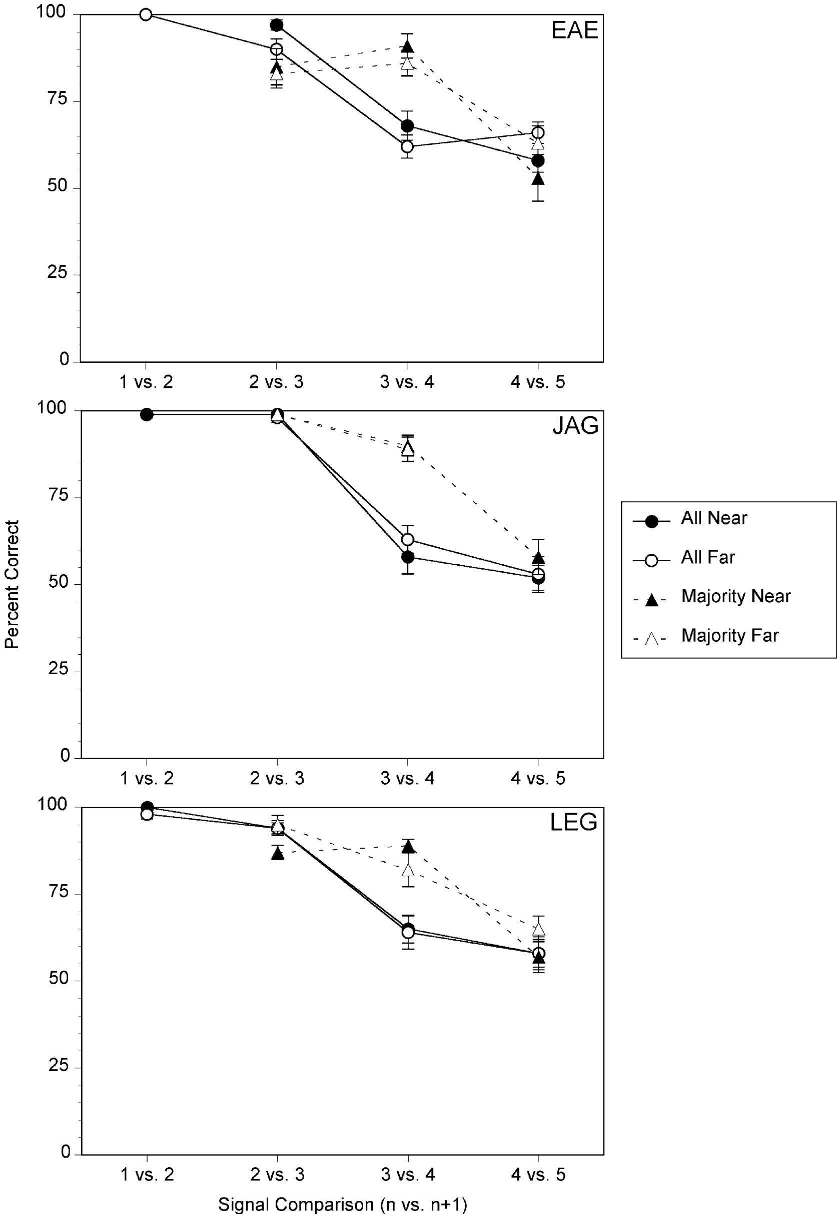

Fig. 2. Mean percent-correct scores as a function of the *n* vs. *n* + 1 transparent-motion signal comparisons. In the same-disparity conditions, dots were presented with either crossed (*all near*, filled circles) or uncrossed disparities (*all far*, open circles). Mixed-disparity conditions contained both disparity planes, with either more near than far signal directions (*majority near*, filled triangles), or vice versa (*majority far*, open triangles). Each data point is the mean of 100 trials; error bars represent 1 *SEM*.

In Experiment 1, we established that noise dots presented on far depth planes have no effect on signal detection within the global-motion system tuned to near depth planes, and vice versa. When transparent-motion signals were distributed across these two depth planes in Experiment 2, observers were able to detect up to three signal directions—an extension of the transparency limit.

4.1. *Basis of the transparent-motion limit extension*

This disparity-based extension of the transparent-motion limit adds further support to the notion that the limit of two arises from high signal-detection thresholds for transparency. The addition of signal directions ordinarily reduces the intensity of all signals present, meaning that the maximum number of signals that can be presented at intensities above the transparent-motion detection threshold is two (Edwards & Greenwood, 2005). This result was replicated in the same-disparity conditions of Experiment 2. Though our results indicate that transparency was not seen within these three-signal intervals, it is possible that the subthreshold signals are represented in some form within motion processing areas. For instance, observers can discriminate between global-flow displays and stimuli with up to nine signal directions covering the same range, despite being unable to detect individual directions within these displays (Williams, Tweten, & Sekuler, 1991). Given the randomised directions of our stimuli, the vector average would not have given any consistent uni-directional percept. However, it is possible that the transparency limit represents a shift from the segmentation of multiple global-motion signals to their integration, as opposed to the stimuli being represented in a manner similar to random noise.

Results from the mixed-disparity conditions of Experiment 2 demonstrate that this limit can be extended. By distributing signals across two independent disparity-tuned systems, signals processed by one system will no longer reduce the intensity of those processed by the other. With fewer signals processed by each system, the effective intensities would rise. This elevation in the effective signal intensity allowed observers to detect up to three signals, with a magnitude of improvement identical to that obtained previously using the speed-tuned global-motion systems (Greenwood & Edwards, 2006).

As well as the extension of the transparency limit, many aspects of transparent-motion processing are facilitated by binocular disparity differences. As discussed earlier, detection thresholds for bi-directional transparent motion are lowered when each signal moves on a distinct depth plane (Hibbard & Bradshaw, 1999). In addition, binocular disparity can restore transparent-motion perception in locally-balanced displays (Qian et al., 1994), improve the detection of specific directions within transparent-motion stimuli (Treue, Hol, & Rauber, 2000), and increase the likelihood of seeing transparency versus coherence in plaid stimuli (Trueswell & Hayhoe, 1993; von Grünau, Dubé, & Kwas, 1993). These interactions also work in the opposite direction, since the addition of transparent motion to cluttered stereo images can facilitate binocular matching (van Ee & Anderson, 2001).

The influence of depth on transparent-motion processing is mirrored in the responses of cells in area V5/MT. When transparent-motion stimuli are presented to these cells, their output is reduced compared with the response to uni-directional motion (Snowden et al., 1991). The responses of a substantial proportion of cells are restored to uni-directional levels when the secondary transparent-motion signal is presented at a distinct binocular disparity (Bradley, Qian, & Andersen, 1995).

This reduced response to transparent-motion stimuli has been taken to suggest inhibitory connections between global-motion units (Snowden et al., 1991), with the addition of disparity differences purported to reduce or eliminate these inhibitory interactions (Bradley et al., 1995). However, as discussed previously (Edwards & Nishida, 1999; Greenwood & Edwards, 2006), this difference in response may relate to the differences in signal intensity between the uni-directional and transparent-motion stimuli used, given the sensitivity of these cells to the signal-to-noise ratio of global-motion stimuli (Britten et al., 1993; Rees, Friston, & Koch, 2000). The increased responsiveness of V5/MT cells that follows the addition of disparity differences may therefore be the result of increasing the effective global-motion signal intensity in the preferred direction of the cell.

Competitive global-motion inhibition has also been proposed to explain the many costs associated with the detection of transparent motion. In our experiments, the high signal-detection thresholds for transparency could certainly be explained in this manner. By distributing the signals across two depth planes, this inhibition would be reduced, elevating the effective signal intensities and improving the detection of multiple transparent-motion signals. However, this account is complicated by the observation that the high signal-detection thresholds for transparent motion appear to be task specific. That is, when only one signal within stimuli needs to be detected, thresholds are identical to those for uni-directional motion (Edwards & Nishida, 1999).

This task dependence suggests that the costs associated with transparency may relate to the demands placed upon attention. Previous work is consistent with this notion. For instance, although the precision of direction judgements is lower for transparent-motion signals than for uni-directional signals, similar costs are observed with multiple spatially segregated global-motion signals (Braddick et al., 2002). Similarly, observers are able to detect individual pre-cued directions within transparent-motion displays containing up to six signal directions, but fail to detect these signals in the absence of such cues (Felisberti & Zanker, 2005). The act of dividing attention between multiple global-motion signals may thus be the cause of the high signal-detection thresholds for transparency. However, regardless of the underlying causes of these thresholds, the interpretation of our current results remains the same—the addition of disparity differences can elevate three signals above the high thresholds for the simultaneous detection of transparent motion.

Though the global-motion stage is the logical starting point in considering the locus of the transparent-motion limit, we cannot rule out the possibility that the limit results from processing within a higher-order stage receiving global-motion input. For instance, cells in MST show disparity selectivity (Roy, Komatsu, & Wurtz, 1992), with further evidence suggesting the existence of at least two independent disparity-tuned systems sensitive to optic flow (Khuu, Li, & Hayes, 2006). These systems may have influenced our results, particularly given the similarities between transparent motion and optic flow patterns such as motion parallax. However, the precise role of optic flow detectors in transparent-motion detection is not presently clear.

4.2. A higher-order limit for transparent-motion detection

Although disparity differences allowed an extension of the transparent-motion limit of two, our results may be indicative of an additional limit on transparency detection. Namely, the same-disparity conditions of Experiment 2 demonstrate that each disparity-tuned system in isolation was capable of detecting two signals simultaneously. With both of these systems operating together in mixed-disparity conditions, up to four signals should have been detected. However, observers consistently failed to detect more than three transparent-motion signals. This was also found in our extension of the limit using the speed-tuned global-motion systems (Greenwood & Edwards, 2006).

Rather than an additional limitation on transparent-motion processing, it could be that the elevation in signal intensities was insufficient to support the detection of four signals. This could have been the case if thresholds for the detection of four signals are higher than those for the detection of three. Such a possibility is consistent with the rise in thresholds from 15% for uni-directional motion to 40% for each transparent-motion signal within bi-directional displays (Edwards & Greenwood, 2005). If this threshold elevation were to continue for three signals, thresholds would have to be below 50% to allow the extension of the transparent-motion limit that was obtained. However, it is possible that the detection of four signals requires signal intensities higher than 50%, which our manipulations of speed and depth did not exceed.

Because both speed and binocular disparity enable an increase in the effective signal intensity, further increases may be possible by combining the two stimulus features. However, at least some independence would be required between these global-motion systems for this to occur. Recent evidence suggests that this may be the case. Namely, although low-speed dots ordinarily alter the perceived speed of mixed-speed stimuli by lowering the vector average, this influence is mitigated by dividing the dots between distinct depth planes (Khuu et al., 2006). Speed processing may thus occur independently within distinct disparity-tuned systems. By dividing transparent-motion signals between both the speed- and disparity-tuned systems, it may then be possible to achieve further increases in signal intensity. With four signals, each within one of these four systems, the effective intensity of each would be 100%. If there are higher thresholds for four signals than for three, this should allow further extension of the transparent-motion limit. However, an additional higher-order limit would restrict any extension of the limit to three.

Because of the large number of trials required to determine speed and disparity values that are independently processed, only J.A.G. was tested. Transparent-motion signals were presented at either the near or far depth planes, with one of two speeds. Slow dots moved at 1.4°/s (0.05° per frame), while fast dots moved at 9.5°/s (0.32° per frame). We have previously shown these speeds to be processed independently for J.A.G. (Greenwood & Edwards, 2006). The remaining stimulus details were identical to those of Experiment 2.

Given the focus on further extension of the transparency limit, only 3 vs. 4 and 4 vs. 5 comparisons were included. For stimuli with three or four signals, no more than one signal was presented to each of the four systems. With five signals, two signals would be presented to one of the systems. Thus, for each of the two signal comparisons, there were four stimulus configurations: *majority-slow* signals with a majority of signals on the near depth plane (*majority near*), *majority slow* with *majority far*, and likewise for the two *majority-fast* conditions.

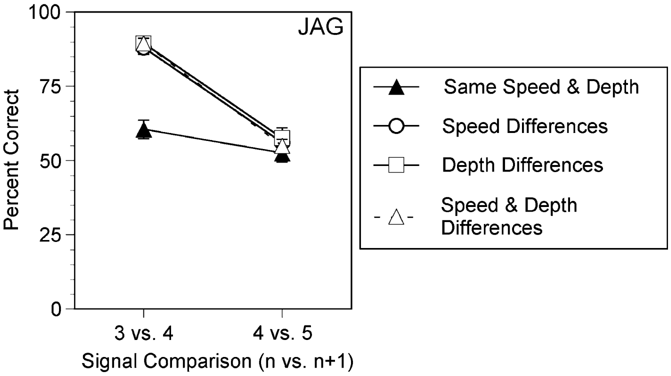

Fig. 3. Data from the follow-up experiment, displaying the mean percent-correct scores for transparent-motion detection. Only the 3 vs. 4 and 4 vs. 5 signal comparisons were conducted. Performance was uniformly poor when signals moved with the same speed and within the same depth plane (filled triangles, derived from the *all near* and *all far* conditions of Experiment 2). Speed differences allow significant improvement on the 3 vs. 4 comparisons (open circles, data averaged across the *majority-slow* and *majority-fast* conditions of Experiment 2 in Greenwood and Edwards (2006)), as do depth differences (open squares, derived from the *majority-near* and *majority-far* conditions of Experiment 2). No further improvement was obtained when both speed and depth differences were used (open triangles, averaged across four stimulus configurations as described in text). Error bars represent 1 *SEM*.

Percent-correct scores from these signal comparisons are presented in Fig. 3. Scores from the four stimulus configurations did not differ significantly, and are presented as a combined average. We have also included the same-disparity and mixed-disparity results from Experiment 2, as well as the mixed-speed results from our previous study (Greenwood & Edwards, 2006) averaged across the different stimulus configurations. Performance with speed and depth differences combined is closely matched with that arising from speed or depth differences alone. With the combination of speed and depth differences, the 3 vs. 4 comparisons were performed at around 90% accuracy. Performance drops to chance for the 4 vs. 5 comparisons, which demonstrates that three signals could be detected, but not four.

When compared with stimuli containing differences in direction only, this is a significant improvement in performance. However, the additional signal intensity elevations did not allow the detection of more than three signals, which gives further evidence for an additional transparent-motion limit of three.

4.3. The basis of the higher-order limit

The transparent-motion limit of two appears to have its origins in global-motion processing. However, the additional limit of three was insensitive to our manipulations of signal intensity, which suggests that it may arise through higher-order processing. The most likely explanation is that this limitation represents a restriction in the number of objects that can be attended to simultaneously. However, research into the limits of object detection and attention typically result in slightly higher capacities than those observed in the present study.

Attentional bottlenecks have been well characterised through processes such as subitizing, which refers to the number of objects that can be enumerated simultaneously without counting (Kaufman, Lord, Reese, & Volkmann, 1949). Up to four items (typically static dots on a uniform background) can be detected in this way, with higher numbers of objects detected with lower accuracy and confidence (Atkinson, Campbell, & Francis, 1976). With respect to moving stimuli, Pylyshyn and Storm (1988) report that up to four or five objects can be tracked attentively within an array of ten distractor objects. This was possible even when serial scanning was minimised through increased dot density and rapid speeds, suggesting that the tracking could occur in parallel. However, this does not represent a fixed upper limit for object tracking. When both target and distractor items are distributed across two depth planes, observers can track up to six or seven items (Viswanathan & Mingolla, 2002).

These processes predict much higher performance in our present task than what was obtained. However, part of this discrepancy may relate to the marked differences in stimuli. Both subitizing and multiple object tracking studies make use of spatially distinct, uncluttered stimuli. Transparent-motion stimuli, by necessity, contain a large number of spatially intermingled dots with global-motion signals that overlap in space. Consistent with this notion, decreasing stimulus size and increasing dot density in multiple object tracking tasks results in significantly poorer performance (Intriligator & Cavanagh, 2001). Performance is again degraded when the trajectories of objects are allowed to overlap (Viswanathan & Mingolla, 2002). Thus, the higher-order limit of three may pertain to attentional limitations within dense, spatially overlapping stimuli.

Alternatively, the higher-order limit may relate to the number of objects that can be manipulated simultaneously within visual working memory. A wide variety of tasks suggest that working memory has a capacity between three and five items, with an overall average of four (Cowan, 2001). Our experiments could be included within the lower end of this spectrum, especially when it is noted that the capacity of visual working memory depends on the complexity of component objects. Complex items, such as random polygons, give a working memory capacity as low as two, while simpler items such as colour patches allow a much higher level of performance (Alvarez & Cavanagh, 2004). Transparent-motion detection is certainly a complex task, as indicated by the many costs associated with transparency when compared to uni-directional motion (e.g., Greenwood & Edwards, 2006; Wallace & Mamassian, 2003). This complexity may lead to the restricted capacity of three signals that can be stored simultaneously within working memory.

Overall, there appear to be two limitations on the detection of transparent-motion signals. When direction is the sole basis for transparency, the first limit restricts the number of transparent-motion signals that can be detected simultaneously to two. This limit reflects high signal-to-noise detection thresholds for transparent motion, and can thus be extended by increasing the effective signal intensity of component transparent-motion signals. There also appears to be a second limit on transparent motion, restricting extensions to no more than three signals. This limit is more rigid than the limit of two and may reflect an absolute bottleneck in either attentional capacity or visual working memory. The precise nature of this higher-order limit remains to be investigated.

Acknowledgments

This work was supported by an Australian Research Council Grant S6505064. An Australian Postgraduate Award also supported the first author.

References

Akerstrom, R. A., & Todd, J. T. (1988). The perception of stereoscopic transparency. *Perception & Psychophysics, 44*, 421–432.

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science, 15*, 106–111.

Atkinson, J., Campbell, F. W., & Francis, M. R. (1976). The magic number 4 ± 0: a new look at visual numerosity judgements. *Perception, 5*, 327–334.

Ball, K., & Sekuler, R. (1980). Models of stimulus uncertainty in motion perception. *Psychological Review, 87*, 435–469.

Braddick, O. J. (1995). Seeing motion signals in noise. *Current Biology, 5*, 7–9.

Braddick, O. J., Wishart, K. A., & Curran, W. (2002). Directional performance in motion transparency. *Vision Research, 42*, 1237–1248.

Bradley, D. C., Qian, N., & Andersen, R. A. (1995). Integration of motion and stereopsis in middle temporal cortical area of macaques. *Nature, 373*, 609–611.

Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1993). Responses of neurons in macaque MT to stochastic motion signals. *Visual Neuroscience, 10*, 1157–1169.

Clarke, P. G. H. (1977). Subjective standstill caused by the interaction of moving patterns. *Vision Research, 17*, 1243.

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87–185.

Edwards, M., & Badcock, D. R. (1994). Global-motion perception: Interaction of the ON and OFF pathways. *Vision Research, 34*, 2849–2858.

Edwards, M., Badcock, D. R., & Smith, A. T. (1998). Independent speed-tuned global-motion systems. *Vision Research, 38*, 1573–1580.

Edwards, M., & Greenwood, J. A. (2005). The perception of motion transparency: A signal-to-noise limit. *Vision Research, 45*, 1877–1884.

Edwards, M., & Nishida, S. (1999). Global-motion detection with transparent-motion signals. *Vision Research, 39*, 2239–2249.

Felisberti, F. M., & Zanker, J. M. (2005). Attention modulates perception of transparent motion. *Vision Research, 45*, 2587–2599.

Greenwood, J. A., & Edwards, M. (2006). An extension of the transparent-motion detection limit using speed-tuned global-motion systems. *Vision Research, 46*, 1440–1449.

Hibbard, P. B., & Bradshaw, M. F. (1999). Does binocular disparity facilitate the detection of transparent motion? *Perception, 28*, 183–191.

Hibbard, P. B., Bradshaw, M. F., & De Bruyn, B. (1999). Global motion processing is not tuned for binocular disparity. *Vision Research, 39*, 961–974.

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

Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkmann, J. (1949). The discrimination of visual number. *American Journal of Psychology, 62*, 498–525.

Khuu, S. K., Li, W. O., & Hayes, A. (2006). Global speed averaging is tuned for binocular disparity. *Vision Research, 46*, 407–416.

Lagae, L., Raiguel, S., & Orban, G. A. (1993). Speed and direction selectivity of macaque middle temporal neurons. *Journal of Neurophysiology, 69*, 19–39.

Lehmkuhle, S., & Fox, R. (1980). Effect of depth separation on metacontrast masking. *Journal of Experimental Psychology: Human Perception and Performance, 6*, 605–621.

Levitt, H. (1971). Transformed up–down methods in psychoacoustics. *Journal of the Acoustical Society of America, 49*, 467–477.

Maunsell, J. H. R., & Van Essen, D. C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity. *Journal of Neurophysiology, 49*, 1148–1167.

Mulligan, J. B. (1992). Motion transparency is restricted to two planes [Abstract]. *Investigative Ophthalmology and Visual Science, 33*, 1049.

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision, 3*, 179–197.

Qian, N., Andersen, R. A., & Adelson, E. H. (1994). Transparent motion perception as detection of unbalanced motion signals. I. Psychophysics. *Journal of Neuroscience, 14*, 7357–7366.

Raymond, J. E. (1993). Movement direction analysers: Independence and bandwidth. *Vision Research, 33*, 767–775.

Rees, G., Friston, K., & Koch, C. (2000). A direct quantitative relationship between the functional properties of human and macaque V5. *Nature Neuroscience, 3*, 716–723.

Roy, J. P., Komatsu, H., & Wurtz, R. H. (1992). Disparity sensitivity of neurons in monkey extrastriate area MST. *Journal of Neuroscience, 12*, 2478–2492.

Smith, A. T., Curran, W., & Braddick, O. J. (1999). What motion distributions yield global transparency and spatial segmentation? *Vision Research, 39*, 1121–1132.

Snowden, R. J. (1989). Motions in orthogonal directions are mutually suppressive. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision, 6*, 1096–1101.

Snowden, R. J., & Rossiter, M. C. (1999). Stereoscopic depth cues can segment motion information. *Perception, 28*, 193–201.

Snowden, R. J., Treue, S., Erickson, R. G., & Andersen, R. A. (1991). The response of area MT and V1 neurons to transparent motion. *Journal of Neuroscience, 11*, 2768–2785.

Treue, S., Hol, K., & Rauber, H. (2000). Seeing multiple directions of motion—physiology and psychophysics. *Nature Neuroscience, 3*, 270–276.

Trueswell, J. C., & Hayhoe, M. M. (1993). Surface segmentation mechanisms and motion perception. *Vision Research, 33*, 313–328.

van Boxtel, J. J. A., & Erkelens, C. J. (2005). Global-motion perception is governed by a single motion system [Abstract]. *Perception, 34*, 178.

van Ee, R., & Anderson, B. L. (2001). Motion direction, speed and orientation in binocular matching. *Nature, 410*, 690–694.

Viswanathan, L., & Mingolla, E. (2002). Dynamics of attention in depth: Evidence from multi-element tracking. *Perception, 31*, 1415–1437.

von Grünau, M., Dubé, S., & Kwas, M. (1993). The effect of disparity on motion coherence. *Spatial Vision, 7*, 227–241.

Wallace, J. M., & Mamassian, P. (2003). The efficiency of speed discrimination for coherent and transparent motion. *Vision Research, 43*, 2795–2810.

Wallace, J. M., & Mamassian, P. (2004). The efficiency of depth discrimination for non-transparent and transparent stereoscopic surfaces. *Vision Research, 44*, 2253–2267.

Watamaniuk, S. N. J., Flinn, J., & Stohr, R. E. (2003). Segregation from direction differences in dynamic random-dot stimuli. *Vision Research, 43*, 171–180.

Williams, D., Phillips, G., & Sekuler, R. (1986). Hysteresis in the perception of motion direction as evidence for neural cooperativity. *Nature, 324*, 235–255.

Williams, D., & Sekuler, R. (1984). Coherent global motion percepts from stochastic local motions. *Vision Research, 24*, 55–62.

Williams, D., Tweten, S., & Sekuler, R. (1991). Using metamers to explore motion perception. *Vision Research, 31*, 275–286.