



City Research Online

City, University of London Institutional Repository

Citation: Freeman, E. D. & Driver, J. (2008). Voluntary control of long-range motion integration via selective attention to context. *Journal of Vision*, 8(11), 18. doi: 10.1167/8.11.18

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/13404/>

Link to published version: <https://doi.org/10.1167/8.11.18>

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Voluntary control of long-range motion integration via selective attention to context

Elliot Freeman (1,2)

Jon Driver (2)

1. Department of Psychology, Brunel University.
2. UCL Institute of Cognitive Neuroscience,
University College London

Correspondence to:

Dr E. Freeman,
University of Brunel,
Uxbridge, UB8 3PH, UK.

Tel : +44 (0)18952 66106

E-mail: elliott.freeman@brunel.ac.uk

Abstract

Ambiguous stimuli can look different in different contexts. Here we demonstrate that subjective appearance of motion depends not only on current visual input but critically on which aspects of the context are attended. Observers fixated a central oblique Test grating flanked by two pairs of orthogonally-oriented Context gratings arranged in a cross (+) configuration. Each context pair could induce the Test stimulus to appear to switch from diagonal motion to either horizontal motion (due to one context pair) or vertical motion (due to the other). Spontaneous switching between these motion states was observed under free-viewing. We demonstrate that observers can voluntarily select between specific states, when cued to attend selectively to one or other context pair in an alternating manner. Concurrent reports of perceived Test stimulus motion depended specifically on which context was currently attended, indicating a high degree of ‘cued-control’ over subjective state via attended context. Further experiments established that the perception was nevertheless still constrained by physical stimulus context, as well as by attentional selection among that context. Moreover, the attentional control evident here did not seem reducible solely to local contrast gain modulation of the attended vs. ignored context elements. Selective attention to different parts of the context can evidently resolve the ambiguity of the Test grating, with integration arising selectively for those components that are jointly attended. Such selective integration can result in substantial voluntarily-controlled changes in phenomenal perception.

Introduction

Our subjective experience of a visual stimulus often depends critically on the context in which we encounter it (e.g. see Figure 1 and further description below). Since early Gestalt studies much research has focused on influences from current *stimulus context* on perception of local static or moving stimuli (e.g. Wallach, 1935/Wuerger, Shapley, & Rubin, [1996](#); Polat & Sagi, [1993](#); Castet & Zanker, [1999b](#); Kim & Wilson, [1997](#); Lorenceau & Zago, [1999](#)), and on corresponding neural responses (Gilbert & Wiesel, [1990](#); Kapadia, Ito, Gilbert & Westheimer, [1995](#); Polat, Mizobe, Pettet, Kasamatsu & Norcia, [1998](#); Duncan, Albright, & Stoner, [2000](#); Huang, Albright, & Stoner, [2007](#)). However there has been mounting evidence that the current *behavioural context* (e.g. including attentional set or task goals) may also play a critical role, affecting both subjective appearance and objective perceptual thresholds for contrast or motion (Braun, [2002](#); Yeshurun & Carrasco, [1998](#); Lee, Itti, Koch et al., [1999](#); Carrasco, Ling, & Read, [2004](#); Raymond, O'Donnell, & Tipper, [1998](#)), plus neural responses in visual cortex (Motter, [1993](#); Moran & Desimone, [1985](#); Luck, Chelazzi, Hillyard & Desimone, [1997](#); Gandhi, Heeger, & Boynton, [1999](#); Reynolds & Chelazzi, [2004](#); Treue, [2001](#)). More recent evidence from electrophysiology (Khoe, Freeman, Woldorff & Mangun, [2006](#); Ito & Gilbert, [1999](#); Casco, Grieco, Campana, Corvino & Caputo, [2005](#)) and psychophysics (Freeman, Sagi, & Driver, [2001](#); Freeman, Driver, Sagi & Zhaoping [2003](#)) suggests that attentional factors may also interact with the current stimulus context, selectively modulating the impact of surrounding context upon a local stimulus. For example, while contrast detection and discrimination thresholds for a static Gabor patch are found to be lower for a static target Gabor patch when presented in the context of collinear patches that together make a continuous global contour (Polat et al., [1993](#)), recent work indicates that this contextual influence arises only when those collinear context patches are selectively attended in preference to other non-collinear patches present in the scene (Freeman et al., [2001](#); Khoe et al., [2006](#)).

Although such attention-to-context effects on objective performance with static Gabor patches were robust in this recent work, they only allowed measurement of relatively subtle effects around contrast threshold, which might go largely unnoticed by an observer, rather than producing

any obvious phenomenal changes in the appearance of the target stimulus. In the present work we introduce a complementary paradigm, which assesses the impact of selective attention to one or other aspect of surrounding context, but now for phenomenal perception of motion, in dynamic displays. To anticipate, we find that attention-to-context can induce striking effects on the appearance of otherwise ambiguous local motion.

Consider the classic aperture motion stimulus depicted in Figure 1a, in which an oblique grating may appear to drift diagonally behind a circular window (e.g. Marr & Ullman, [1981](#); Wallach, 1935/Wuerger et al., [1996](#)). The appearance of this stimulus is highly context-sensitive: for example, adding a pair of orthogonally-oriented gratings is sufficient to induce a switch in perceived motion-direction from diagonal to either horizontal or vertical (compare Figure 1a with 1b and 1c; see also Alais, van der Smagt, van den Berg & van de Grind, [1998](#); Lorenceau & Shiffrar, [1992](#)). In Figure 1d, which depicts the new situation we studied here, there are now two such sets of contextual gratings, each inducing orthogonal global motion percepts (see also animation and c.f. Castet & Zanker, [1999b](#), for a related example). Either of the two pairs of peripheral context-gratings in the example of Fig 1d could in principle be integrated with the central grating, but each would potentially result in different, incompatible global motions (i.e. horizontal versus vertical). This situation may exemplify other more complex situations, in which available local motion samples may each belong to different objects present in the same scene, either moving in similar or dissimilar directions (Albright & Stoner, [1995](#), e.g. see their Fig. 4). A veridical percept may only be achieved if the brain integrates those disparate local motion samples that truly belong to the same source or object (Grossberg, Mingolla, & Viswanathan, [2001](#)), thus requiring selective integration. More generally, when ambiguous stimuli appears in the context of several potentially disambiguating stimuli, selective attention to one or other aspect of the surrounding context may be critical in determining which perception dominates, as we seek to establish here.

> FIGURE 1 (see also attached animation file expt1demo.mov) <

Most work to date on this issue of selective contextual integration, at least for motion, has focused on stimulus-based segmentation cues, rather than attentional factors. For example, binocular disparity or occlusion relations that indicate the 3D arrangement of surfaces can strongly determine which samples of local motion are perceived as belonging to the same surface, and thereby which get integrated into a global-motion percept (e.g. see Grossberg et al., [2001](#); Nakayama, Shimojo, & Silverman, [1989](#)). Neural correlates of such motion integration have now been measured in awake monkeys (e.g. Duncan et al., [2000](#)), even without any explicit task-set for perceiving one interpretation or another. Stimulus-driven global-motion phenomena can arise spontaneously in human observers, without requiring any prior exposure to fully disambiguated examples (e.g. Lorenceau & Alais, [2001](#)). Such observations might be taken to support the traditional notion that early perceptual processing processes is primarily stimulus-driven and perhaps preattentive.

In contrast, there is some evidence that manipulation of task instruction can influence the reported directions of motion for a single ambiguous ‘barber-pole’ grating (Castet, Charton, & Dufour, [1999a](#)), or can affect motion-coherence thresholds for a previously attended direction (Raymond et al., [1998](#)). Attention may also modulate the neural correlates of compound plaid-motion versus component-motion in human MT+ (Castelo-Branco, Formisano, Backes, Neuenschwander, Singer & Goebel, [2002](#)). Furthermore, observers can reportedly amplify a sub-threshold motion percept (Cavanagh, [1992](#)), or even voluntarily influence the perceived rotation of an ambiguous motion display, by attentionally tracking the motion of features in the desired direction (Lu & Sperling, [1995b](#); Lu & Sperling, [1995a](#); Verstraten, Cavanagh, & Labianca, [2000](#)). Such results could be taken to indicate that attention may have some general influences over how motion is perceived. But what they do not reveal as yet is whether top-down factors can play a specific role in modulating selective integration of local contextual motion components into global motion, via attention to one or other aspect of surrounding context, as addressed here.

The present work tests whether voluntary attention (to one or other aspect of surrounding context) is

sufficient to determine which of several conflicting contexts will be integrated with a local central stimulus. Given a context-sensitive local stimulus such as the motion aperture described above (Fig 1), such attentionally-determined selective contextual integration might result in a rather gross change in subjective appearance for the central stimulus (unlike the rather subtle contrast-threshold effects studied for static Gabors by Freeman et al., [2001](#), [2003](#)). Moreover, by attending to one or other aspect of the surrounding context here (see Fig 1d), observers should thereby have some control over their perception for the multistable central stimulus, thus producing a new way to switch voluntarily between different subjective states for that stimulus, simply by attending to one or other aspect of its surrounding context.

Many past studies have sought to measure possible voluntary control over subjective perception (Leopold & Logothetis, [1999](#); Long & Toppino, [2004](#)), while using multistable stimuli. But none to our knowledge has done so by means of cuing attention to one or other aspect of the surrounding context, as here. In one popular paradigm observers are instructed to try to ‘speed up’ or ‘slow down’ ongoing fluctuations in perception that are otherwise spontaneous (van Ee, van Dam, & Brouwer, [2005](#)), but without emphasis on switching to or maintaining one particular state versus another (e.g. see criticism by Meng & Tong, [2004](#)). In another approach, observers are requested to ‘hold’ a pre-specified state for as long as possible for a stimulus that otherwise spontaneously changes in appearance (Meng & Tong, [2004](#)). None of these previous approaches (which have been utilized in many prior studies of possible voluntary control over multistable percepts, and have typically revealed rather little such control, if any) have specifically examined whether selective attention to one or other aspect of surrounding context can provide a powerful mechanism for control over otherwise multistable perception. Our new approach might therefore be of some potential general utility, for research on possible control over fluctuating subjective perceptions, in addition to addressing the more specific issues about motion integration as described below.

Our new paradigm used a multi-stable motion display as introduced above (see Figure 1d and

animation), comprising a single central ‘Test’ grating, surrounded by two pairs of ‘Context’ gratings, arranged on separate vertical and horizontal axes, to form a ‘cross’ (+) configuration. The left and right gratings can appear to form an integrated chevron-pattern or ‘zig-zag’ together with the central pattern, that all appear to drift vertically together (see blue arrow in Fig. 1c and 1d). Alternatively, the upper and lower gratings can appear to form an integrated zig-zag pattern with the central grating, which then appear to drift all horizontally together (see yellow arrow in Fig 1b and 1d). Note that either interpretation can apply to the common central grating, but this creates a potential conflict because the central component cannot be seen to move in both orthogonal directions (vertically and horizontally) at once. The result, under free viewing without attentional constraints, is a vigorously multistable display, which under passive viewing switches spontaneously between (primarily) horizontal and vertical global motion; see radial histograms of perceived motion-directions under free-viewing in bottom row of Figure 1 (e-h), as explained further in the Methods below (see also Fig 1 legend).

This new ‘dual-axis’ global motion display was designed by analogy to the static Gabor displays used in our previous work on attentional modulation of contextual integration (Freeman et al., [2001](#), [2003](#)). Here, we assessed whether simply directing the observer’s covert selective attention to one or other context (vertical or horizontal pair of flankers) would lead to a corresponding phenomenal switch between motion-direction percepts for the central moving grating. To manipulate attention, we cued the pair of Context gratings on either the horizontal or vertical axis, while instructing subjects to report the perceived direction of central grating motion (in the context of the possible ‘zigzag’ patterns formed with the flanker gratings).

The attentional cues alternated predictably between the horizontal or vertical pair of Context stimuli, every six seconds. In contrast with the classical ‘hold’ paradigm described above (e.g. Meng & Tong, [2004](#)), which seeks to measure how long an observer can hold a given state before they are interrupted by an *involuntary* switch, our goal here was to measure the possible *voluntary*

switching behaviour that might be enabled by selective attention to one of other aspect of the surrounding context (i.e. by attentional cues to the ‘columnar’ or ‘row’ flanking context, see Fig 1d). The sequence of cuing-epochs here was by design made entirely regular and predictable for the observer, in order to maximise the opportunity for voluntary control, via systematic attention to one or other aspect of the context.

Observers were required to indicate continually the direction in which the central Test grating currently appeared to drift (by smoothly adjusting a pair of peripheral cursor-dots to indicate the drift angle, using a computer mouse). With the stimulus configuration illustrated in Fig. 1d, perception of horizontal motion would be indicated in Experiment 1 by positioning the cursor dots along the horizontal axis, i.e. orthogonal to the cued axis when the vertical pair of flankers was cued, with these flankers being expected to induce a percept of horizontal motion for the central Target (though in later experiments this mapping of cued-axis with respect to expected-central-response could vary unexpectedly). We predicted that if observers can voluntarily modulate integration of the Context gratings with the Test, by selectively attending to one or other Context pair, then the reported direction of perceived Test drift should switch in temporal correlation with the attention cues for the Context.

Using this new paradigm, we found a robust ability for voluntary control of motion integration and consequently phenomenal perception in all observers, via selective attention. To our knowledge this provides the first direct demonstration for the role of selective attention to context, in perceptual integration of local motion to produce global-motion percepts. Further control experiments confirmed that these effects reflected genuine changes in perception rather than demand characteristics, and identified stimulus-based constraints on these phenomena that constrain the possible mechanisms involved.

General Methods

Subjects

8 male and female subjects aged between 21 and 35 each took part in one or more of the experiments, as described below. All reported normal or corrected acuity. All participated with informed consent and were paid for their time. Each experiment had three to five observers, apart from the free-viewing situation (see below) in which seven observers participated as part of their familiarization with the displays and response options.

Stimuli and apparatus

Stimuli were presented on a CRT display (19" Mitsubishi Diamond Pro CRT display for Expt. 1, and 21" Sony GDM-F520 CRT for Expts. 2&3), viewed from a distance of 1 metre in a darkened room. Video mode was 1600x1200, with a screen refresh rate of 75hz. In Expts. 1 and 2, displays were linearized using 8-bit software gamma-transformation. In Expt. 3, a BITS++ digital video processor (Cambridge Research Systems) provided 14-bit grey-level resolution (in MONO++ mode), allowing optimal rendering for both the low and high-contrast gratings. In all cases, mid-grey luminance was 40cdm^{-2} . Stimulus control was provided by a PC running Psychophysics Toolbox (Brainard, [1997](#)) under Matlab 6.1. Responses were made using a mouse via the standard PC keyboard, see below

Each local motion stimulus was composed of a gray-level-modulated grating of wavelength 0.75 degrees of visual angle. In experiments 1 and 2, this was sharply windowed by a circular aperture with diameter 4.59 degrees. In experiment 3, the grating had a Gaussian window with standard deviation of 0.75 degrees. Gratings drifted with a maximum temporal frequency of 1.5 wavelengths per second (i.e. 1.125 degrees per second). By default, grating contrast was set to 100%, though this was modulated in Experiment 3. Displays typically comprised five such gratings in a cross (+) configuration of intersecting vertical and horizontal axes (see Figure 1d), comprising: one circular grating centred at fixation; one pair along the horizontal axis, left and right of the central grating;

and the other pair along the vertical axis, above and below the central grating. Centre-to-centre grating separation was 6.29° . The central grating always had an orientation of 45° anti-clockwise from vertical. The four Context gratings always had the same orientation as each other, orthogonal to the central testing (e.g. 45° clockwise from vertical), but pairs on either axis usually drifted in *opposite* directions. For example, in Fig. 1d, diagonal rightward and upward motion of the centre grating (as indicated for it with red arrow) could be integrated with the horizontally flanking gratings (drifting leftward and upward) to result in the appearance of a chevron (zigzag) pattern with upward global motion (see blue arrow in Fig 1d). By contrast, integrating the same central grating with its vertical flankers instead (drifting downward and rightward) should result in the appearance of rightward global motion instead (orange arrow in Fig 1d). In such an ‘Orthogonal’ display the two potentially available directions of global motion were thus both orthogonal to the orientation of the particular inducing axis (horizontal or vertical flankers) across which local motion could be integrated.

Alternatively, simply reversing the direction of all the flankers to generate an alternative ‘Aligned’ stimulus could result in the appearance of a horizontally sliding pattern when integrating the central grating with the horizontal flankers, or a vertically sliding pattern when integrating with the vertical flankers. In this case, the perceived global motion would then be *aligned* with, rather than orthogonal to, the inducing axis. Both patterns were compared in Experiment 3. A third control pattern was tested in Experiment 2 (see below), where global motion was now aligned with one axis but orthogonal with the other, so that the resulting direction of global motion was the same regardless of which pair of context flankers were integrated. In Experiments 1 and 2, all gratings drifted smoothly at the same temporal frequency of 1.5 wavelengths per second (i.e. 1.125 degrees per second). All gratings reversed drift direction between experimental runs in order to refresh any adaptation state. In Experiment 3, the motion was oscillatory rather than unidirectional, with motion speed and direction modulated by a sinusoid with frequency of 0.6 cycles per second, thus minimising any directional aftereffects within each experimental run. Note that, in any case, any

adaptation or aftereffects should be unrelated to the critical attentional-cuing manipulation here.

Superimposed on the central grating was a small white fixation spot (diameter 0.24°), surrounded by a larger blank disk of background luminance (diameter 0.97°), which masked-off motion around the foveal region to reduce any tendency for involuntary eye-movements (note also that formal eye-tracking was implemented in Experiment 3, see below). In Experiments 1 and 2, identical white cue spots were also displayed at the centres of either the two Context gratings in the horizontal pair, or those in the vertical pair, to provide the attentional cues. In Experiment 3, these cue dots were placed in a row or column configuration each flanking the fixation point at a distance at 0.5° , within the central fixation area (so that these rather ‘central’ attentional cues should act pure as instructions for endogenous top-down attention, rather than as peripheral attention-attracting events at the same location as one pair of context gratings)

There were also two cursor spots, displayed on opposite sides of the fixation point at an eccentricity of 9.65° . These were yoked to rotate together in opposite directions around a virtual circle under the smooth control of a computer mouse. Observers used this mouse to make smooth adjustments, in real-time, to the radial angle of these dots with respect to the centre, to indicate the direction in which the central grating currently appeared to be drifting, in the context of the attentionally-cued gratings. Users quickly learned the calibration between mouse and cursor position so could select positions rather quickly and accurately (see Appendix 1 in Supplementary Materials for a sample timecourse of cursor positions during one typical experimental run). With the Orthogonal motion configuration (see Fig 1d), for example, attending the horizontal axis was expected to result in perceived vertical motion, which would be indicated by rotating the cursors to indicate top and bottom positions. Conversely, for the alternative Aligned configuration (see above), attention to the horizontal axis may now result in horizontal motion percepts, indicated by positioning the cursors to the left and right of the display. Note that the global perceived motion associated with a given attended axis (horizontal or vertical) could thus vary, either predictably in Experiment 1, or

unpredictably in Experiments 2 and 3, due to the relation of the central target with the surrounding context pairs, as described in outline above and in more detail below.

As in any experimental paradigm the range of possible psychological experiences might in principle be wider than it is feasible or relevant to measure. Here for example, our cursor-positioning instructions assume that motion percepts were frontoparallel translations (as we ourselves observed for these displays, and as our observers reported on debriefing). Though we cannot rule out that other percepts may sometimes be experienced, orthogonal to the cursor position measure, any such unusual percepts would presumably not lead to a systematic relation between attentional-cuing and recorded cursor-position, which is the key finding here.

Design and procedure

Experimental sessions comprised a series of up to 40 runs, each of 50 seconds duration, during which the stimulus was displayed continuously. In all but the initial free-viewing condition, runs were divided into alternating cueing-epochs of 6 seconds (see below). Between runs only the central fixation point was displayed until the next run was initiated by a keypress from the observer.

In the first training session, observers were initially shown simplified displays comprising either the central target alone (as in Figure 1a), or just the horizontal or vertical axes (i.e. Figs. 1b & 1c), and for five runs of each of these display types they reported their perceptions under free-viewing (i.e. no attentional constraints or cueing). Observers were thus familiarised with the appearance of vertical and horizontal ‘zigzag’ motion respectively, and the method of indicating the perceived direction for the central target using the cursor spots. All subjects reported no difficulty in seeing the global motion of the single ‘zigzag’ pattern. They then performed a further 30 free-viewing runs with the complete five-grating displays (Fig 1d). The proportional distribution of motion-direction responses in the free-viewing session is shown averaged across seven observers, as radial histograms with 95% confidence intervals (Fig 1e-h). The presence of horizontal and vertical components in the responses clearly reflects the global motions available in the stimulus (note the

different outcomes for display types Fig 1a-c, as shown in Fig 1e-g respectively). The free-viewing data also confirm that the novel five-grating stimulus (Fig 1d) was multistable, switching spontaneously between two global motion directions (and a third local motion direction) with comparable frequencies (Fig 1h). In this free-viewing condition mean durations across observers for each percept in the dual-axis display were as follows: Horizontal 1.26s (SE 0.38), Oblique 0.90s (0.12), Vertical 0.95 (0.14). Mean (and Standard Deviation) parameters of the least-squares fitted log normal distribution of dominance durations averaged across observers were as follows: Horizontal -0.49 (1.36); Oblique -0.97 (1.57); Vertical -0.67 (1.35).

In the subsequent experimental sessions, attentional-cue dots switched between vertical and horizontal flanker-pairs in alternating *cueing-epochs*, by default every 6 seconds (i.e. much longer than the above mean dominance durations for free-viewing), beginning with the vertical axis. Observers were told that their task was to attend to the cued pair of context gratings while reporting via the cursors the currently perceived direction of drift for the *central* grating in particular. Observers were also encouraged to indicate the diagonal direction (i.e. the normal component commonly seen if the central stimulus is presented in isolation, see Fig 1a and 1e) if that ever became perceptually dominant for the central grating. Care was taken to avoid suggesting that any particular response or outcome was 'correct' (see also below, for further demonstrations that mere response-bias or demand characteristics did not determine our most critical effects). Observers were emphatically instructed to maintain central fixation (as formally confirmed by an eye-tracker in Experiment 3).

Data analysis

> FIGURE 2 <

During each run, the radial angle of the cursor was recorded continuously, along with the current binary state of the attention-cues. Angles were binned (but see also Supplementary Materials, for analysis in terms of finer bins) into one of three main perceived drift trajectories: vertical global motion (angles between $\pm 22.5^\circ$ from vertical), horizontal global motion ($\pm 22.5^\circ$ from horizontal)

and diagonal local motion ($\pm 22.5^\circ$ from the normal component trajectory). A fourth bin accepted other angles that were unlikely to be related to any perception (i.e. $\pm 22.5^\circ$ from the angle orthogonal to the normal component), but that might occasionally be generated if the observer accidentally rotated the cursors in the ‘wrong’ direction to that they intended (in practice this arose on average in only 0.02% of all trials). The distribution of responses over multiple epochs and runs in Experiment 1 is visualised in Figure 2a, as colour-coded ‘raster plots’ of perceptual report (blue for report of vertical movement for the central grating; yellow for horizontal; red for the diagonal normal component; any others as cyan) as a function of time (x-axis) and run (y-axis). Any coupling of perceptual states to attentional cueing may be assessed by noting how this compares to an analogously colour-coded representation of the attentional cue states (see top panel of Figure 2a, where ‘R’ for ‘row’ means that the horizontal pair of flankers was cued for attention, which was expected to induce vertical motion percepts with the orthogonal-configuration displays; while ‘C’ for ‘column’ means that the vertical pair of flankers was cued for attention, expected to induce horizontal motion percepts in Experiment 1). Radial histograms (e.g. see Figure 2b) were then constructed for each of the two attentional-cue states separately, using the same bins as described above, with counts now converted into the proportion of the total number of response samples obtained in each cue-state. In these proportional histograms, the value of 0.33 indicates the proportion predicted on the null hypothesis of equal distribution of responses over three principal vertical, horizontal and diagonal motion bins. This equiprobable chance value seems justifiable given that the fourth bin was used only very rarely (0.02% of trials, see above), and also that under free viewing the remaining three bins were indeed visited with equal frequency (see Fig 1h).

Quantification of ‘cued-control’ was then obtained by taking the proportion of global motion reports corresponding (as described in the Introduction) to the *cued* context in one of the cueing epochs, then subtracting the proportion of reports of the same motion direction in the other epoch, in which the cue was to the other axis. To take the case of the Orthogonal configuration (Fig 1d) as an example, if the proportion of horizontal-motion reports when cued to the vertical pair of flankers

was 1.0, but was 0.8 when given cue to the horizontal pair of flankers, this would result in a value of 0.2 cued-control (i.e. the *change* in observed proportions, due to which context pair was cued for attention).

The resulting values (reported below as averages across epochs, with a corresponding estimate of standard error) can in principle range between 1 to -1 , with high positive values indicating a strong dominance of responses in the motion-direction expected given the currently cued context (i.e. good cued control). Values close to zero indicate that there were few consistent shifts in reported motion direction between cueing epochs.

Experiment 1

This experiment measured reports of three observers for the perceived direction of motion in the central grating, when cued to attend to one or other axis of context gratings (i.e. the vertical or horizontal pair), for the Orthogonal display configuration (Fig 1d).

Results

Figure 2 illustrates the typical pattern of cue-controlled effects obtained over time, for each of the three observers. Fig. 2a shows ‘raster’ plots of responses, with time on the horizontal dimension and experimental runs on the vertical, colour-coded according to which direction of motion was being reported. The panel at the top of Figure 2a indicates, using the same colour-coding scheme, which direction of motion is predicted given the currently cued context (‘R’ for the ‘row’ of horizontal flankers, expected to induce vertical motion percepts when attended; or ‘C’ for the ‘column’ of vertical flankers, expected to induce horizontal motion percepts instead, see above). This expected direction correspondingly alternates for successive cue epochs. The phenomenal report for the central target clearly depended on the attentional-cue states for the context. For example, blue colours in the top sequence of Fig 2a indicate periods where observers were cued to attend to the horizontal pair of Context stimuli and were thus predicted to perceive vertical chevron motion of an

apparent global ‘zig-zag’ pattern (see Introduction). This did indeed lead to subjective reports in all three observers of primarily vertical motion for the central Test grating (as also coded blue in the raster plots below). Yellow in the top sequence of Fig 2a indicates instead that observers were cued to attend to the vertical pair of Context stimuli. As expected, this led to subjective reports of primarily horizontal motion for the central grating for all three observers, coded as yellow in the raster plots below. Periods where the diagonal normal component was reported for the central grating (coded red) were more common in observer RD, who tended to see vertical global motion less frequently and report oblique instead. However RD’s behaviour was nonetheless not so different from the other observers in the sense of showing some reliable ‘cued control’, as the raster plot still reveals clear dependency on the attentional cue, as reflected also in her significantly above-zero measure of cued-control. Moreover, consistency over subjects in this respect is also confirmed in our subsequent experiments here.

In these ‘raster-plot’ representations of the data, there is a notable rightward-offset of the response-state transitions relative to the cue-states. To avoid grossly underestimating cued-control, this response delay was compensated for before subsequent analysis, by searching for the temporal offset between cue onset and response that maximised our measure of cued-control. This procedure, performed for individual subjects on data pooled across conditions, is unlikely to grossly overestimate cued-control for any response sequences that were genuinely not phase-locked to the cue sequences; nor is it likely to bias the pattern of differences between conditions in any systematic way. We found that an average offset of 1600ms (s.d. 480ms) was sufficient to optimise phase-alignment of cue and perceptual-state transitions.

This delay between cue switching and response is unlikely to be attributed entirely to the mechanics of cursor adjustment, which were typically fast and smooth once initiated (e.g. see Figure 2). While there may be some internal delay associated with recognizing a subjective switch when it occurs, part of the delay may also reflect sluggish dynamics of the internal perceptual transition. This

would concur with subjective reports of observers on debriefing, who typically remarked that the perceptual switch did not occur immediately following the attentional cue but only after a delay. We note also that in another recent study from our group (Freeman & Driver, [2006](#)), where periodic veridical switches of an unambiguous structure-from-motion stimulus triggered coupled switches in a similar but fully ambiguous stimulus, there was again a surprisingly long delay (approx. 750ms) before the subjective switch was reported.

The data from Fig 2a are replotted as radial histograms in Fig. 2b, with analogous colour-coding. The circle marked with the value of .33 on these histograms signifies the ‘chance-level’ performance that would be expected if observers had randomly reported all three principal binned directions with equal frequency. The histograms indicate that, as predicted, all 3 subjects predominantly reported vertical motion of the central grating when the horizontal pair of Context gratings was cued for attention (blue), but predominantly horizontal motion when the vertical pair of flankers was cued for attention (yellow). The relationship between reports and cueing is quantified in Fig. 2c for individual observers, with 95% confidence intervals based on the standard error between measurements obtained across cueing-epochs. Significant positive values in all observers for the cued-control measure confirm that observers usually reported the direction corresponding to the global motion induced by the currently cued context, and did so significantly more often than for the other (uncued) directions.

The readiness with which observers reported global motion may appear to contrast with reports from some past studies in which other displays comprising gratings that drift behind circular windows produced only weak coherent global motion, especially at high contrast (Lorenceanu & Zago, [1999](#); Majaj, Carandini, & Movshon, [2007](#)). Several factors might account for the apparent difference to our observations. Here the greater eccentricity of our surrounding gratings may have lowered their effective contrast, while the periodic switching of attention may have continually refreshed any adaptation state to global motion. Finally, expansion/contraction percepts has

sometimes been reported informally under free viewing with stimuli comprising alternating arrays of orthogonally-oriented gratings (e.g. Lorenceau & Zago, [1999](#); Alais et al., [1998](#)). Here our use of specific attentional instructions to attend to one or other flanking axis may have reduced the likelihood of any such percepts during formal testing. In any case, the clear result from Experiment 1 is that significant ‘cued-control’ is possible under instruction to attend to one or other axis of flanking context. Our next experiments further explored the reliability of this and possible boundary conditions upon it.

> FIGURE 3 <

Experiment 2

In Experiment 1, phenomenal reports of perceived motion direction for the central target were systematically related to which pair of context gratings was cued for attention. We next addressed whether this substantial effect might somehow reflect mere demand characteristics. Perhaps a given cue (e.g. to attend the horizontal pair of context gratings) might somehow come to be considered compatible with indicating a given cursor position (i.e. alignment of the cursors for central report along the orthogonal vertical axis), regardless of what was actually perceived. Such an extreme demand characteristic seems unlikely, especially given that it would need to be orthogonal in nature. Nevertheless, the possibility merits a decisive experimental test. As a control, we therefore now tested the extent to which observers’ tendency to indicate horizontal or vertical motion depends on whether such global-motion perceptions are actually available in the stimulus. In randomly-intermingled runs, we interleaved our standard orthogonal-motion stimulus configuration (as in Experiment 1, see Fig 1d, and as also illustrated schematically in Figure 3a), for comparison with displays in which only a single direction of global motion was now available, so that our phenomenon would no longer be expected to arise if genuine. One such ‘uni-directional’ display (horizontal only) was generated by simply reversing the drift direction of the horizontal pair of contextual flankers (see Figure 3b). This eliminated all possibility of seeing a horizontally extended chevron or ‘zigzag’ drifting vertically. Instead the horizontal flankers now reinforce the perception

of horizontal motion provided by the other pair of flankers also. In a second uni-directional motion stimulus (now inducing vertical-motion only, Figure 3c), the drift direction of the vertical pair of flankers was reversed, producing a display in which all flankers are now only consistent with vertical global motion.

The present experiment deliberately maximised the opportunity for demand characteristics to influence responses, in order to determine whether such factors can indeed affect performance in the present paradigm (and if so, to what extent relative to the critical experimental findings with the bistable displays). First, the three observers (LB, LC, and MG) who participated in Experiment 2 had already performed in Experiment 1 (LB and MG) and/or pilot experiments (LC), and therefore already had ample opportunity to attune to any demand characteristics, potentially learning for instance that cueing of the vertical axis often led to reports of horizontal global motion. Second, the new uni-directional displays were interleaved randomly with the standard orthogonal-motion displays in which both global motions were available. Our observers had never seen the new (uni-directional) configurations prior to this experiment, and thus had no prior knowledge of which global motions should or should not be available for those displays. If the previously observed phenomenon of apparent switches in perceived global motion due to attentional cuing was driven solely by demand characteristics, observers should presumably continue to show similar behaviour with the new stimuli, reporting ‘percepts’ that now had no stimulus support.

Results

A clear difference in cued-control was found between the old orthogonal-motion and the new uni-directional stimulation conditions. Figure 3d displays results for each individual observer, with error bars indicating 95% confidence intervals. Importantly, values of cued-control were significantly higher when both horizontal and vertical directions of global motion were available in the stimulus (leftmost datapoints in Fig 3d, arising from the orthogonal-motion displays illustrated in Fig 1d and 3a), dropping very sharply when only one global motion direction was available (i.e.

either just horizontal as in Fig 3b, or just vertical motion, as in Fig 3c on both axes; see central and rightmost datapoints in Fig 3d respectively). The frequency of reported global motion in the standard orthogonal-motion condition was overall somewhat lower than in the first experiment. This may reflect some loss of confidence in perceptual reports with the intermixing of old and new stimulus types. But the important outcome was that observers scored significantly higher on cued-control when the stimulus objectively contained two possible directions of global motion, than in the new control stimuli which contained only one such direction.

The results so far demonstrate (for the orthogonal-motion displays, as introduced in Experiment 1, and as replicated in Experiment 2 while controlling further for potential demand characteristics) that perceived motion direction for the central Test grating was strongly modulated by which of the two pairs of surrounding Context gratings were selectively attended. This accords with our prediction that selective attention to one or other disambiguating context should constrain motion integration. Moreover, Experiment 2 further confirmed that the influence of attended context on perceived motion direction for the central Test grating was constrained by the presence or absence of specific global motion possibilities in the context stimuli.

Experiment 3

Thus far we have attributed the attentional ‘cued-control’ effect in Experiments 1 and 2 to covert selective attention, but for further confirmation of this, we next included formal eye-tracking in Experiment 3, to ensure that eye-position did not differ systematically between our different cue conditions (and that any occasional trials with substantial gaze-deviation were eliminated from our perceptual measures). Experiment 3 also implemented a visual-contrast manipulation for the context gratings, to further explore the possible mechanisms for our newly uncovered attentional effect. One possibility is that attending to a particular pair of context gratings, as instructed, might act equivalently to an increase in the effective contrast for those attended gratings (see Carrasco et al., [2004](#); Doshier & Lu, [2000](#); Lee et al., [1999](#); Martinez-Trujillo & Treue, [2002](#); McAdams &

Maunsell, [1999](#); Reynolds, Pasternak, & Desimone, [2000](#); Zenger, Braun, & Koch, [2000](#)); or analogously, as an effective decrease for the ignored gratings. If such a modulation of effective contrast by selective attention applied, one could argue that this alone might *indirectly* promote the selective influence of the attended context gratings on perception of the central grating (in effect, by making these gratings stronger). An alternative hypothesis that we would favour, given some of our prior research on attentional modulation of contextual integration (Freeman et al., [2003](#)), is that selective attention more *directly* modulates integration between the attended context with the central grating. Such a mechanism may be analogous to that proposed in our earlier work on integration between aligned static Gabor patches (Freeman et al., [2003](#)), where the observed attentional effects involved selective modulation of perceptual contour integration, and were found *not* to be equivalent to merely increasing the effective contrast of one set of context stimuli versus another.

To assess this issue for the present motion paradigm, we next manipulated relative contrast for the pairs of contextual moving gratings, independent of which pair of gratings was cued for attention. If our previous selective-attention effects were equivalent to a manipulation of effective contrast, then it should in principle be possible to mimic those effects of attention by simply manipulating the physical contrast of selected contexts. Thus, context gratings of substantially higher *physical* contrast should now have more impact on perceived direction for the central Test grating than lower-contrast Context gratings, analogously to the cued-attention findings in Experiment 1 and 2 (for the orthogonal-motion displays). Indeed, if the physical contrast difference introduced was larger than the documented upper-limits for attentional modulation of effective contrast (see Carrasco et al., [2004](#); Lee et al., [1999](#); Reynolds et al., [2000](#)), this should presumably then override any effects from attentional influences reflecting contrast-gain modulation (i.e. the physically higher-contrast context should now always dominate).

To manipulate context salience we presented very high contrast (100%) context gratings on the vertical axis, and low contrast gratings (2.12%) on the horizontal axis, surrounding a central grating

of 14% contrast, with these contrasts remaining the same for the duration of an experimental run. This seven-fold contrast difference between contexts is much larger than any documented estimate of effective-contrast modulation due to attention (e.g. a 51% estimated increase reported by Reynolds et al., [2000](#), and even a two-fold increase reported by Carrasco et al, [2004](#), at low contrast). It should thus provide a robust test for the potential account of selective integration due indirectly to contrast gain alone (see above), which now predicts a strong impairment of cued-control in the new situation, with dominance of the much higher contrast context instead. However, if cued-control is unaffected by our contrast manipulation, this would suggest that cued-control and the corresponding selective integration is not achieved merely as an indirect outcome of effective contrast-gain modulation by attention for the contexts, but rather by more direct modulation of selective context integration (Freeman et al., [2001](#), [2003](#)).

A further condition was included in which we presented a maximum-contrast (100%) context pair together with another context pair of such low contrast (1%) that it was barely visible in the periphery (near contrast threshold, though we did not assess that threshold formally), to provide an even more extreme manipulation of relative salience. In this situation we could again test whether motion perception now becomes tonically dominated by the high-contrast context (which now has only a *very* weak competing context); or whether instead observers can still reduce the influence of the much higher-contrast context, by switching attention away from the high-contrast axis when the very low-contrast axis is cued instead.

Finally in addition to testing the original orthogonal-motion displays (as in Experiment 1, in which any global motion should be *orthogonal* to attended axis orientation), we also now generalized to a stimulus in which the directions of global motion were each *aligned* with their inducing axis. This was achieved by simply reversing the drift-directions of all context gratings relative to the central patch. The subjective result, for a given pairs of context grating stimuli when considered together with the central target stimulus, is then of global motion sliding non-rigidly along the inducing axis.

We also included two further ‘uni-directional’ stimulus configurations equivalent to those used as controls in Experiment 2, in which context pairs on each axis now both moved in the same direction (either *all horizontal* or *all vertical*). Schematics of all these different configurations are shown in the leftmost column of Figure 6, further below. By presenting these four configurations interleaved in pseudo-randomly ordered runs, we were able to eliminate any overall contingency between the currently cued axis and the ‘correct’ direction of reported global motion. As in Experiment 2 also, this should prevent observers from making stereotyped responses.

Methods

Four observers participated in this experiment, two of whom (MG, RD) had performed in Experiment 1. Observers sat in a dark sound-proof cubicle, with head stabilisation provided by a chin rest. Eye position was monitored using an ASL5000 infra-red eyetracker, sampling at 120Hz. Gratings were now Gaussian windowed (i.e. Gabor patches) with edges fading smoothly into the background, rather than sharply as in the preceding experiments. This was done to prevent possible horizontal or vertical local motion at the edges from influencing perception of drift direction (c.f. the Barber-Pole illusion, Wallach, 1935/Wuerger et al., [1996](#)). To further encourage adherence to central fixation, gratings no longer drifted smoothly, but now oscillated back and forth through one cycle over a period of 1.6 sec, with a sinusoidal velocity profile. Observers were again repeatedly told to maintain rigorous central fixation throughout the experimental runs. As a final modification, the axis cues were now positioned within the central masking disk 0.5° from fixation (either left and right or above and below fixation), rather than at the edges of the display. This was intended to eliminate any possibility of the peripheral cues causing exogenous shifts of attention, and thus to ensure that any observed attentional control was purely endogenous, or ‘top-down’ (see Figure 4 and associated animation).

> FIGURE 4 (see also animation file expt3demo.mov) <

There were three levels of contrast manipulation in this experiment. First, in the ‘Equal’ baseline

condition, both target and context gratings were presented at 14% contrast. In the intermediate contrast-modulated condition, the contrast of the context gratings was increased to the maximum 100% on the vertical axis and reduced to 2.12% on the horizontal axis ('Contrast A' condition), while the central grating remained at 14%, thus approximating for both contexts a seven-fold contrast difference relative to the central grating. In the more extreme 'Contrast B' condition, the contrast of gratings on the horizontal axis was dropped even further to 1%, rendering these barely visible (i.e. close to contrast threshold, although that threshold was not formally assessed).

Analysis of eye-position

We conducted an event-related (i.e. based on cuing-epoch) analysis of eye-position data (X and Y coordinates) to test whether observers were using any systematic eye strategy to induce global motion percepts in response to the attention-cues, by saccading or smoothly tracking in the desired direction. For example, whenever the context stimuli reinforcing vertical global motion were cued, observers might conceivably have tended to saccade or pursue vertically in order to induce or sustain perception of vertical global motion. For that cueing epoch, such behaviour should temporarily increase the variance of Y components in the time series relative to that for the X components.

> FIGURE 5 <

Using the iLab toolbox for Matlab (Gitelman, [2002](#)), eye position data for each experimental run (X and Y coordinates) were first screened for blinks and gross outliers (i.e. eye positions apparently greater than 8 degrees from fixation, which may arise during partial signal dropout). Perceptual report data corresponding to these frames were also completely excluded from further analysis (accounting for 7.5% of data on average across observers, SE 2.8%), so that none of our psychophysical measures could be contaminated by outliers in eye-position or losses of eye signal. Eye positions / saccadic endpoints for all the remaining frames were then identified (using an algorithm published by Fischer, Biscaldi, & Otto, [1993](#) with default iLab parameters). Eye data were pooled across contrast conditions and across orthogonal and aligned stimulus configurations,

but split into two sets according to whether the expected global motion percept (given the configuration and the attentional cues in a particular epoch) was either horizontal or vertical in motion-direction. Fig. 5a summarizes the saccade data split according to whether horizontal or vertical global motion was cued (blue and red segments respectively), in the form of a radial histogram. The distribution of saccade angles over 8 bins is shown as proportions of the total number of saccades, with 0.125 predicted in all bins for an isotropic distribution. The close overlap of red and blue segments suggests there was no consistent tendency for gaze to follow the cued global motion direction.

We also compared the standard deviations of X and Y components calculated for each cued global motion direction (see Figure 4b, with bars indicating means with 95% confidence limits based on the within-subjects standard error of the means, see Cousineau, [2005](#), and symbols indicating individual subject values). If observers tended to saccade (or pursue) in the direction of the cued global motion, this should appear as an increase in variance along the Y dimension relative to X for cued vertical motion, and a reversed relationship (X variance greater than Y) for cued horizontal motion. There was no such pattern observed, neither for the group means nor for the individual data.

Note that saccade data from the entire duration of each epoch were used to compute the above standard deviations. In addition, we undertook a further more detailed analysis of the epoch-related time-course of raw eye-positions, to test for any evidence of pursuit of the oscillating gratings. For X and Y coordinates in each epoch, we subtracted the mean and then rectified the resulting timecourse, so that any displacement of gaze from the central tendency of fixation should register as positive values on average. To highlight possible pursuit movements (tracking the oscillating motion) we excluded those occasional epochs where the standard deviation of eye displacements was greater than the oscillatory trajectory of the gratings (usually due to occasional signal loss). Importantly, we also excluded any such epochs from our psychophysical data, so that our

perceptual measures are for exactly the same epochs to which we applied stringent eye-position analysis. Figure 5c reveals no systematic effects of cueing on the magnitude of eye displacements in X relative to Y directions, contrary to the expectation if horizontal versus vertical pursuit behaviour arose.

While it may sometimes be difficult to maintain perfect fixation throughout this kind of experiment, the important point is that we find no consistent patterns of oculomotor behaviour that can account fully for our observed pattern (see below) of cue-controlled subjective reports (i.e. no systematic horizontal or vertical differences between the two cue conditions). Although we did not record eye-position formally in Experiments 1 and 2, the cued-control effects found in Experiment 3 (see below) do at least allow the conclusion that systematic differences in oculomotor behaviour do not appear to be necessary for the present attentional cued-control to emerge.

> FIGURE 6 <

Results

We first present results for the Orthogonal and Aligned stimulus configurations, which were potentially bi-stable, such that the context pairs on either of the two axes could induce different directions of global motion (unlike the uni-directional displays; see schematic illustrations of all display types in leftmost column of Fig. 6, plus main text below). As in Experiment 1 and 2, for all four observers the perceptual reports of movement direction for the potentially bi-stable displays strongly depended on which axis of context gratings was currently attended (see radial histograms of reported directions presented centrally in Figs. 6a and 6b, which represent data pooled across observers). Results were very similar for the Equal and Contrast A conditions (left and middle radial histograms in 6a and 6b). For example, when displaying the orthogonal-motion configuration (Figure 6a, c.f. the same configuration used in Experiments 1 and 2), observers tended to report horizontal motion when cued to the vertical context axis (yellow segments in radial histograms), but vertical motion when the horizontal axis was cued (blue segments), even when the context gratings on that cued axis were much lower in contrast (Contrast A condition) but clearly visible. In the

Contrast B condition, however, where the contrast of the horizontal axis gratings was now lowered to 1% so that they were barely visible (rightmost radial histograms in Figs 6a and 6b), observers now tended to report diagonal (i.e. local) motion for the central grating when the horizontal axis was cued (see blue segment in the rightmost radial plot of Fig 6a), rather than vertical global motion (which would normally be seen when the horizontal axis contrast was higher). This presumably arose because the horizontal flankers were now too weak to induce global motion percepts, with perception of the central grating now reverting to local perception. But critically, these results from Contrast B still show that attending to the very low-contrast context (even when shown at only 1% contrast, evidently too low to induce global motion as indicated from the prevalence of diagonal reports now) could still greatly reduce the impact of the high-contrast context, despite the latter now having 100 times greater contrast (compare yellow and blue segments in rightmost radial histograms of Fig. 6a).

The complementary pattern was found for the new aligned-motion condition (Figure 6b), except of course that (with the exception of the 1% contrast context in condition B, which again now led to diagonal reports when the very low contrast axis was attended) horizontal motion was now generally reported when the horizontal axis was cued and vertical reports when the vertical axis was cued (i.e. the opposite to what was found with the orthogonal-motion condition, but the analogous outcome when considering the global-motion possibilities in the stimulus and the currently attended pair of context stimuli). This further confirms the specificity of our results, and their dependence on the actual display configuration presented, in addition to which aspect of the context in that display was attended.

As pointed out by an anonymous reviewer, the use of contextual gratings with unequal contrasts (in our Contrast A and Contrast B conditions) might in principle lead perceived direction of global motion to be biased in the direction of the higher contrast grating, as observed in the case of ‘plaids’ composed of superimposed gratings (Stone, Watson, & Mulligan, [1990](#)). But an additional analysis

of our data plotting radial histograms but now using much finer bins showed no evidence for such a bias with our displays (see Appendix 2 in Supplementary Materials) and in any case cued-control was still found clearly here.

The results from the new Aligned-Orthogonal manipulation (Fig 6a vs 6b) provides some evidence that suggest further points concerning possible mechanisms (and brain regions) underlying motion integration. One likely possibility is that dorsal motion-sensitive areas (such as MT) are involved in the kind of long-range motion integration studied here (e.g. Huang et al., [2007](#)), but it is also possible in principle that form-based (possibly more ventral) mechanisms might also contribute to formation of a global pattern, such as a chevron (Lorenceanu & Alais, [2001](#)). If the latter applied, one might anticipate relatively decreased global-motion and increased local-motion reports for the Aligned condition, where the sliding motion of the stimulus did not accord with a rigid form interpretation. But in fact we observed comparable results for the Aligned and Orthogonal configurations, indicating that rigid form may not constrain this particular case of long-range motion integration. On the same form-based issue, we consider it unlikely that lateral interactions of the specific kind we studied previously using static Gabors (e.g. Freeman et al., [2001](#), [2003](#)) could on their own explain possible grouping of the dynamic gratings into a potential zig-zag contour here, as the well-established lateral interactions with static stimuli are maximally sensitive to collinear arrangements of contour segments (with common local and global orientations), but minimally sensitive when the display elements differed in local orientation by 90°, as they did for the dynamic Orthogonal conditions here. We thus suggest that while the general principle from our static studies (Freeman et al., [2001](#), [2003](#)) may accord with the present dynamic study in emphasizing the role of selective attention to one or other aspect of surrounding context, the importance of form constraints may differ between the two cases. The specific involvement of particular brain areas (e.g. MT+, and/or further dorsal or ventral areas) is clearly an issue for a future research using fMRI or other neurophysiological measures in our new paradigm, and cannot be resolved by psychophysics alone..

For the new uni-directional configurations (where both axes reinforce the same direction of either horizontal or vertical global motion, see leftmost schematic illustrations in Fig 6c and 6d), radial histograms for the Equal and Contrast A conditions both indicate that cueing had no effect on direction of global motion reported (Figure 6c and 6d), thus further confirming that subjects' reports faithfully reflected the global motion available in the stimulus, with a very different outcome to the orthogonal or aligned configurations (whose results were shown in Fig 6a and 6b, as discussed above). However in the extreme Contrast B condition, an effect of cueing became apparent even for the uni-directional configurations, with local diagonal motion tending to be reported only when the low-contrast gratings were cued (which themselves were barely visible, and thus evidently ineffective in inducing their corresponding global motion). This is again consistent with high-contrast gratings having a greatly reduced impact on perceived motion when attention is withdrawn from them.

To further visualise the distribution of motion-direction reports as a function of cueing and contrast, Figure 6e shows separate graphs for epochs with cues to the vertical (high contrast) versus cues to the horizontal (low contrast) axes respectively (left or right parts of Fig. 6e, respectively). Results are now expressed in terms of the relative proportion of responses in three directions: local motion, global motion in the direction induced by the cued context, or global motion in the direction induced by the uncued context. As this coding of the data ignores the actual motion direction and context configuration (i.e. measuring only whether the motion is or is not in the direction expected given the current context and cueing state), these data could now be pooled across Orthogonal and Aligned stimulus configurations (c.f. a similar pooling procedure used for the eye-data above). The relative proportions, averaged across observers, are visualized as stacked bar graphs with the red shading representing the proportion of diagonal local motion; the white intermediate region represents the average proportion of reports in the direction induced by the cued axis; and the green bars at the top represents the remaining average proportion of epochs with the orthogonal or uncued global motion being reported (i.e. in the direction orthogonal to that expected given the currently

cued axis for the current display configuration). Superimposed on this are the results for individual observers, visualized as cumulative proportions of the three motion directions, thus directly representing the breakdown of the averaged results displayed in the background. The filled red datapoints show the proportion of oblique local motion reported; while the open red/white datapoints represent the same local motion proportion *plus* the proportion of global motion reports in direction induced by the *cued* axis. Distance from the top of the graph to the latter red/white datapoints then represents the remaining proportion of *uncued* motion orthogonal to that expected given the cue and stimulus.

It can be seen that when the vertical (high-contrast axis) was cued (leftmost half of Figure 6e), the corresponding global motion contributed by far the greatest proportion of responses (note the height of the open red/white datapoints on the plots, and also the substantial extent of the grey stack in the background representing the group data). Moreover, this proportion was fairly stable for all three contrast conditions. However when the horizontal low-contrast axis was cued (rightmost half of Figure 6e), the proportion of local motion (solid red datapoints, red stacks at bottom) dramatically increases as the context contrast drops (along the X-axis of the plot, for the different contrast conditions). The increasing area of the top green region with more extreme contrast manipulations indicates there was also a slight increase in the proportion of uncued global motion (corresponding to the salient high-contrast axis), but even at maximum this still accounts for on average only ~25% of the responses.

The uni-directional global motion conditions present a very similar pattern (Fig 6f) for the extreme Contrast B condition. Thus again there is a strong dominance of the motion corresponding to the cued context (white intermediate region and open red/white symbols), with the exception of the extreme Contrast B condition where again the local motion (red region, solid red points) dominates whenever the low-contrast axis is cued.

Note that Figures 6e & f also show a slightly wider distribution of datapoints for the most extreme contrast B condition but only when the low contrast gratings were cued, with responses tending to reflect more of a combination of local motion, uncued global motion and a residual proportion of the cued global motion. This increase in variability would be consistent with the low contrast gratings occasionally becoming visible and consequently inducing global motion.

> FIGURE 7 <

Figure 7a graphs a quantitative measure of cued-control for orthogonal- and aligned-motion configurations, computed (as described in Methods above) by comparing the proportion of global motion reports associated with the high-contrast axis (relative to the other global and local motion directions) between the two cueing-epochs. Positive values indicate that this global motion direction was reported frequently when the high-contrast vertical axis was cued, but infrequently when the other axis was cued; conversely low values would indicate that the distribution of motion direction reports did not change consistently between epochs. Mean results across the four observers are shown, with errorbars reflecting 95% confidence intervals based on the within-subjects standard error (Cousineau, [2005](#)). Superimposed on this are results for individual observers (filled symbols for Orthogonal configuration, open for Aligned). Values of cued-control were significantly greater than zero across all contrast conditions. Critically, there was no consistent drop in cued-control between Equal and Contrast A conditions. Even for Contrast B, cued-control values mostly remained on average significantly above zero, for the four observers taken together. These results confirm that in the majority of cases responses were still yoked to cueing-epochs, even under conditions where the high-contrast contexts might otherwise have been expected to invariably dominate perception, regardless of cueing, due to their higher salience.

An anonymous reviewer suggested that integration be easier with lower rather than higher contrast gratings (see Lorenceau & Zago, [1999](#)), thus actually predicting an opposite tendency for responses

to be dominated by the lower contrast gratings. Alternatively, integration might arguably become harder to achieve if large differences in contrast induce perceptual segmentation of the centre from specific peripheral gratings. But the preservation of cued-control across the Equal and Contrast A conditions seems to argue against either of those possibilities (see also the results from a dominance analysis, below), and certainly means that our cued-control effect cannot be reduced to contrast per se.

For the unidirectional motion conditions (Figure 7b), cued-control values were, as expected, close to zero in the Equal and Contrast A condition. This reflects the fact that only one direction of global motion was present in the display, and that observers reported this faithfully in both cueing epochs, while reporting local motion only infrequently. In the Contrast B condition only, however, cued-control values are on average significantly higher than zero (rightmost datapoints in Fig 7b). This increase in cued-control mostly reflects the tendency to switch to reporting oblique local motion whenever the very low-contrast (1%) axis was attended, with these weak gratings evidently having insufficient contrast to induce global motion. This outcome nevertheless still shows that the influence of the much higher-contrast context gratings could be successfully overcome by attending away from them.

A further analysis was conducted to quantify the extent to which global motion reports are tonically dominated by the axis with the higher contrast (i.e. vertical), regardless of cueing-epoch. This measure was obtained by taking the overall proportion (i.e. across all cueing-epochs) of global motion reports in the direction expected to be induced by the vertical axis, and subtracting the proportion of reports in the other directions. If there were any net tendency for the high-contrast vertical axis to dominate over the other directions, the resulting value should be greater than zero (as might be predicted if the impact of a given context automatically depended on its contrast relative to the other context axis). On the other hand, values close to zero should obtain if the other two directions were reported with a frequency equalling that of the high-contrast direction, as might

be expected if observers can completely override the influence of the high-contrast context and report the other directions when attention is directed away from the high-contrast flankers.

Results for the bi-directional displays (Figure 7c) confirmed the latter prediction, with estimates of vertical-axis dominance remaining close to zero for all contrast conditions. This indicates, critically, that the global motion associated with the high-contrast axis did not generally dominate responses due to its greater salience, even in the Contrast B condition where the gratings on the low-contrast axis were barely visible. Likewise there is no evidence here for low-contrast contexts dominating due to being more easily integrated (cf. Lorenceau & Zago, [1999](#)). Results for the unidirectional displays (Fig. 7d) show that responses were dominated by the global motion that both axes equally induced when both were clearly visible, however this dominance is almost eliminated in the Contrast B condition. This again reflects the tendency for observers to switch away from this global motion to reporting local motion during the epochs where the low-contrast axis was cued.

These results suggest that cued-control effects on motion perception are unlikely to reflect solely an effective increase in contrast for attended contexts relative to unattended. Such an account would presumably predict that very large physical contrast modulations, as we introduced in this experiment, should tonically bias perception in favour of the higher-contrast context, outweighing any attentional influences if those only reflect relatively modest changes in effective contrast due to attention (c.f. Carrasco et al., [2004](#); Lee et al., [1999](#); Reynolds et al., [2000](#)). In the Contrast A condition there was a substantial difference between high and low-contrast contexts here, but this did not change the impact of cued-control, compared to the equal contrast condition. Even when one context was made so low in contrast as to be barely visible (contrast B condition, where the contrast difference between context gratings was ~100-fold), cued-control still remained strong in most cases. These latter results further suggest that cued-control does not necessarily depend on having two clearly visible alternative contexts in competition with each other, but rather that observers can suppress the inducing effects of a single pair of high-contrast context gratings simply

by attending away from them.

General discussion

It is well established that the classical aperture problem of motion perception (Marr & Ullman, 1981; Wuerger et al., 1996) may be resolved in the context of other moving contextual stimuli elsewhere in the scene (Mingolla, Todd, & Norman, 1992; Alais et al., 1998; Lorenceau & Alais, 2001; Lorenceau & Shiffrar, 1992). However when there is more than one possible context present, multistability can arise if each supports a different interpretation of motion. Here we show that observers could readily switch between specific subjective states in accord with an attentional cue to select to one pair of context gratings or the other, which then changed the appearance of the central grating correspondingly.

We used a stimulus that could induce (primarily) three very different states of subjective drift direction, depending on the selective long-range integration of a central grating with its surrounding context. In such a stimulus (potentially analogous to any real-world scene in which there may be separate objects moving across each other in different directions, e.g. see Albright & Stoner, 1995), achieving a coherent and stable percept requires *selective* integration of only the relevant directions of motion present in the scene (e.g. those belonging to the same object, or to the same attended sample) separately from those for motions belonging to other sources or samples (e.g. see Grossberg et al., 2001).

Under free-viewing our new stimulus was vigorously multistable (see Fig 1h), switching chaotically between three modes of perceived motion, consistent with there being multiple perceptual interpretations in mutual competition. In Experiment 1 we showed for the first time that when observers were instructed to attend in alternation to one context or the other (i.e. the vertical or horizontal pair of flankers), as instructed by a periodically changing attentional cue, they reported a concurrent 90° switch in the perceived direction of global motion.

This phenomenon of cued-control goes beyond past demonstrations of some top-down control over multistable perception, as when observers can successfully ‘hold’ an explicitly pre-specified perceptual state (Long & Toppino, [2004](#); Meng & Tong, [2004](#); van Ee et al., [2005](#); Verstraten & Ashida, [2005](#)). In such ‘hold’ paradigms, dominance phases typically cannot be extended indefinitely, but are inevitably curtailed unpredictably by an unwilled spontaneous switch to an alternative perceptual state. Thus, when a switch does eventually arise it may appear to be driven by autonomous mechanisms. The critical difference in the present paradigm is that observers can actively initiate a switch to a specific new state in response to a predictable attentional cue.

This aspect of the present results may be of relevance to a long-standing debate over whether subjective switching involves passive or active processes (Leopold et al., [1999](#); Long & Toppino, [2004](#)). For example, uncontrolled switches in many previous paradigms have been attributed to passive stochastic noise or ‘satiation’ in competitive networks at the sensory-encoding level (e.g. Kohler & Wallach, [1944](#); Attneave, [1971](#)), or to semi-autonomous executive processes whose function may be to actively search for and impose new perceptual interpretations on the perceptual system (Leopold et al., [1999](#)). Frontal activations in fMRI during spontaneous switching have sometimes been cited as possible support for the latter proposal (Lumer, Friston, & Rees, [1998](#); Sterzer & Kleinschmidt, [2007](#)), while potential contributions from eye-movements have also been suggested. Our analysis of eye-movements here help to rule out the possibility that the present observations of cued-control might be accounted for solely by consistent patterns of ocular-motor behaviour (for Expt. 3 in particular). Our results lend clear new behavioural support to the notion of active top-down state selection, but now for the specific case of covert selective attention to one or other aspect of an inducing context. Future studies (including fMRI and other neural measures) could now capitalize on our new design for studying control of multistable perception by attention to one or other aspect of surrounding context, to determine whether successful cue-controlled subjective switches (as opposed to unsuccessful or unwilled switching) might be accompanied by

specific patterns of, say, pre-frontal activity.

The present paradigm inherited some of the critical features of our earlier paradigm for investigating selective contour integration in static displays (Freeman et al., [2001](#), [2003](#)), such as the separation of surrounding context from the central target to be judged, and cued attentional selection of one aspect of surrounding context versus another. These aspects allowed the context to be manipulated both physically, and in terms of attention, without changing the central target itself or its task-relevance, yet while measuring the effects of the contextual manipulations on perception of the same central target.

For instance, this allowed us to manipulate the context in Experiment 2, so that only one rather than two conflicting directions of global motion might be available, while independently measuring the consequences of this for perceiving the target. This approach contrasts with investigations of motion integration that physically manipulate the target stimulus itself (Alais et al., [1998](#); Castet et al., [1999a](#); Castet & Zanker, [1999b](#); Lorenceau & Alais, [2001](#)). While our Experiment 1 showed that the effect of contextual attention on test perception can be strong, Experiments 2 & 3 demonstrated that this is not absolute, but is constrained by the objective availability of specific global motion vectors in the display. This dependence on the global flanking context helps exclude a simple ‘feature-based’ attention account for attentional control, as an explanation for the present results. Note that owing to the aperture problem (Marr & Ullman, [1981](#)), the windowed gratings that comprised the local components of our stimuli were always in principle just as locally ambiguous when arranged in one stimulus configuration as in another, thus in principle always representing motion components spanning a full 90° range of directions. Theoretically, one possible means for attentional control might therefore have been to simply enhance the desired motion component within this range at the expense of others (c.f. a feature-based attention mechanism of gain control, Treue & Martinez Trujillo, [1999](#)). But our observations do not accord with such a simple mechanism for local motion modulation, which would have predicted no effect of global

display configuration, but imply instead that top-down control mechanisms interface with stimulus-driven (configuration-dependent) mechanisms to constrain how local motions are integrated over space into representations of global motion.

The distinction between context and target also provided an additional advantage for the subjective measure used here, as observers were never directly asked to try to switch between specific directions of global motion, but merely to attend selectively to different contexts while reporting whichever direction of motion was currently seen in the central grating. Consequently, we were able to present the new configurations in Experiments 2 and 3 in an unpredictable and interleaved fashion, to naïve observers, without ever indicating to them what directional reports were expected. This indirect approach differs from past studies where the explicit instructions were to hold a specific perceptual state for as long as possible (Long & Toppino, [2004](#); Meng & Tong, [2004](#); van Ee et al., [2005](#); Verstraten & Ashida, [2005](#)). In addition, by testing naïve observers who had previously experienced only one mapping between cue and report, we could test the extent to which their reports were merely based on a previously learned contingency between attentional cue and reported direction. Despite deliberately maximizing the likelihood of such learned responses, the observed stimulus-dependence in Experiments 2 and 3 confirmed here that observers' reports indeed faithfully reflected the perceptual alternatives afforded by the stimulus, even while being strongly modulated by selective attention to one or other aspect of the surrounding context.

In Experiment 3 we further tested the hypothesis that attention might operate as a local gain control (Martinez-Trujillo & Treue, [2002](#); Reynolds et al., [2000](#)), simply boosting the effective contrast of the local representation of the selected context, and thereby indirectly promoting its integration. If this was the only mechanism involved, then presumably gross manipulation of physical contrast for one or other component of the context should function to override the effects of attention, causing the context with the higher contrast to dominate perception regardless of attentional cuing.

However, in fact we found that cued-control did not suffer as a function of such contrast

manipulations. In common with our previous study using a similar logic to investigate attentional modulation of interactions between static Gabor patterns (Freeman et al., [2003](#)), the persistence of attentional effects despite gross differences in context contrast may be taken to support an alternative hypothesis: namely that attention might gate the global integration between the local components that are currently selected, rather than just modulating the response to the local components themselves.

According to a biased-competition account (Desimone & Duncan, [1995](#)), alternative motion perceptions may actively compete for dominance when faced with a multistable stimulus. Selection may then be achieved by boosting one perception in a top-down manner via selective attention, which consequently suppresses the other. While some aspects of our results may accord with this very general competitive account, we note one apparent exception: in the ‘Contrast B’ condition (with one context now barely visible) cued-control remained high even though the barely-visible low-contrast context was evidently too weak to elicit much competing global motion. This suggests that as well as facilitating integration of a relevant context, attention may also function effectively to segment out motion that is not relevant, even when there is no strongly competing global organization. Though evidence of attentional modulation in the apparent absence of competition may seem at odds with the general biased-competition account (c.f. Treue & Martinez Trujillo, [1999](#)), it remains possible that some form of ‘competition’ is still provided by the local motion interpretation. This could explain why local oblique motion was typically perceived whenever the high-contrast context was ignored, as this would become dominant when the competing global organization was suppressed.

Taken together, our previous work on attentional modulation of contour integration between static Gabors (Freeman et al., [2001](#), [2003](#)) and the present work with dynamic gratings, while differing in detail both suggest a rather general role for selective attention that may previously have been overlooked, namely modulation of perceptual integration by selective attention to one or other

aspect of potentially disambiguating context. The implementation of this general principle may nevertheless differ in detail for the two cases, with integrative form mechanisms sensitive to collinear configurations being implicated for the static Gabors (Freeman et al., [2001](#), [2003](#)), but context-dependent mechanisms sensitive to global patterns of motion (though less affected by collinearity) applying here.

Our new evidence for a role of attention in selective motion integration and segmentation (as for the reduced impact of the high-contrast flankers when ignored in Experiment 3) may accord with models of motion processing in which these commonly opposed functions (Braddick, [1993](#)) might actually reflect aspects of the same mechanism (Grossberg et al., [2001](#)) rather than necessarily two independent processes. Mechanisms of motion integration have been extensively studied in physiology, typically with stimuli comprising overlapping motion components (Albright & Stoner, [1995](#); Rust, Mante, Simoncelli & Movshon, [2006](#)), or as a function of contextual cues for occlusion (Duncan et al., [2000](#)). However, a recent single-cell study (Huang et al., [2007](#)) provided new evidence consistent with motion integration in MT, showing apparent vector summation between non-overlapping stimulus components inside and outside the classical receptive field (CRF). Such integration was found only when the stimulus within the CRF was itself ambiguous; but when the stimulus in the CRF was rendered physically unambiguous, segmentation from the context arose instead (i.e. apparent vector repulsion). Relatively little is known as yet about the exact mechanisms involved in such interactions, and less still about how spatial attention to selected contexts might bias such mechanisms at the cellular level, as might be studied in future by applying invasive methods to the attentional-context paradigm introduced here. One speculative possibility, given the present results, is that selective attention to a given aspect of the context may facilitate integration of that selected context with the ambiguous grating in the receptive field; the latter having thus been disambiguated, may then segment itself from the other unattended context. This would still require some explanation for the role of attention in Experiment 3, however, where segmentation (decreased impact) from an unattended context occurred even when the contrast of the attended

context was itself too low to fully disambiguate the target.

For more than a century psychologists have pondered the extent to which the perception of multistable stimuli may be influenced by voluntary attention. In the conclusion of their recent study, van Ee et al ([2005](#), p. 50) wrote: '*Voluntary control in perceptual bi-stability is clearly limited. Although we can modify the perceptual reversal process, we are often not able to choose the moment of reversal*' (pg. 50). In the present study we demonstrated that, when provided with appropriate contexts for attention to select between, observers can not only voluntarily switch between alternative perceptions with remarkable facility, but that the timing of this can be controlled with appropriate attentional cuing. These new results implicate a top-down mechanism that can selectively integrate specific combinations of local stimulus components in a controlled manner, and make their emergent global properties available to phenomenal perception. While remaining firmly grounded on the available stimulus evidence, this enables us to voluntarily switch between two profoundly different ways of seeing the same physical stimulus, depending on which aspect of the context is currently attended.

Acknowledgements

This research was funded by a BBSRC research grant S20366 to JD & EF. JD holds a Royal Society - Leverhulme Trust Senior Research Fellowship.

References

- Alais, D., van der Smagt, M. J., van den Berg, A. V., & van de Grind, W. A. (1998). Local and global factors affecting the coherent motion of gratings presented in multiple apertures. *Vision Res.*, 38, 1581-1591. [[PubMed](#)]
- Albright, T. D. & Stoner, G. R. (1995). Visual motion perception. *Proc.Natl.Acad.Sci.U.S.A*, 92, 2433-2440. [[PubMed](#)] [[Article](#)]
- Attneave, F. (1971). Multistability in Perception. *Scientific American*, 225, 61-71. [[PubMed](#)]
- Braddick, O. (1993). Segmentation Versus Integration in Visual-Motion Processing. *Trends in Neurosciences*, 16, 263-268. [[PubMed](#)]
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spat.Vis.*, 10, 433-436. [[PubMed](#)]
- Braun, J. (2002). Visual attention: light enters the jungle. *Curr.Biol.*, 12, R599-R601. [[PubMed](#)]
- Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nat.Neurosci.*, 7, 308-313. [[PubMed](#)]
- Casco, C., Grieco, A., Campana, G., Corvino, M. P., & Caputo, G. (2005). Attention modulates psychophysical and electrophysiological response to visual texture segmentation in humans. *Vision Research*, 45, 2384-2396. [[PubMed](#)]
- Castelo-Branco, M., Formisano, E., Backes, W., Zanella, F., Neuenschwander, S., Singer, W. Goebel, R. (2002). Activity patterns in human motion-sensitive areas depend on the interpretation of global motion. *Proc.Natl.Acad.Sci.U.S.A*, 99, 13914-13919. [[PubMed](#)] [[Article](#)]
- Castet, E., Charton, V., & Dufour, A. (1999a). The extrinsic/intrinsic classification of two-dimensional motion signals with barber-pole stimuli. *Vision Res.*, 39, 915-932. [[PubMed](#)]
- Castet, E. & Zanker, J. (1999b). Long-range interactions in the spatial integration of motion signals. *Spat.Vis.*, 12, 287-307. [[PubMed](#)]
- Cavanagh, P. (1992). Attention-based motion perception. *Science*, 257, 1563-1565. [[PubMed](#)]
- Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorial in Quantitative Methods for Psychology*, 1, 42-45.
- Desimone, R. & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18, 193-222. [[PubMed](#)]
- Dosher, B. A. & Lu, Z. (2000). Mechanisms of perceptual attention in precuing of location. *Vision Research*, 40, 1269-1292. [[PubMed](#)]
- Duncan, R. O., Albright, T. D., & Stoner, G. R. (2000). Occlusion and the interpretation of visual motion: perceptual and neuronal effects of context. *J.Neurosci.*, 20, 5885-5897. [[PubMed](#)]
- Fischer, B., Biscaldi, M., & Otto, P. (1993). Saccadic eye movements of dyslexic adult subjects. *Neuropsychologia*, 31, 887-906. [[PubMed](#)]
- Freeman, E. D. & Driver, J. (2006). Subjective appearance of ambiguous structure-from-motion can be driven by objective switches of a separate less ambiguous context. *Vision Res.*, 46, 4007-

4023. [[PubMed](#)]

- Freeman, E. D., Driver, J., Sagi, D., & Zhaoping, L. (2003). Top-down modulation of lateral interactions in early vision: does attention affect integration of the whole or just perception of the parts? *Curr.Biol.*, *13*, 985-989. [[PubMed](#)]
- Freeman, E. D., Sagi, D., & Driver, J. (2001). Lateral interactions between targets and flankers in low-level vision depend on attention to the flankers. *Nat.Neurosci.*, *4*, 1032-1036. [[PubMed](#)]
- Gandhi, S. P., Heeger, D. J., & Boynton, G. M. (1999). Spatial attention affects brain activity in human primary visual cortex. *Proc.Natl.Acad.Sci.USA*, *96*, 3314-3319. [[PubMed](#)] [[Article](#)]
- Gilbert, C. D. & Wiesel, T. N. (1990). The influence of contextual stimuli on the orientation selectivity of cells in primary visual cortex of the cat. *Vision Research*, *30*, 1689-1701. [[PubMed](#)]
- Gitelman, D. R. (2002). ILAB: a program for postexperimental eye movement analysis. *Behav.Res.Methods Instrum.Comput.*, *34*, 605-612. [[PubMed](#)]
- Grossberg, S., Mingolla, E., & Viswanathan, L. (2001). Neural dynamics of motion integration and segmentation within and across apertures. *Vision Res.*, *41*, 2521-2553. [[PubMed](#)]
- Huang, X., Albright, T. D., & Stoner, G. R. (2007). Adaptive surround modulation in cortical area MT. *Neuron*, *53*, 761-770. [[PubMed](#)]
- Ito, M. & Gilbert, C. D. (1999). Attention modulates contextual influences in the primary visual cortex of alert monkeys. *Neuron*, *22*, 593-604. [[PubMed](#)]
- Kapadia, M. K., Ito, M., Gilbert, C. D., & Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: Parallel studies in human observers and in V1 of alert monkeys. *Neuron*, *15*, 843-856. [[PubMed](#)]
- Khoe, W., Freeman, E., Woldorff, M. G., & Mangun, G. R. (2006). Interactions between attention and perceptual grouping in human visual cortex. *Brain Res.*, *1078*, 101-111. [[PubMed](#)]
- Kim, J. & Wilson, H. R. (1997). Motion integration over space: interaction of the center and surround motion. *Vision Res.*, *37*, 991-1005. [[PubMed](#)]
- Kohler, W. & Wallach, H. (1944). Figural aftereffects: An investigation of visual processes. *Proceedings of the American Philosophical Association*, *88*, 269-357.
- Lee, D. K., Itti, L., Koch, C., & Braun, J. (1999). Attention activates winner-takes-all competition among visual filters. *Nat.Neurosci.*, *2*, 375-381. [[PubMed](#)]
- Leopold, D. A. & Logothetis, N. K. (1999). Multistable phenomena: changing views in perception. *Trends Cogn Sci.*, *3*, 254-264. [[PubMed](#)]
- Long, G. M. & Toppino, T. C. (2004). Enduring interest in perceptual ambiguity: alternating views of reversible figures. *Psychol.Bull.*, *130*, 748-768. [[PubMed](#)]
- Lorenceau, J. & Alais, D. (2001). Form constraints in motion binding. *Nat.Neurosci.*, *4*, 745-751. [[PubMed](#)]
- Lorenceau, J. & Shiffrar, M. (1992). The influence of terminators on motion integration across space. *Vision Res.*, *32*, 263-273. [[PubMed](#)]

- Lorenceanu, J. & Zago, L. (1999). Cooperative and competitive spatial interactions in motion integration. *Vis.Neurosci.*, 16, 755-770. [[PubMed](#)]
- Lu, Z. L. & Sperling, G. (1995a). Attention-generated apparent motion. *Nature*, 377, 237-239. [[PubMed](#)]
- Lu, Z. L. & Sperling, G. (1995b). The functional architecture of human visual motion perception. *Vision Res.*, 35, 2697-2722. [[PubMed](#)]
- Luck, S. J., Chelazzi, L., Hillyard, S. A., & Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *J.Neurophysiol.*, 77, 24-42. [[PubMed](#)]
- Lumer, E. D., Friston, K. J., & Rees, G. (1998). Neural correlates of perceptual rivalry in the human brain. *Science*, 280, 1930-1934. [[PubMed](#)]
- Majaj, N. J., Carandini, M., & Movshon, J. A. (2007). Motion integration by neurons in macaque MT is local, not global. *J.Neurosci.*, 27, 366-370. [[PubMed](#)]
- Marr, D. & Ullman, S. (1981). Directional selectivity and its use in early visual processing. *Proc.R.Soc.Lond B Biol.Sci.*, 211, 151-180. [[PubMed](#)]
- Martinez-Trujillo, J. & Treue, S. (2002). Attentional modulation strength in cortical area MT depends on stimulus contrast. *Neuron*, 35, 365-370. [[PubMed](#)]
- McAdams, C. J. & Maunsell, J. H. R. (1999). Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *Journal of Neuroscience*, 19, 431-441. [[PubMed](#)]
- Meng, M. & Tong, F. (2004). Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *Journal of Vision*, 4(7):2, 539-551, <http://journalofvision.org/4/7/2/>, doi:10.1167/4.7.2. [[PubMed](#)] [[Article](#)]
- Mingolla, E., Todd, J. T., & Norman, J. F. (1992). The perception of globally coherent motion. *Vision Res.*, 32, 1015-1031. [[PubMed](#)]
- Moran, J. & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, 229, 782-784. [[PubMed](#)]
- Motter, B. C. (1993). Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *J.Neurophysiol.*, 70, 909-919. [[PubMed](#)]
- Nakayama, K., Shimojo, S., & Silverman, G. H. (1989). Stereoscopic depth: its relation to image segmentation, grouping, and the recognition of occluded objects. *Perception*, 18, 55-68. [[PubMed](#)]
- Polat, U., Mizobe, K., Pettet, M., Kasamatsu, T., & Norcia, T. (1998). Collinear stimuli regulate visual responses depending on cell's contrast threshold. *Nature*, 391, 580-584. [[PubMed](#)]
- Polat, U. & Sagi, D. (1993). Lateral interactions between spatial channels: Suppression and facilitation revealed by lateral masking experiments. *Vision Research*, 33, 993-999. [[PubMed](#)]
- Raymond, J. E., O'Donnell, H. L., & Tipper, S. P. (1998). Priming reveals attentional modulation of human motion sensitivity. *Vision Res.*, 38, 2863-2867. [[PubMed](#)]

- Reynolds, J. H. & Chelazzi, L. (2004). Attentional modulation of visual processing. *Annu.Rev.Neurosci.*, 27, 611-647. [[PubMed](#)]
- Reynolds, J. H., Pasternak, T., & Desimone, R. (2000). Attention increases sensitivity of V4 neurons. *Neuron*, 26, 703-714. [[PubMed](#)]
- Rust, N. C., Mante, V., Simoncelli, E. P., & Movshon, J. A. (2006). How MT cells analyze the motion of visual patterns. *Nat.Neurosci.*, 9, 1421-1431. [[PubMed](#)]
- Sterzer, P. & Kleinschmidt, A. (2007). A neural basis for inference in perceptual ambiguity. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 323-328. [[PubMed](#)] [[Article](#)]
- Stone, L. S., Watson, A. B., & Mulligan, J. B. (1990). Effect of contrast on the perceived direction of a moving plaid. *Vision Res.*, 30, 1049-1067. [[PubMed](#)]
- Treue, S. (2001). Neural correlates of attention in primate visual cortex. *Trends Neurosci.*, 24, 295-300. [[PubMed](#)]
- Treue, S. & Martinez Trujillo, J. C. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature*, 399, 575-579. [[PubMed](#)]
- van Ee, R., van Dam, L. C., & Brouwer, G. J. (2005). Voluntary control and the dynamics of perceptual bi-stability. *Vision Res.*, 45, 41-55. [[PubMed](#)]
- Verstraten, F. A. & Ashida, H. (2005). Attention-based motion perception and motion adaptation: what does attention contribute? *Vision Res.*, 45, 1313-1319. [[PubMed](#)]
- Verstraten, F. A., Cavanagh, P., & Labianca, A. T. (2000). Limits of attentive tracking reveal temporal properties of attention. *Vision Res.*, 40, 3651-3664. [[PubMed](#)]
- Wuerger, S., Shapley, R., & Rubin, N. (1996). On the visually perceived direction of motion" by Hans Wallach: 60 years later. *Perception*, 25, 1317-1367.
- Yeshurun, Y. & Carrasco, M. (1998). Attention improves or impairs visual performance by enhancing spatial resolution. *Nature*, 396, 72-75. [[PubMed](#)]
- Zenger, B., Braun, J., & Koch, C. (2000). Attentional effects on contrast detection in the presence of surround masks. *Vision Research*, 40, 3717-3724. [[PubMed](#)]

Figure legends

Figure 1

Examples of aperture motion and long-range context effects: a) Single stimulus with diagonal perceived drift (e.g. in direction of red arrow). b) In the presence of orthogonally oriented contextual gratings (which each on their own may appear to drift diagonally in the direction indicated by the red arrows), all gratings appear to drift together either horizontally (yellow arrow), or (c) vertically (blue arrow), depending on the local drift direction of the context. d) A combination of both types of context gratings produces spontaneous ‘multistable’ switching between percepts of horizontal, vertical motion (here, in directions orthogonal to the two intersecting axes along which the context-pairs are arranged), or local diagonal motion, under free-viewing. In the experiments, white dots over the context gratings could be used as attentional cues (as illustrated in d), while more peripheral white dots served as cursors rotating around the centre, used by the observer to indicate the vector of currently perceived central motion (in this example diagonal). e-h) radial histograms showing the proportional distribution of cursor positions over three bins during free-viewing of the stimulus shown above each plot (averaged across 7 subjects with 95% confidence intervals). Note how the presence of horizontal and vertical components in the responses reflect the physical availability of such global motion vectors in the stimulus. Note also the ‘multistable’ outcome for Fig 1d, as shown in Fig 1h.

Figure 2

Results for the four observers in experiment 1 tested with standard displays in which the context gratings had the same drift speed as the test, and which produced global motion orthogonal to their context-pair axes (as in Fig 1d). a) ‘raster plots’ of cued and reported drift directions (with successive 75hz display frames ordered left to right for each run, and with successive runs shown on separate rows). Blue colouring and the letter ‘R’ in the top panel indicates epochs during which the cue indicated that attention to the contexts on the horizontal axis was required (i.e. the Row of the cruciform configuration); in the lower panels the same colour blue corresponds to frames where

the observer indicated the direction of global motion that would be expected given this cued context, i.e. vertical global motion. Yellow colouring (and the letter ‘C’ in the top panel) corresponds to cueing of the vertical axis (i.e. the Column of the cruciform), while yellow in the lower panels corresponds to reported horizontal global motion. Red colouring in the raster plots indicates that the observer indicated perception of diagonal motion. Any other reported directions are coloured in cyan. b) Radial histograms plotting the distribution of reported directions (across the three bins of vertical, horizontal or diagonal) as a function of cue state (blue for vertical cued motion, yellow for horizontal, using the same scheme as above). Values plotted are proportions of the total number of display frames in each cue state. c) estimates of cued-control for each observer, with 95% confidence intervals based on standard error across epochs.

Figure 3

Experiment 2 stimuli and results. a-c) Schematics of stimulus configurations. Gray diagonals pointing out of each circular grating indicate local motion directions; black arrows indicate global motion directions induced by each axis. Three different stimulus configurations are compared: (a) horizontal and vertical directions of global motion are both available (‘H+V’); (b) just horizontal global motion available (‘H only’); (c) vertical global motion only (‘V only’). d) Measure of ‘Cued-control’ for three observers averaged over epochs, with 95% confidence limits based on standard error across epochs; Values greater than zero indicate that observers tended to report the direction of perceived motion that was expected to be induced by the currently cued flanker axis.

Figure 4

Example of stimulus used in the ‘contrast A’ condition of Experiment 3. Note that gratings are now Gabor patches. Attentional cues are presented close to fixation, in this example directly left and right the fixation point to indicate that the contexts on the horizontal axis should be attended. The cursor dots are illustrated here in the upper and lower periphery, indicating vertical motion. See also animation file **expt3demo.mov**.

Figure 5

Results of eye-tracking in Experiment 3. a) Radial histograms showing the distribution of saccade angles over 8 bins as proportions of the total number of saccades. Results are averaged over contrast conditions and stimulus configurations, but split by the direction of global motion expected to be induced given the current attentional cues (blue segments: horizontal global motion expected; red: vertical motion expected). Note that eye-position did not vary systematically with perceived motion direction. b) Standard deviations of horizontal (X) components of saccades, averaged over cueing epochs and contrast conditions. Blue and red bars show means over 4 observers, for horizontal and vertical cued global motion respectively, with errorbars indicating 95% confidence intervals based on the within-subjects standard error of the means (see main text for further details). Symbols show values for individual observers. Note no systematic relation to the direction of motion expected given the attentional cues. c) Event-related timecourse of eye displacement magnitudes relative to the onset of a cue switch (at time zero on the x-axis), averaged across epochs but split by the direction of global motion that would be consistent with the cued axis (blue: horizontal-motion cued; red: vertical-motion cued). In order to compare conditions where different cueing conditions were consistently associated with reports of horizontal versus vertical motion, the lowest-contrast conditions were omitted from this analysis. Left and right columns of graphs show X and Y coordinates averaged across all epochs respectively for each of the four observers, in rows. Any tendency to pursue or saccade in horizontal or vertical directions dependent on the cue type should appear in all graphs as a transient elevation of the blue trace in the left X-coordinate graphs, and/or a similar elevation of the green trace in the right Y-coordinate graph. However, no such systematic differences are seen.

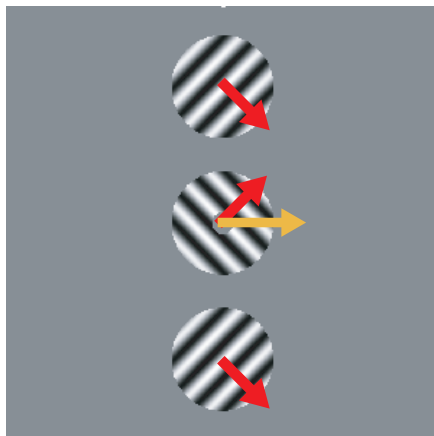
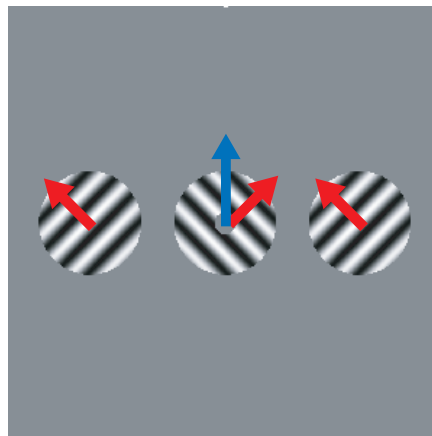
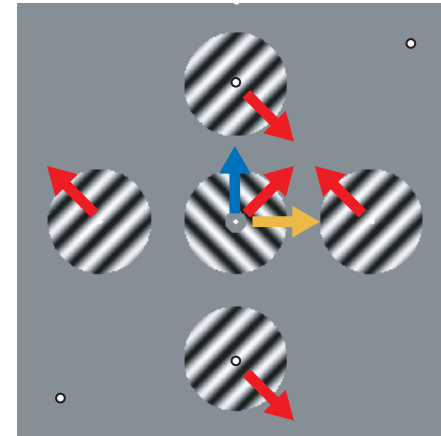
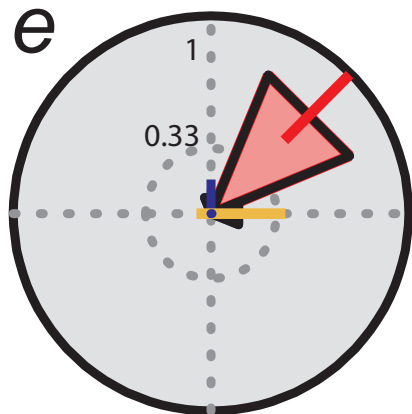
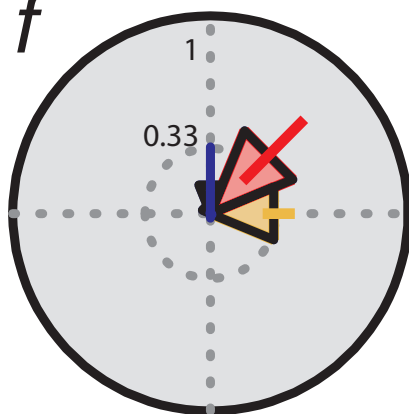
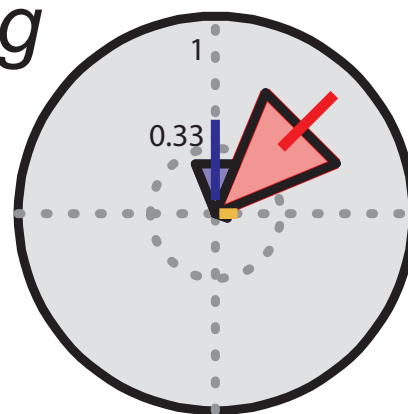
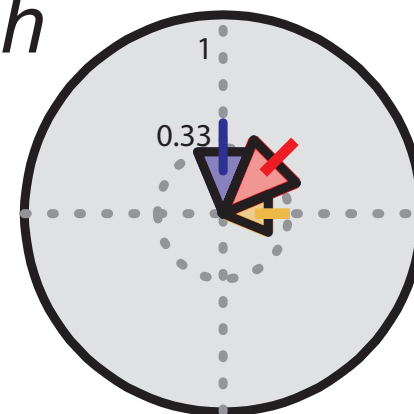
Figure 6

Experiment 3 results pooled across observers, for each of the stimulus configurations (a-d). Schematic representations of the display configurations are shown in the leftmost column, with local motion-directions indicated by gray diagonals pointing out of each circular grating, and global motion directions expected to be induced by attention to one or other pair of context stimuli

indicated by black arrows attached to one or other context axis. Along each row (a-d), radial histograms display the distribution of reported global motion directions for each contrast condition (Equal, Contrast A, or Contrast B). Yellow segments indicate the proportion of responses during cueing to the vertical axis; the blue are for cueing to the horizontal axis. (e-f) relative proportions of responses in three directions visualized as stacked bar charts summarizing the relative proportions averaged across subjects. Separate panels show responses during attentional cueing of the high-contrast axis (left half of graph), versus cueing to the low-contrast axis (right half). Red bars indicate the proportion of local motion reports; gray middle region above red bar indicate global motion reports in the direction expected to be induced by the currently cued axis; green bars at top of graph indicate the remaining proportion of reports reflecting the global motion induced by the uncued axis (this may be quantified by measuring from the top of the graph). Results for individual observers are superimposed as cumulative proportions, representing the breakdown of the mean proportions shown in the background, now shown as individual datapoints. Filled red points indicate proportion of local motion reports; open red/white points indicate proportion of cued global motion reports (*plus* the local motion proportion indicated by the lower solid points). Distance from the top of graph to the open red/white points thus reflects the remaining proportion of responses in the uncued global direction.

Figure 7

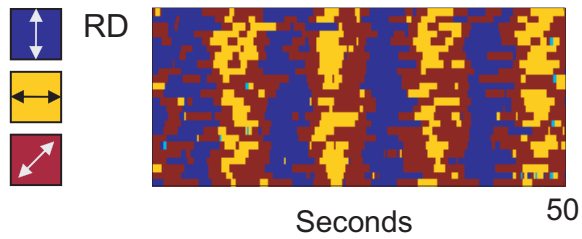
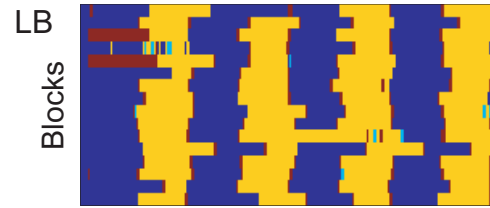
Experiment 3 results for four observers. a-b) Cued control for bi-directional and uni-directional displays respectively; c-d) Dominance of global motion corresponding to the high-contrast (vertical) axis. Blue and red lines (or superimposed filled and open symbols) indicate mean (or individual observers' results) for orthogonal and aligned-motion configurations respectively. Error-bars indicate within-subjects 95% confidence intervals for the averaged results.

a*b**c**d**e**f**g**h*

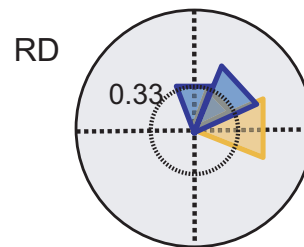
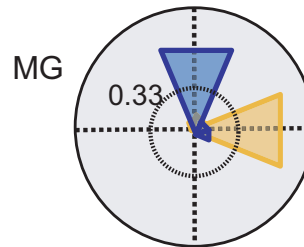
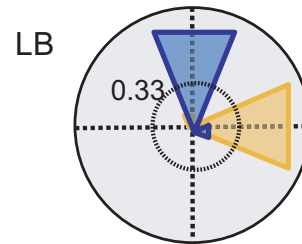
a Cued axis (Row / Column)

• R • C • R • C • R • C • R • C

Reported direction



b Response histograms



c Cued-control

