



City Research Online

City, University of London Institutional Repository

Citation: Freeman, E. D., Macaluso, E., Rees, G. & Driver, J. (2014). fMRI correlates of object-based attentional facilitation versus suppression of irrelevant stimuli. *Perception*, 43(10), p. 1116. doi: 10.1068/ava13xm

This is the accepted version of the paper.

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

Permanent repository link: <http://openaccess.city.ac.uk/13397/>

Link to published version: <http://dx.doi.org/10.1068/ava13xm>

Copyright and reuse: 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.

City Research Online:

<http://openaccess.city.ac.uk/>

publications@city.ac.uk

fMRI correlates of object-based attentional facilitation versus suppression of irrelevant stimuli, dependent on global grouping and endogenous cueing.

Elliot D. Freeman¹

Emiliano Macaluso²

Geraint Rees^{3, 4}

Jon Driver^{4*}

1. City University London
2. Neuroimaging Laboratory, Fondazione Santa Lucia, I.R.C.C.S. Rome
3. Wellcome Trust Centre for Neuroimaging, University College London
4. Institute of Cognitive Neuroscience, University College London

Correspondence to:

Elliot D. Freeman
City University
Northampton Square
London EC1V 0HB
Telephone: +44 (0)20 7040 0102
Fax: +44 (0)20 7040 8580
email: elliott.freeman@city.ac.uk

Abstract

Theories of object-based attention often make two assumptions: that attentional resources are facilitatory, and that they spread automatically within grouped objects. Consistent with this, ignored

* Deceased

visual stimuli can be easier to process, or more distracting, when perceptually grouped with an attended target stimulus. But in past studies, the ignored stimuli often shared potentially relevant features or locations with the target. In this fMRI study, we measured the effects of attention and grouping on Blood Oxygenation Level Dependent (BOLD) responses in the human brain to entirely task-irrelevant events.

Two checkerboards were displayed each in opposite hemifields, while participants responded to check-size changes in one pre-cued hemifield, which varied between blocks. Grouping (or segmentation) between hemifields was manipulated between blocks, using common (versus distinct) motion cues. Task-irrelevant transient events were introduced by randomly changing the colour of either checkerboard, attended or ignored, at unpredictable intervals.

The above assumptions predict heightened BOLD signals for irrelevant events in attended versus ignored hemifields for ungrouped contexts, but less such attentional modulation under grouping, due to automatic spreading of facilitation across hemifields. We found the opposite pattern, in primary visual cortex. For ungrouped stimuli, BOLD signals associated with task-irrelevant changes were lower, not higher, in the attended versus ignored hemifield; furthermore, attentional modulation was not reduced but actually inverted under grouping, with higher signals for events in the attended versus ignored hemifield.

These results challenge two popular assumptions underlying object-based attention. We consider a broader biased-competition framework: task-irrelevant stimuli are suppressed according to how strongly they compete with task-relevant stimuli, with intensified competition when the irrelevant features or locations comprise the same object.

Introduction

In our complex environment, with many different competing stimuli and goals, coherent behaviour demands attentional selection. Such selection may both enhance relevant information and suppress irrelevant information. One important constraint on this selection is the perceptual organisation of the stimulus into groups or proto-objects (Driver & Baylis, 1989; Driver, Davis, Russell, Turatto & Freeman, 2001; Palmer & Rock, 1994; Scholl, Pylyshyn & Feldman, 2001). Classical studies (Baylis & Driver, 1993; Egly, Driver & Rafal, 1994; Kahneman & Henik, 1981) show that it is easier to process, or harder to ignore visual stimuli when they are perceptually grouped with another stimulus at the current focus of attention. A popular explanation is that attention involuntarily spreads within the bounds of an object, automatically facilitating processing of all of its constituent parts and features (Chen, 2012; Driver *et al.*, 2001; Duncan, 1984; Vecera & Behrmann, 2001; Watson & Kramer, 1999; but see Davis & Holmes, 2005). Such spreading may allow focusing activity on a specific object to the exclusion of others, for example eating off our own plate rather than our neighbours. But objects are commonly associated with other objects in a hierarchical structure (Baylis & Driver, 1993; Logan, 1996; Watson & Kramer, 1999), and sometimes we need to select just one individually, for example if we want to eat the peas on our plate but avoid carrots. If spreading of attention were always mandatory and facilitatory, we might find it difficult to 'drill down' to specific sub-objects within a hierarchy, or their specific features, while ignoring others.

Some past research has examined how goal-driven attention or 'perceptual set' (Neisser & Becklen, 1975; Vecera & Behrmann, 2001) might interact with grouping processes to parse the local constituents of a scene into task-relevant global structures (Baylis & Driver, 1993; Freeman & Driver, 2008; Freeman, Sagi & Driver, 2001; Khoe, Freeman, Woldorff & Mangun, 2006; Logan, 1996; Vecera & Behrmann, 2001; Watson & Kramer, 1999). However the extent to which object-based allocation of attention is automatic or more intelligently prioritised continues to be debated (Chen & Cave, 2006; Chen & Cave, 2013; Richard, Lee & Vecera, 2008; Shomstein, 2012; Yeari & Goldsmith, 2010; Zhao, Kong & Wang, 2013). Research on object-based facilitatory spreading of attention also often neglects the 'dark side' of visual attention (Chun & Marois, 2002; Fuentes, Humphreys, Agis, Carmona & Catena, 1998; Tipper, Driver & Weaver, 1991), namely suppression of irrelevant information. This paper examines the factors that may determine whether attentional selection facilitates or suppresses irrelevant features, or events at an irrelevant location of a scene, and how this may depend on the task-driven control of endogenous spatial attention, and stimulus cues for grouping.

Many past results seem consistent with the notion that excitatory attention spreads within the boundaries of a continuous object, but less readily across the gap between separate objects. For example, studies based on the popular cueing paradigm introduced by Posner, Snyder and Davidson (1980) show that performance in discriminating a salient target on one end of a shape is improved if a previous cue has invalidly directed attention to the other end of the same shape, compared to when it cues the second shape (Egly, Driver & Rafal, 1994). This is consistent with attentional spreading within the boundaries of the shape defining the 'grouped array' (Avrahami, 1999; Hollingworth, Maxcey-Richard & Vecera, 2012; Vecera & Behrmann, 2001). Other studies, using variants of the Eriksen flanker paradigm (Eriksen & Eriksen, 1974) show that stimuli seen as belonging to the same object tend to be processed automatically, leading to response competition (Baylis & Driver, 1992; Driver & Baylis, 1989; Kramer & Jacobson, 1991; Kahneman & Henik, 1981; Zhao *et al.*, 2013). Several studies (He, Fan, Zhou & Chen, 2004; Martinez, Ramanathan, Foxe, Javitt & Hillyard, 2007; Martinez, Teder-Sälejärvi, Vazquez, Molholm, et al., 2006; Müller & Kleinschmidt, 2003) found complementary effects in EEG and fMRI using similar stimuli and or cueing, for example showing an increased neural response to the unattended end of a rectangle when the other end was attended, relative to when another separate rectangle was cued.

A common claim in many of the above studies is that the grouped stimulus benefits from attentional spreading despite being irrelevant or even potentially disruptive to the task at hand. Such irrelevance is essential to the assumption that such spreading is based on automatic pre-attentive grouping processes (Chen, 2012; Davis, Driver, Pavani & Shepherd, 2000; Kramer, Weber & Watson, 1997; Vecera & Behrmann, 2001; Vecera & Farah, 1994; Weber, Kramer & Miller, 1997; Yeari & Goldsmith, 2010; Zhao *et al.*, 2013). This claim may be challenged on two fronts. Firstly, past behavioural studies based on the Posner cueing paradigm have used cues that were intentionally unreliable, with the result that a task-relevant target could sometimes appear on uncued parts of the stimulus (i.e. in invalid-cue trials). In such situations of unreliable cueing, any evidence of apparent object-based attentional spreading might reflect a prior attentional set (or 'attentional prioritisation') that preferentially includes grouped regions of the cued stimulus where targets are expected to appear, even if infrequently (Müller & Kleinschmidt, 2003; Shomstein & Yantis, 2002; Shomstein, 2012). Secondly, in paradigms based on the Eriksen interference, the irrelevant flankers necessarily share potentially task-relevant properties with the target in order to provide measurable response interference. Such stimuli might attract attention due to their similarity (Baylis & Driver, 1992; Harms & Bundesen, 1983; Kim & Cave, 2001), perhaps via feature-based attention (Martinez-Trujillo & Treue, 2004; Saenz, Buracas & Boynton, 2002) or template-matching mechanisms (Duncan & Humphreys, 1989), while grouping processes might constrain

such feature-based selection (Festman & Braun, 2012; Melcher, Papathomas & Vidnyánszky, 2005), resulting in what appears as involuntary attentional facilitation. It therefore remains possible that under the right circumstances, facilitation might be eliminated or even become inhibitory if the probe stimuli are perfectly irrelevant and have nothing in common with the target stimulus.

Behavioural paradigms often suffer the limitation that any effect on 'irrelevant' stimuli can only be assessed through overt responses, which may thus prime or direct attention to them. While some neurophysiological studies have just replicated the classic paradigms (e.g. He *et al.*, 2004; Müller & Kleinschmidt, 2003), others have taken advantage of the possibility of measuring the brain's implicit responses to irrelevant stimuli. For example, Martinez *et al.* (2006, 2007) found enhanced evoked potentials for an irrelevant probe transient when participants were attending similar target transients on the opposite side of a rectangle figure. However, facilitation might still in principle have been caused by the activation of a 'template' feature for detecting target transients. Another recent EEG study used frequency tagging to track flicker-evoked neural activity associated with a central target and a surround which either formed a continuous grating pattern with the centre, or was segmented by a gap or phase offset (Kim & Verghese, 2012). Participants judged threshold increments in the contrast of the central component. This did not vary in the surround, which was thus entirely irrelevant. Surprisingly, activity to the surround was actually lower for the continuous pattern compared to discontinuous. This result would be consistent with suppression, rather than activation of wholly irrelevant grouped areas, perhaps as a result of more focal spatial attention (Kim & Verghese, 2012). However, the gap or phase manipulation may have introduced low-level features that modulated the amount of reciprocal centre-surround inhibition, which in turn may have been amplified by a spread of attention across the surface. Manipulation of global rather than local grouping cues may be preferable to avoid this ambiguity.

While the above studies typically manipulated grouping between patterns occupying different regions of space, other studies have attempted to control spatial attention to obtain a purer measure of the capabilities of object-based attention in selecting one group in the presence of a second overlapping pattern (e.g. Duncan, 1984). Many studies manipulated grouping via spatio-temporal cues such as common-fate motion (e.g. Jarmasz, Herdman & Johannsdottir, 2005; O'Craven, Downing & Kanwisher, 1999), for example using transparent coherently-moving dots (Ciaramitaro, Mitchell, Stoner, Reynolds & Boynton, 2011; Mitchell, Stoner & Reynolds, 2004; Schoenfeld, Tempelmann, Martinez, Hopf, et al., 2003; Valdes-Sosa, Bobes, Rodriguez & Pinilla, 1998). Here selection of one feature tends to activate selection of other features that are bound to the same object, even if they are too faint to be consciously perceived (Melcher & Vidnyánszky, 2006), and

may strongly suppress physiological responses to the irrelevant object (Valdes-Sosa *et al.*, 1998). Again it is often claimed from this that the spread of attention is inevitable, and facilitatory to all features belonging to the object that is attended. In support of this a recent fMRI study found selective enhancement of frequency-tagged irrelevant features belonging to a relevant dot pattern, in context of irrelevant overlapping dot pattern (Ernst, Boynton & Jazayeri, 2013). However, a behavioural study found that cueing of a feature results in selective speeding of responses to it, but did not facilitate responses to other irrelevant features belonging to the same object (Wegener, Ehn, Aurich, Galashan & Kreiter, 2008). One critical difference between this study and others, as Wegener *et al.* (2008) suggest, may be that the target stimulus did not overlap with a distractor stimulus (see also Davis & Holmes, 2005, for consideration of further stimulus factors which may determine within-object benefits versus costs). In cases of overlap (e.g. Ernst *et al.*, 2013; Ciaramitaro *et al.*, 2011; Jarmasz *et al.*, 2005; Mitchell *et al.*, 2004; O'Craven *et al.*, 1999; Schoenfeld *et al.*, 2003; Valdes-Sosa *et al.*, 1998), we might accrue evidence from any available redundant cues, such as a contrasting colour or sudden change in motion trajectory of the targets dots, that might help to uniquely distinguish the target features from features belonging to an overlapping distractor, even if they are not consciously detected (cf. Melcher & Vidnyánszky, 2006). Such additional cues are of less relevance when there are no confusable features within the spotlight of spatial attention, and might thus be ignored more effectively when they are irrelevant. Thus, while object-based spreading of attention may appear mandatory in the above studies using overlapping stimuli, there might be less attentional spreading if the stimuli did not need to be segmented in order to be selectively attended.

Commented [EF1]: Revision 3

It might be concluded from this previous work that facilitation across features within an object is not mandatory but dependent on the need to segment a target from competing overlapping or surrounding features when present. However it is not yet clear whether object-based attentional allocation to completely irrelevant stimuli is also optional across space to similar features of different stimuli, whether it is always facilitatory, and how this might depend on global (rather than local) grouping cues. To address these issues, we used fMRI to measure neural activity associated with any interactions between global grouping (by common-fate motion) between two stimuli in opposite visual hemifields. We measured the effect of attending to subtle targets in one versus the other hemifield, while also independently measuring BOLD signals evoked by highly salient but completely irrelevant transient colour changes within the attended or ignored hemifield. We focused our analyses on relevant areas of motion-sensitive or retinotopic early visual cortices (see Methods) and assessed whether effects of attention and grouping mostly affected feature-specific activity in colour and motion related areas (i.e. V4 and V5/MT respectively), and/or whether there were more

general effects on early visual cortices (e.g. see Ciaramitaro *et al.*, 2011). If object-based attentional spreading is generally mandatory and facilitatory both within and between objects, irrelevant transients should always evoke an increase in signals associated with attended stimuli, and also in the opposite hemifield specifically when it is grouped with the attended hemifield. More generally, attention-related activations (associated with the blocked manipulation of hemifield cueing) should leak over to the unattended hemifield under grouping. The contrasting hypothesis is that attentional spreading may be less facilitatory or even inhibitory towards features that are entirely irrelevant to the task and of no use for image segmentation.

Commented [EF2]: Revision 2

Methods

Participants

Eight participants aged between 25 and 35 participated with their informed consent. All had normal or corrected-to-normal vision, and had previous experience of imaging experiments, but were naïve to the purpose of the present study. The experiment was approved by the local ethics committee.

Figure 1

Stimuli

An LCD projector back-projected stimuli onto a screen at the rear of the magnet bore. Video mode was 640 x 480 with screen refresh rate of 60Hz, and output was linearised using 8-bit software gamma-transformation. Observers lay supine in the scanner, and viewed the screen via a mirror mounted on the head coil, across a total viewing distance of 62cm. Stimulus presentation and timing was controlled by a PC running MATLAB (Mathworks Inc.) and COGENT 2000 toolbox (<http://www.vislab.ucl.ac.uk/Cogent2000/>). The visible display subtended visual angles of 31deg horizontal by 14deg vertical. Displays were composed of two light and dark diagonally oriented checkerboards (each composed of the product of two orthogonal oblique sinusoidal gratings of wavelength 4.84deg, thresholded with light colouring for positive values and black for negative). This resulted in diamond-shaped checks with edges measuring 2.40deg in length. Checkerboards were presented on a mid-grey screen on each side of the vertical midline, visible through 90° segments of a central annulus-shaped sharp-edged window, with inner and outer diameter of 2.82deg and 9.8deg respectively (see Figure 1). Each checkerboard translated behind the window along a circular path of radius 1.5deg, taking 2 seconds to complete each cycle. Grouping was manipulated by moving left and right hemifields either in phase with each other or 90° out of phase. In-phase motion produced the impression of a continuous checkerboard surface passing behind left and right apertures (see Movie 1); out-of-phase motion gave the appearance of two independent checkerboards (Movie 2).

Fixation stimuli consisted of small '<' and '>' characters at the centre of the visible display, one red and the other green interchangeably. On every 2 second motion cycle, synchronously but independently on both sides, the light checks on either the upper or lower quadrant smoothly expanded in size by a maximum of 0.2deg, while the dark checks contracted in size by the same amount, with maximum size change at mid-cycle, returning again to their original size at the end of the cycle. This was achieved by adding a 2D Gaussian function to the combined oblique gratings

composing the chequerboard (SD 1.2deg, positioned ± 1.7 deg horizontally and ± 0.68 deg vertically relative to the fixation point), prior to thresholding, and modulating the amplitude of this function with a Gaussian temporal profile (SD .33 sec). Participants were required to discriminate between upper and lower check-size changes on the hemifield pointed to by the red fixation arrow, while ignoring all changes on the opposite hemifield (which were anyway uninformative). Participants made 'up' or 'down' responses using one of two keys on an optical fibre button-box. The importance of responding on every motion-cycle 'trial' was emphasised. Eye position data were sampled at a frequency of 60Hz during scanning using remote-optics infrared eye tracking (ASL 504, Applied Science Laboratories, Bedford, MA). The importance of maintaining fixation on the central arrow stimuli was emphasised to participants.

For most of the time, the checkerboards were coloured green on a black background, but would occasionally flash red on one hemifield or the other for a duration of 500ms. These events were not temporally correlated with the check-size changes. Subjective red-green isoluminance was established for each participant prior to scanning, using method of adjustment to minimise perceived flicker of a 30Hz alternating red-green checkerboard.

Design and procedure

Participants first attended a half-hour training session in a psychophysics laboratory. They were familiarised with the task and were given verbal feedback on their eye-movements during the task. The maximum check-size change determined task difficulty, and a method of constant stimuli was used to find the 85% accuracy level for each participant, which was used throughout the scan. In the scanner, participants first completed the isoluminance adjustments and eye-tracker calibration. There then followed 10 four-minute scans. Each of these runs was divided into four one-minute blocks, presented in counterbalanced order.

Each block represented a different crossing of two independent variables: Grouping (in-phase or 'Grouped' motion versus out-of-phase or 'Ungrouped' motion), and Attention (to left versus right hemifields). This 2x2 block design was superimposed on an event-related design, in which transient red flashes occurred independently on left and right hemifields, every 2 to 12 seconds for a period of 500ms (e.g. Fig 1b). In each block, five flashes would occur on each hemifield, independently and unpredictably. A given flash, therefore, could be classified as occurring on an attended side or an unattended side, and on a checkerboard that was either grouped or segmented from its opposite counterpart.

Neuroimaging

BOLD contrast fMRI images were acquired in a Siemens Allegra 3 Tesla MRI scanner (Siemens, Erlangen, Germany), using an EPI sequence. Slices were positioned to cover the whole brain. Voxel size was 2x2x2mm. There were 10 scanning runs for each participant, each lasting 4 minutes 40 seconds, and consisting of 85 volumes sampled with repetition time of 3.12 sec and 48 slices per volume. Volumes had 48 slices of 2mm thickness with a 1mm gap between slices, giving a resolution of 3x3x3mm. We also acquired T1-weighted MPRAGE images for structural analysis with a resolution of 1x1x1mm.

Localization

Retinotopic visual areas (i.e. V1, V2, V3, V3a, and V4) were each identified on the basis of standard rotating-wedge scans conducted in a prior session, with segmentation and cortical flattening using MrGray software (Teo, Sapiro & Wandell, 1997; Wandell, Chial & Backus, 2000). These retinotopic regions of interest (ROIs) were then inclusively masked by t-maps representing voxels that were significantly activated (at $p < .05$ uncorrected) by the stimuli across all blocked conditions. We identified regions of interest corresponding putatively to motion-sensitive areas, for individual participants based on voxels showing significant activation ($p < .001$ uncorrected) across all blocked conditions, whose coordinates were consistent with the published location of area hMT/V5+ (e.g. Hasnain, Fox & Woldorff, 1998; Tootell, Reppas, Kwong, Malach, et al., 1995; Watson, Myers, Frackowiak, Hajnal, et al., 1993). As the above localisation analyses were based on BOLD signal averaged across all block-related conditions (attention left vs. right, and grouped versus ungrouped) this method of localization could not bias the outcome of our tests for hypothesised differences between conditions, either block-related or event-related. This method of defining a region of interest, based on a contrast that is orthogonal to those used to test an experimental hypothesis, is an established approach in the literature (Friston, Rotshtein, Geng, et al., 2006). We used our own circularly translating stimuli, rather than a traditional independent motion localiser based on moving random-dot kinematograms, as this could isolate regions sensitive to the specific type of motion used in our main experiment, providing a principled and statistically independent way to identifying relevant voxels that might be subject to our particular modulations of spatial attention and stimulus grouping. However, given that these areas were not identified using standard functional localisers, the label 'hMT/V5+' is used tentatively.

Commented [EF3]: Revision 5

Commented [EF4]: Revision 1

Figure 2

Eye-movement control

Prior to fMRI analysis, we used eye position data from eye tracking to control for the possibility that attentional cueing to left and right hemifields, and our manipulation of grouping, could systematically affect participants' fixation patterns. For each run, eye-tracker data (X and Y coordinates for each 16.6ms acquisition frame) were processed to remove any linear trend, and filtered to exclude blinks or signal drop-outs. Frames in which there was a horizontal deviation from central fixation of greater than 2 degrees were then identified, which might be caused if participants made saccades towards one of the hemifield stimuli. From these we derived a measure of fixation bias towards the attended hemifield, for each scanning run in each participant, by subtracting the proportion of fixation deviations away from the cued hemifield, from the proportion of deviations towards the cued hemifield. This bias measure was therefore positive when participants made more saccades of greater than 2 degrees horizontally towards the cued hemifield. The distribution of this measure over runs and participants had a long tail towards higher values (mean 0.0084, SD 0.188, skewness 3.63), consistent with the occasional tendency to peek at the hemifields containing the task-relevant targets. We attempted to correct this by excluding individual runs in which this fixation bias measure had values of greater than 0.01. This resulted in omission of 23% of runs on average across participants (SD 18%), and a more symmetrical distribution of gaze bias scores (mean 0.0017, SD 0.0048, skewness -1.77) and mean gaze locations (see Figure 2, plotting frequency of horizontal gaze locations towards versus away from the cued hemifield, before and after the above correction for bias, for the two grouping conditions separately). Following this correction for bias, we compared the effects of left/right cueing and grouping on mean horizontal gaze coordinates based on ten scanning runs per subjects, in a two-way repeated measures ANOVA. Results showed no significant bias towards the cued hemifield [$F(1,7) = 3.32$, $p = .11$, no main effect of grouping [$F(1,7) = 0.43$, ns] and no significant interaction [$F(1,7) = 0.002$, ns].

Commented [EF5]: Revision 4: clarifications in the following paragraph

fMRI Analysis

Preprocessing and analysis of fMRI images was conducted using SPM2 (<http://www.fil.ion.ucl.ac.uk/spm>). The first five images of each scanning run were discarded to allow for magnetic saturation effects. The remaining images were realigned and coregistered to the individual participants' structural scans for analysis of early retinotopic areas. A high-pass filter was applied at 0.0078 Hz to remove low-frequency signal drifts. For whole-brain analysis, images were spatially normalised into standard space (MNI) and spatially smoothed with a Gaussian kernel of 8mm FWHM.

Commented [EF6]: Revision 6

For each participant, data were entered into a general linear model (Friston, Holmes, Worsley, Poline, et al., 1994) specifying blocked variables and transient events as separate regressors convolved with a canonical HRF, within the same model. Each model therefore had four regressors for each of the four block types in the 2x2 (attention x grouping) design, in addition to eight further event-related regressors corresponding to the four conditions of the 2x2 design for the left and right transient events independently. Whole-brain random-effect analyses were then performed using one-sample t-tests, to assess the statistical significance of selected contrasts across participants. Block-related contrasts compared BOLD signals for left versus right attention, and grouped versus ungrouped stimuli. Event-related contrasts compared left versus right transients, in the context of grouped versus ungrouped stimuli. We also tested contrasts which assessed the hypotheses that the difference between left and right blocked attention, or signal evoked by left versus right transients, was greater (or smaller) under grouped versus ungrouped conditions.

Results

Behavioural data

One participant failed to respond on 16% of trials, compared to an average failure rate across the remaining participants of only 0.75% (SD 0.5%). Proportion correct was calculated for up-down discrimination of check-size changes after filtering out missed trials. Mean accuracy across all participants was 91% (SD 3%). The same participant with high miss-rates also had the poorest accuracy (87%). After excluding this participant, the filtered accuracy data were analysed in a repeated-measures ANOVA, with attended hemifield (Left vs. Right) and Grouping (Grouped versus Ungrouped stimuli) as repeated-measures factors. There was a significant main effect only for Attention, with higher accuracy for discriminating check-size changes in the left hemifield [$F(1,7) = 9.75, p=.02$]. Mean (and SD) for left was 94% (0.3%) and for right, 91% (1%). There was no significant interaction with grouping. A similar pattern was observed with the full data set.

Figure 3

Whole brain analyses

Statistical contrasts of left versus right cued attention revealed significant activations (family-wise corrected for multiple comparisons at $p<.05$) contralateral to the cued hemifield in cuneus and lingual gyrus (see Table 1 for coordinates). Lowered thresholds ($p<.001$ uncorrected, see Figure 3, left) revealed widespread activations only in posterior visual areas contralateral to the cued hemifield. Event-related contrasts of left versus right transients showed significant activations contralateral to the transient (corrected for multiple comparisons at $p<.05$) in lingual gyrus and fusiform gyrus (Table 1). At lower thresholds (e.g. $p<.001$ uncorrected, Figure 3, right) activations were seen in occipital inferior and superior occipital areas putatively within area V3a (Tootell *et al.*, 1995). Contrasts of grouped versus ungrouped stimuli revealed no notable activations even at lowered thresholds ($p<.001$ uncorrected), for either event-related or block-related analyses. There were also no significant results for whole-brain block and event-related analyses of specific interactions between grouping and attention.

Retinotopy: Blocked analyses

Beta weights (representing the overall level of BOLD activation) for each of the block-related conditions were estimated from each of the ROI's for each participant (i.e. masked retinotopic visual areas including hMT/V5+, see Methods). As we had no specific hypotheses about the laterality of attentional effects, data from left and right hemispheres were pooled according to

whether the respective contralateral visual hemifield was attended or ignored. Data were entered into a 3-way repeated measures ANOVA with the following factors: Attention (whether a given ROI was contralateral versus ipsilateral to the attended hemifield), Grouping, and cortical Area.

Initial analysis revealed a main effect of Attention, with larger BOLD signals when attention was cued to the contralateral versus ipsilateral hemifield [$F(1,7) = 70.69, p=.0001$]; and a main effect of Grouping, with greater signals for ungrouped than grouped stimuli [$F(1,6) = 9.29, p=.019$]. There was a main effect of Area [$F(5,35) = 23.86, p<.0001$], which was partially accounted for by higher signal estimates in hMT/V5+ (9.87, SD 0.14) compared to the other areas (mean 4.86, SD 0.87). Variability was also much higher in MT/V5+ (SD 0.14) compared to other areas (mean SD 0.022, SD 0.017). The only significant interaction was between Attention and Area [$F(5,35) = 4.11, p=.005$], with larger effects of attention to the contralateral versus ipsilateral hemifield in hMT/V5+ than in the other visual areas (see Supplementary Figure 1).

Figure 4

All further analyses excluded the participant with poor hit rates, in case this was indicative of poorly controlled attention. To render the variances between all ROI areas more uniform, we normalised each participant's beta estimates for each given condition from each given ROI (pooling data across hemispheres, as described above), by subtracting the average (and dividing by the standard deviation) of block-related beta estimates obtained across all conditions from the same bilateral ROIs. Normalised data now had similar means and ranges across ROIs and participants, while the detailed pattern of results across conditions remained unchanged within each ROI. An ANOVA based on the normalised data now revealed a significant interaction of Attention and Grouping [$F(1,6) = 8.06, p=0.03$], with greater attentional modulation in the grouped compared to grouped condition (see Figure 4). These analyses also confirmed the main effect of Attention [$F(1,6) = 705, p<.0001$] and Grouping [$F(1,6) = 22.12, p=.003$], and no main effect or interaction for Area.

Commented [7]: Revision 10

Commented [8]: Revision 7

Figure 5

Separate analyses were conducted for each visual area showed that the interaction between grouping and attention was significant in V2 [$F(1,6) = 13.62, p=.01$], and borderline significant in V1 [$F(1,6) = 5.40, p=.059$] (see Figure 5). A similar pattern was observed for the non-normalised

data [V1: $F(1,6) = 5.18, p=.063$; V2: $F(1,6) = 9.05, p=.023$] (Supplementary Figure 1). Including the participant with high miss-rates rendered these effects non-significant.

Event-related analyses

In the event-related analyses, beta weights were estimated from each ROI for each participant. As we had no specific hypotheses about the left versus right location of the transient stimuli, transient events were coded according to whether they appeared in the attended versus ignored hemifields, and whether they appeared in the hemifield contralateral or ipsilateral to the ROI. A four-way ANOVA included these two factors (Attention and Transient Location) along with Grouping and cortical Area.

Initial analysis of the whole sample (excluding the participant with poor hit rates) revealed only the main effect of contralateral versus ipsilateral Transient Location as significant. Examination of the raw data revealed one participant with highly disparate beta estimates in particular ROIs and conditions (contributing to an increased range of betas ± 15 compared to ± 5 for other participants, and increased standard deviation of 3.5, while others varied between 1 and 2.3, resulting in a z-score of 2.13 with respect to the whole sample).

Similar to the block-related analysis, we normalised the data for each ROI to render the variances between participants and areas more uniform. This was done for each participant, and for each pair of bilateral ROIs, by taking the beta estimates obtained for each condition, subtracting the average beta across all conditions obtained from the same bilatero ROIs (pooling data across hemispheres, as before), and then dividing by the standard deviation of that same sample. All participants and ROIs now had data varying across a similar range, while the fine pattern of results across conditions within ROIs was not affected. ANOVA based on these data showed a significant main effect of Transient Location, where contralateral transients produced greater activation than ipsilateral transients [$F(1,6) = 82.44, p=.0001$]. There was a significant interaction between Cortical Area and Transient Location, with apparently less difference between the response to ipsilateral and contralateral events in V5 relative to other areas [$F(5,30) = 4.60, p=.003$]. The only other significant interaction was between Area, Grouping and Attention [$F(5,30) = 2.77, p=.036$].

Commented [9]: Revision 10

Figure 6

To further explore this latter interaction we analysed normalised results for each visual area (see Figure 6), in separate three-factor ANOVA's for attention, grouping, and contralateral versus

ipsilateral hemifield (excluding the participant with high miss-rates). The main effect of contralateral versus ipsilateral hemifield was significant in all areas except V5 [V5: $F(1,6) = 0.29$, ns; other areas: $F(1,6) \geq 43.80$, $p \leq .0006$]. The interaction between Grouping and Attention was significant in V1 only [$F(1,5) = 9.51$, $p = .022$]. As shown in Figure 6, this interaction took the form of a cross-over interaction: under grouped conditions, transients evoked stronger BOLD signals when occurring within the attended versus unattended hemifield; conversely in the ungrouped condition, transient events evoked more response in the unattended compared to the attended hemifield. To take a specific example, when participants attended left, a left transient was associated with a stronger response than a right transient, in the grouped context, but a weaker response compared to a right transient in the ungrouped context (with the complementary situation occurring for right transients under attention to the right hemifield).

Figure 7

Similar but non-significant trends [$F \leq 1.38$] for the interaction between Attention and Grouping were apparent in areas V2, V3 and V4 (Figure 7). There were no other significant effects, and in particular no significant trends from the two-way or three-way interactions, to indicate that the spread of activations between contralateral and ipsilateral ROIs was significantly dependent on grouping [all $p \geq .1$].

A similar pattern of results and in particular the significant V1 interaction between Grouping and Attention were also observed in the non-normalised data [$F(1,5) = 7.82$, $p = .038$] (Supplementary Figure 2), after excluding the participant with unusual variance.

General Discussion

To test the predictions of object-based attention theories, here we measured the BOLD signals evoked by a completely irrelevant transient colour change, as a function of grouping between attended and to-be-ignored hemifields. We observed a cross-over interaction between grouping and attention-related effects on BOLD signals evoked by the task-irrelevant transients, which was significant in retinotopic area V1. When both hemifields were grouped by common-fate motion, the transient stimulus evoked greater activation when it appeared in the attended hemifield compared to unattended. Conversely, when each hemifield moved independently, the transient evoked greater activity when it appeared in the ignored hemifield compared to attended. As explained below, our

findings challenge the assumptions often made in studies of object-based attention, that attentional spreading within and between objects is usually automatic and facilitatory.

Many studies of object-based attention have concurred that spatial attentional influences spread automatically between parts of a display which are grouped, and that the effect of this spreading is largely a facilitation of the processing of these parts (Driver & Baylis, 1989; Egly *et al.*, 1994; Hollingworth *et al.*, 2012; Kramer & Jacobson, 1991; Martinez *et al.*, 2007; Martinez *et al.*, 2006; Müller & Kleinschmidt, 2003; Zhao *et al.*, 2013). If facilitation were generally the case, we should have observed an increase in activation evoked by transients in the unattended hemifield specifically when this hemifield was grouped with the attended hemifield, compared to when it was moving independently. We found the opposite: a decrease in activation under grouping compared to no grouping. Furthermore, if object-based spreading were always automatic, there should have been no effect of manipulating attention to different parts of a grouped display, yet V1 showed a clear benefit for transients which were part of the attended versus ignored hemifield.

Our results further challenge the common assumption that facilitatory attentional influences spread automatically between features bound *within* a single object (Ciaramitaro *et al.*, 2011; Duncan, 1984; Ernst *et al.*, 2013; Jarmasz *et al.*, 2005; Mitchell *et al.*, 2004; O'Craven *et al.*, 1999; Schoenfeld *et al.*, 2003; Valdes-Sosa *et al.*, 1998). If always facilitatory, no specific difference in the response to transients due to grouping should have been expected for stimuli currently under the focus of attention, as the grouping manipulation should not have affected local binding of target and transient features. However, in our results, facilitation of irrelevant within-object features was observed only when the stimulus was grouped with the opposite hemifield, with eliminated facilitation when the attended stimulus was segmented from its counterpart by motion cues. Furthermore, if spreading were always automatic, it should also not depend on the allocation of spatial attention to different parts of a group, but here we observed that facilitation to transients increased under spatial attention with grouped stimuli (and decreased under spatial attention, without grouping). We next consider the factors that might underlie this pattern of results.

Comparison with previous studies

Our study has much in common methodologically with previous studies discussed above, which generally required an attentionally demanding discrimination of subtle features belonging to one object, while ignoring other features belonging to the same and/or different objects. For example, we used a spatial discrimination of check-size modulations in upper versus lower quadrants of the cued hemifield, to encourage subjects to spread their attention over the whole hemifield stimulus

rather than focusing exclusively on one location. This is analogous to paradigms based on Egly et al (1994) in which subjects must discriminate an event such as a luminance increment, that can occur unpredictably on one or other end of a stimulus shape. Note that the luminance modulations defining the target and distractor check-size events in our experiment could not have confounded our event-related measures of the response to transient colour-flash events, which were presented bilaterally at temporally uncorrelated periods during the trial, rather than synchronised with the check-size changes.

Commented [EF10]: Revision 9

One critical difference with previous work is that here we measured the effects of attentional spreading to completely irrelevant transients. In our paradigm, the transient red flashes were entirely irrelevant to the task of discriminating subtle upper versus lower quadrant check-size changes, and contained no features that were confusable with the target. This contrasts with many of the above studies, in which the stimulus to which attention spreads is either not entirely irrelevant, or shares some common features with the relevant target. This methodological difference might help to explain why the present results did not show any evidence of facilitatory attentional spreading between grouped stimulus parts, but rather indicated a reduction of facilitation (or suppression) under specific attention and grouping conditions.

Another methodological contrast is that previous studies indicating automatic and facilitatory spreading of attentional resources between features within an attended object have often presented targets stimuli transparently overlapping with distractor stimuli (Ernst *et al.*, 2013; Melcher *et al.*, 2005; O'Craven, Downing & Kanwisher, 1999; Valdes-Sosa *et al.*, 1998), thus creating a segmentation problem. To resolve this problem, one strategy might be to recruit information from other (nominally irrelevant) features of the target stimulus that are uniquely associated with the target (Wegener *et al.*, 2008). Here, our use of opaque stimuli removes this particular segmentation problem, at least in the ungrouped condition. This could account for the lack of facilitatory spreading between relevant and irrelevant features within the attended hemifield, observed in the ungrouped condition in the form of a reduced response to transients in the attended hemifield (e.g. see Fig. 6). However in the grouped condition, an analogous segmentation problem might arise, where the target hemifield must be distinguished from the opposite hemifield with which it is grouped. In this case selective attention to the target in one hemifield may be aided by spreading facilitatory attention to the transient cues, which help to define the context of the hemifield in which the target appears. Such attentional facilitation of irrelevant features belonging to the target may thus help to resolve the segmentation problem and improve selection of the task-relevant features,

and could explain the observed attentional spreading to transients within the attended hemifield, specifically under grouping (Fig. 6).

Commented [EF11]: Revision 8

The need for segmentation might also explain the apparently similar findings of Kim & Verghese (2012), who reported a reduction rather than an increase of the physiological response to a completely irrelevant surround stimulus in the presence of a grouped versus segmented central target stimulus. They proposed that the demanding central contrast detection task required a withdrawal of spatial attention from the irrelevant surround specifically under conditions of grouping, where presumably the surround becomes more distracting. A similar account might also explain the present block-related results (Fig. 4), showing specifically decreased BOLD signal in the unattended hemifields under grouping compared to no grouping. However an alternative account of Kim & Verghese's (2012) result is that attention modulated inhibitory lateral interactions between the closely abutting centre and surround stimulus. In common with some other studies manipulating local stimulus features (e.g. Murray, Kersten, Olshausen, Schrater & Woods, 2002; Altmann, Bühlhoff & Kourtzi, 2003; Egly *et al.*, 1994; Martinez *et al.*, 2007; Martinez *et al.*, 2006), this creates a potential ambiguity over whether top-down or local ('horizontal') interactions between stimulus parts may be the neural substrate of attentional spreading. Local spreading might function via a mechanism of 'incremental grouping' (Roelfsema, 2006) whereby signals are transmitted between visual areas along object boundaries via excitatory horizontal lateral interactions (Avrahami, 1999). These lateral interactions may themselves be gated by attention (De Meyer & Spratling, 2009; Freeman *et al.*, 2001). Such local interactions were controlled in the present study because hemifields were always separated by a gap (2.66 deg), while the local structure and motion of each hemifield was identical under both grouping conditions. By manipulating only the spatio-temporal relationship between the hemifields, any resulting differences in BOLD response in early visual cortex may be more readily attributable to top-down grouping mechanisms associated with differences in perceived grouping between hemifields, rather than the local structure of visual stimulation presented within each hemifield.

An account based on top-down influences receives further support from our observed grouping-by-attention interaction even within visual areas ipsilateral to the transient stimulus. This effect was independent from a significant effect of contra>ipsilateral areas for transient events. Thus while the stimulus-driven response to transients remained strongly localised to contralateral areas, this was observed in the context of general increases or decreases of BOLD signal in contralateral and in unstimulated ipsilateral regions, which depended on grouping and whether the transient was part of an attended or unattended stimulus. This pattern is consistent with a combination of localised

bottom-up activation and spatially undifferentiated top-down feedback. Such global attentional effects have often been observed in studies of feature-based attention, where attention to specific features modulates brain activity across the visual field (Martinez-Trujillo & Treue, 2004; Saenz, Buracas & Boynton, 2002).

Kim & Verghese (2005) found widespread EEG correlates throughout visual cortex, while here the effects were significant only in V1. The involvement of object-based effects in such an early retinotopic area concurs with another recent fMRI study (Ciaramitaro *et al.*, 2011), and is also consistent with another report that attentional spreading across motion and colour features may depend on primary representations of spatiotemporal correspondences between these features, which could be represented at early stages of visual processing (Melcher *et al.*, 2005). We observed similar trends in visual areas V4 and V5/MT, but these were non-significant, possibly due to insufficient statistical power. Such trends might reflect specific modulation of colour and motion representations, or more general feed-forward effects of selection imposed in V1.

Theoretical proposals

If the classical assumptions of facilitatory and automatic object-based attention cannot explain our findings, what are the alternatives? One possibility is that grouping cues can constrain not only attentional facilitation of potentially relevant stimuli, but also suppression of non-target features when they are completely irrelevant (Fuentes *et al.*, 1998; Tipper *et al.*, 1991). For example, when the stimuli are segmented by motion into two separate objects, relevant target features of the currently attended object can be selected, while the irrelevant transient belonging to the same object is suppressed; however this suppression may be constrained by the boundaries of the attended object and thus does not spread to the transient belonging to the ignored object, which still evokes a cortical response. Consistent with this, suppression did also affect the irrelevant hemifield in the grouping condition. However puzzlingly, transient stimuli in the relevant hemifield now evoked stronger responses than with ungrouped stimuli. This apparent loss of within-object feature selectivity is difficult to explain under the above account of object-based suppression alone. This discrepancy might be explained with the additional assumption (discussed earlier) that the grouping condition in our study creates a segmentation problem (analogous to that encountered with overlapping stimuli; see Wegener *et al.*, 2008). In this case selective attention to the target on one hemifield may also be aided by nominally irrelevant transient cues, which define the context of the hemifield in which the target appears.

Commented [EF12]: Revision 8

The observed combination of facilitatory and suppressive pattern of results is also consistent with the theory of biased competition (Desimone & Duncan, 1995) between 'objects', as defined by grouping cues and including the component parts and features (Vecera & Behrmann, 2001). According to this framework, different objects compete for representation, and top-down bias can be applied to allow one selected object to win this competition (while the competitors are simultaneously suppressed). Our results might be explained by additionally assuming that *competition is stronger between stimuli comprising a group, than between stimuli associated with separate groups*. On this assumption, the irrelevant hemifield and its transient flashes compete for attention more vigorously in the grouped condition compared to the ungrouped condition. To reduce this distraction, stronger top-down bias is needed in favour of the relevant hemifield. This hemifield bias simultaneously explains the increase in response to transients (which may help to define the relevant stimulus area, see the segmentation argument above), and the decrease in activity for transients in the irrelevant hemifield (blue lines in Fig. 6). In the ungrouped case (green lines in Fig. 6), there would be less competition between hemifields, so that relevant target features can be selected without much interference from the opposite hemifield. Thus less top-down bias is needed to select the relevant hemifield. However in the ungrouped condition there would still be strong competition within the attended hemifield from its irrelevant transients, compared to transients in the opposite hemifield. Top-down bias might then be needed to suppress these non-target features specifically whenever they occur within the attended hemifield. This would explain why the response to the transient appeared lower in the attended hemifield than in the unattended hemifield, in the ungrouped condition. **It might be advantageous to apply this bias as early as possible in the processing stream, to achieve maximum leverage over the balance of competition between features, and hemifields, which could explain why the most robust effects of modulation were observed in primary visual cortex.**

Commented [EF13]: Revision 2

Increased within-group competition would also be consistent with two aspects of the block-related results (Fig. 4): firstly, the apparent enhancement of the difference between hemispheres contralateral and ipsilateral to the attentional cue, under grouping compared to no grouping, is consistent with stronger effects of bias in the former case; and secondly, the generally lower block-related activation across visual areas for grouped versus ungrouped stimuli, consistent with greater mutual inhibition between the hemifield representations (c.f. Kastner, De Weerd, Desimone & Ungerleider, 1998).

The proposed assumption of intensified competition within groups offers an alternative explanation of some more classical findings from object-based attention research, for example why 'flanker'

stimuli belonging to the same group as a target compete more vigorously for control over responding, than when displayed in a segmented context (Baylis & Driver, 1992; Zhao *et al.*, 2013). Enhanced competition for attention within groups could be of functional benefit by promoting rapid redeployment of attention to different locations within the same object. This would explain the lower response times observed in Posner-cuing paradigms to targets following invalid cueing to a location elsewhere within a closed contour (Egly *et al.*, 1994), which is more commonly explained in terms of facilitatory attentional spreading. However, in contrast to an attentional mechanism based only on automatic mutual facilitation of object parts and features, this enhanced inhibition might also allow the perceiver to ‘drill down’ to just one component of a group when this is uniquely relevant (e.g. the check-size targets in one hemifield), while suppressing other irrelevant component when they are fully irrelevant (e.g. the transient events, when they are truly of no use for segmenting the hemifields).

To conclude, the results of this study confirm that allocation of attentional resources, as indexed by changes in the BOLD signals in V1 evoked by task-irrelevant stimulus flashes, is strongly dependent on global cues for grouping. However a complex pattern of apparently suppressive as well as facilitatory effects reveals a wider gamut of behaviour than expected by current theories of object-based attention. In contrast to many previous findings, our results suggest that allocation of attention within the bounds of an object is neither always automatic nor always facilitatory, but can be task-dependent and suppressive for truly irrelevant stimuli. Such new patterns highlight the great flexibility, rather than the limitations, of our ability to selectively ignore irrelevant information, and to drill down to the level of detail at which specifically task-relevant information may be found.

Acknowledgements

This research was supported by project grants S13736 and S20366 from the Biotechnology and Biological Sciences Research Council (UK) and by the Wellcome Trust (GR). J.D. held a Royal Society Wolfson Research Merit Award). We thank John-Dylan Haynes for assistance with retinotopic mapping.

References

- Altmann, C. F., Bühlhoff, H. H., & Kourtzi, Z. (2003). Perceptual organization of local elements into global shapes in the human visual cortex. *Current Biology* 13, 342-349.
- Avrahami, J. (1999). Objects of attention, objects of perception. *Perception & Psychophysics* 61, 1604-1612.
- Baylis, G. C., & Driver, J. (1993). Visual attention and objects: evidence for hierarchical coding of location. *Journal of Experimental Psychology: Human Perception and Performance* 19, 451.
- Baylis, G. C., & Driver, J. (1992). Visual parsing and response competition: the effect of grouping factors. *Percept Psychophys* 51, 145-62.
- Chen, Z. (2012). Object-based attention: a tutorial review. *Attention, Perception and Psychophysics* 74, 784-802. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22673856>.
- Chen, Z., & Cave, K. R. (2013). Perceptual load vs. dilution: the roles of attentional focus, stimulus category, and target predictability. *Front Psychol* 4, 327.
- Chen, Z., & Cave, K. R. (2006). Reinstating object-based attention under positional certainty: the importance of subjective parsing. *Percept Psychophys* 68, 992-1003.
- Chun, M. M., & Marois, R. (2002). The dark side of visual attention. *Current Opinion in Neurobiology* 12, 184-189.
- Ciaramitaro, V. M., Mitchell, J. F., Stoner, G. R., Reynolds, J. H., & Boynton, G. M. (2011). Object-based attention to one of two superimposed surfaces alters responses in human early visual cortex. *J Neurophysiol* 105, 1258-65.
- Davis, G., Driver, J., Pavani, F., & Shepherd, A. (2000). Reappraising the apparent costs of attending to two separate visual objects. *Vision Research* 40, 1323 - 1332. Available at: <http://www.sciencedirect.com/science/article/pii/S0042698999001893>.
- Davis, G., & Holmes, A. (2005). Reversal of object-based benefits in visual attention. *Visual Cognition*, 12(5), 817-846.
- De Meyer, K., & Spratling, M. W. (2009). A model of non-linear interactions between cortical top-down and horizontal connections explains the attentional gating of collinear facilitation. *Vision Res* 49, 553-68.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience* 18, 193-222.
- Driver, J., & Baylis, G. C. (1989). Movement and visual attention: the spotlight metaphor breaks down. *Journal of Experimental Psychology: Human Perception and Performance* 15, 448-56. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/2527954>.
- Driver, J., Davis, G., Russell, C., Turatto, M., & Freeman, E. (2001). Segmentation, attention and phenomenal visual objects. *Cognition* 80, 61-95.
- Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General* 113, 501.
- Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review* 96, 433.
- Egley, R., Driver, J., & Rafal, R. D. (1994). Shifting visual attention between objects and locations: evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General* 123, 161.
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics* 16, 143-149.
- Ernst, Z. R., Boynton, G. M., & Jazayeri, M. (2013). The spread of attention across features of a surface. *Journal of Neurophysiology*. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23883860>.
- Festman, Y., & Braun, J. (2012). Feature-based attention spreads preferentially in an object-specific manner. *Vision Research* 54, 31-8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22227158>.
- Freeman, E., & Driver, J. (2008). Voluntary control of long-range motion integration via selective attention to context. *J Vis* 8, 18.1-22.

- Freeman, E., Sagi, D., & Driver, J. (2001). Lateral interactions between targets and flankers in low-level vision depend on attention to the flankers. *Nat Neurosci* 4, 1032-6.
- Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J-P., Frith, C. D., & Frackowiak, R. S. J. (1994). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping* 2, 189-210.
- Friston, K. J., Rotshtein, P., Geng, J. J., Sterzer, P., & Henson, R. N. (2006). A critique of functional localisers. *Neuroimage*, 30(4), 1077-1087.
- Fuentes, L. J., Humphreys, G. W., Agis, I. F., Carmona, E., & Catena, A. (1998). Object-based perceptual grouping affects negative priming. *Journal of Experimental Psychology: Human Perception and Performance* 24, 664.
- Harms, L., & Bundesen, C. (1983). Color segregation and selective attention in a nonsearch task. *Perception & Psychophysics* 33, 11-19.
- Hasnain, M. K., Fox, P. T., & Woldorff, M. G. (1998). Intersubject variability of functional areas in the human visual cortex. *Human Brain Mapping* 6, 301-315.
- He, X., Fan, S., Zhou, K., & Chen, L. (2004). Cue Validity and Object-Based Attention. *Journal of Cognitive Neuroscience* 16, 1085-1097.
- Hollingworth, A., Maxcey-Richard, A. M., & Vecera, S. P. (2012). The spatial distribution of attention within and across objects. *J Exp Psychol Hum Percept Perform* 38, 135-51.
- Jarmasz, J., Herdman, C. M., & Johannsdotir, K. R. (2005). Object-based attention and cognitive tunneling. *J Exp Psychol Appl* 11, 3-12.
- Kahneman, D., & Henik, A. (1981). Perceptual organization and attention. *Perceptual Organization* 1, 181-211.
- Kastner, S., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1998). Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. *Science* 282, 108-111.
- Khoe, W., Freeman, E., Woldorff, M. G., & Mangun, G. R. (2006). Interactions between attention and perceptual grouping in human visual cortex. *Brain Res* 1078, 101-11.
- Kim, M. -S., & Cave, K. R. (2001). Perceptual grouping via spatial selection in a focused-attention task. *Vision Research* 41, 611-624.
- Kim, Y. -J., & Verghese, P. (2012). The selectivity of task-dependent attention varies with surrounding context. *The Journal of Neuroscience* 32, 12180-91. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22933800>.
- Kramer, A. F., & Jacobson, A. (1991). Perceptual organization and focused attention: The role of objects and proximity in visual processing. *Perception & Psychophysics* 50, 267-284.
- Kramer, A. F., Weber, T. A., & Watson, S. E. (1997). Object-based attentional selection—Grouped arrays or spatially invariant representations?: Comment on Vecera and Farah (1994).
- Logan, G. D. (1996). The CODE theory of visual attention: an integration of space-based and object-based attention. *Psychological Review* 103, 603.
- Martinez, A., Ramanathan, D. S., Foxe, J. J., Javitt, D. C., & Hillyard, S. A. (2007). The Role of Spatial Attention in the Selection of Real and Illusory Objects. *The Journal of Neuroscience* 27, 7963-7973.
- Martinez-Trujillo, J. C., & Treue, S. (2004). Feature-based attention increases the selectivity of population responses in primate visual cortex. *Curr Biol* 14, 744-51.
- Martinez, A., Teder-Sälejärvi, W., Vazquez, M., Molholm, S., Foxe, J. J., Javitt, D. C., Di Russo, F., Worden, J. S., & Hillyard, S. A. (2006). Objects Are Highlighted by Spatial Attention. *Journal of Cognitive Neuroscience* 18, 298-310.
- Melcher, D., Papatomas, T. V., & Vidnyánszky, Z. (2005). Implicit attentional selection of bound visual features. *Neuron* 46, 723-9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15924859>.
- Melcher, D., & Vidnyánszky, Z. (2006). Subthreshold features of visual objects: Unseen but not unbound. *Vision Research*, 46(12), 1863-1867.
- Mitchell, J. F., Stoner, G. R., & Reynolds, J. H. (2004). Object-based attention determines

- dominance in binocular rivalry. *Nature* 429, 410-413.
- Murray, S. O., Kersten, D., Olshausen, B. A., Schrater, P., & Woods, D. L. (2002). Shape perception reduces activity in human primary visual cortex. *Proceedings of the National Academy of Sciences* 99, 15164-15169.
- Müller, N. G., & Kleinschmidt, A. (2003). Dynamic interaction of object- and space-based attention in retinotopic visual areas. *J Neurosci* 23, 9812-6.
- Neisser, U., & Becklen, R. (1975). Selective looking: Attending to visually specified events. *Cognitive Psychology* 7, 480-494.
- O'Craven, K. M., Downing, P. E., & Kanwisher, N. (1999). fMRI evidence for objects as the units of attentional selection. *Nature* 401, 584-587.
- Palmer, S., & Rock, I. (1994). Rethinking perceptual organization: The role of uniform connectedness. *Psychonomic Bulletin & Review* 1, 29-55.
- Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General* 109, 160.
- Richard, A. M., Lee, H., & Vecera, S. P. (2008). Attentional spreading in object-based attention. *J Exp Psychol Hum Percept Perform* 34, 842-53.
- Roelfsema, P. R. (2006). Cortical algorithms for perceptual grouping. *Annu Rev Neurosci* 29, 203-27.
- Saenz, M., Buracas, G. T., & Boynton, G. M. (2002). Global effects of feature-based attention in human visual cortex. *Nature Neuroscience* 5, 631-632.
- Schoenfeld, M. A., Tempelmann, C., Martinez, A., Hopf, J.-M., Sattler, C., Heinze, H., & Hillyard, S. A. (2003). Dynamics of feature binding during object-selective attention. *Proceedings of the National Academy of Sciences* 100, 11806-11811.
- Scholl, B. J., Pylyshyn, Z. W., & Feldman, J. (2001). What is a visual object? Evidence from target merging in multiple object tracking. *Cognition* 80, 159-77. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11245843>.
- Shomstein, S. (2012). Object-based attention: strategy versus automaticity. *Wiley Interdisciplinary Reviews: Cognitive Science* 3, 163-169.
- Shomstein, S., & Yantis, S. (2002). Object-based attention: Sensory modulation or priority setting? *Perception & Psychophysics* 64, 41-51.
- Teo, P. C., Sapiro, G., & Wandell, B. A. (1997). Creating connected representations of cortical gray matter for functional MRI visualization. *IEEE Transactions on Medical Imaging*, 16:852-863.
- Tipper, S. P., Driver, J., & Weaver, B. (1991). Short report: Object-centred inhibition of return of visual attention. *The Quarterly Journal of Experimental Psychology* 43, 289-298.
- Tootell, R. B., Reppas, J. B., Kwong, K. K., Malach, R., Born, R. T., Brady, T. J., Rosen, B. R., & Belliveau, J. W. (1995). Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *Journal of Neuroscience* 15, 3215.
- Valdes-Sosa, M., Bobes, M. A., Rodriguez, V., & Pinilla, T. (1998). Switching attention without shifting the spotlight object-based attentional modulation of brain potentials. *Journal of Cognitive Neuroscience* 10, 137-51. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9526088>.
- Vecera, S. P., & Behrmann, M. (2001). Attention and unit formation: A biased competition account of object-based attention. *Advances in Psychology* 130, 145-180.
- Vecera, S. P., & Farah, M. J. (1994). Does visual attention select objects or locations? *Journal of Experimental Psychology: General* 123, 146.
- Wandell, B. A., Chial, S., & Backus, B. T. (2000). Visualization and measurement of the cortical surface. *Journal of Cognitive Neuroscience*, 12:739-752.
- Watson, J. D. G., Myers, R., Frackowiak, R. S. J., Hajnal, J. V., Woods, R. P., Mazziotta, J. C., Shipp, S., & Zeki, S. (1993). Area V5 of the Human Brain: Evidence from a Combined Study Using Positron Emission Tomography and Magnetic Resonance Imaging. *Cerebral Cortex* 3, 79-94. Available at: <http://cercor.oxfordjournals.org/content/3/2/79.abstract>.
- Watson, S. E., & Kramer, A. F. (1999). Object-based visual selective attention and perceptual

- organization. *Perception and Psychophysics* 61, 31-49.
- Weber, T. A., Kramer, A. F., & Miller, G. A. (1997). Selective processing of superimposed objects: an electrophysiological analysis of object-based attentional selection. *Biological Psychology* 45, 159 - 182. Available at:
<http://www.sciencedirect.com/science/article/pii/S0301051196052271>.
- Wegener, D., Ehn, F., Aurich, M. K., Galashan, F., & Kreiter, A. K. (2008). Feature-based attention and the suppression of non-relevant object features. *Vision Research* 48, 2696-2707.
- Yeari, M., & Goldsmith, M. (2010). Is object-based attention mandatory? Strategic control over mode of attention. *J Exp Psychol Hum Percept Perform* 36, 565-79.
- Zhao, J., Kong, F., & Wang, Y. (2013). Attentional spreading in object-based attention: the roles of target-object integration and target presentation time. *Atten Percept Psychophys* 75, 876-87.

Table 1

MNI coordinates (mm) of regions identified in whole-brain contrasts, significant at $p < .05$ corrected.

Contrast	Area	x¹	y	z	T
<u>Block-related</u>					
L-R	Cuneus	24	-94	20	6.72
R-L	Cuneus	-26	-92	29	8.22
	Lingual gyrus	-26	-80	-10	5.19
<u>Event-related</u>					
L-R	Fusiform gyrus	32	-64	-18	7.12
R-L	Lingual gyrus	-26	-84	-12	4.31

1 MNI coordinates (mm)

Figure legends

Figure 1

Sample stimulus displays. Left panel: No transients, no targets; central red arrow is cueing leftwards attention. Right panel: Transient colour change on right hemifield, and a target check-size change is also shown on in both hemifields in the lower quadrant. Upper versus lower location was random and independent for each hemifield, and participants had to indicate the location of the target in the cued hemifield (here on the bottom left).

Figure 2

Distribution of horizontal gaze locations relative to central fixation (0), towards the cued hemifield (positive values) or away (negative), in degrees of visual angle. Height of bars indicates the number of scanning runs associated with each gaze value. Distributions are shown separately for grouped and ungrouped conditions (upper and lower graphs respectively). Red and blue colouring depicts unfiltered and filtered eye-data respectively (see Methods for details). Symbols mark mean horizontal gaze locations for individual subjects.

Figure 3

Whole-brain analyses. Left: Block-related contrast of left versus right-cued conditions, under similar stimulus conditions. These attention-driven areas were used to mask visual areas distinguished using retinotopy, to define our regions of interest. Right: Event-related contrast of left versus right colour-change transients. Results are superimposed on a standard template, with a threshold of $p < .001$ uncorrected.

Figure 4

Results from retinotopic analysis of Block-related BOLD, averaged across all visual areas. Normalised (see Methods) beta weights are plotted for hemispheres contralateral versus ipsilateral to the cued visual hemifield, for grouped (blue circles and solid lines) versus ungrouped stimuli (green squares and dashed lines). In all graphs, error-bars indicate one unit of within-subjects standard error. Inset shows predictions assuming facilitatory and automatic spreading of object-based attention.

Figure 5

Normalised block-related results for each area of visual cortex.

Figure 6

Results from retinotopic analysis of Block-related BOLD, showing the interaction between grouping and attention, which was significant in area V1 only. Normalised beta-weights (see Methods) are plotted as a function of whether the colour-change transients were on the attended or the unattended hemifield, for grouped (blue circles) versus ungrouped stimuli (green squares). Results are averaged across areas contralateral and ipsilateral to the location of the transient.

Figure 7

Event-related normalised results for each visual region of interest, and for contralateral versus ipsilateral transients.

Movie attachments

Movie1 and *Movie2* show excerpts of typical stimulus sequences from the Grouped and Ungrouped blocked conditions, respectively.

Supplementary Figures

Results for non-normalised analysis of Block-related (Supplementary Figure 1) and Event-related (Supplementary Figure 2) results.