Microsaccadic correlates of covert attention and crowding

Krishnamachari S. Prahalad

College of Optometry, University of Houston, Houston, TX, USA ✉

Daniel R. Coates

College of Optometry, University of Houston, Houston, TX, USA ✉

Spatial crowding occurs when an object is cluttered among other objects in space and is a ubiquitous factor affecting object recognition in the peripheral visual field. Crowding is typically tested by presenting crowded stimuli at an eccentric location while having observers fixate at a point in space. However, even during fixation, our eyes are not perfectly steady but instead make small-scale eye movements (microsaccades) that have recently been suggested to be affected by shifts in attentional allocation. In the current study, we monitored microsaccadic behavior (a possible attentional correlate) to understand naturally occurring shifts in attention that occur following the presentation of a crowded stimulus. A tracking scanning laser ophthalmoscope (TSLO) was used to image the right eye of each observer during a psychophysical task. The stimuli consisted of Sloan numbers (0–9) presented briefly, either unflanked or surrounded by Sloan numbers at one of four nominal spacings. The extent of crowding was found to decrease by 26% on trials with the presence of incongruent microsaccades (proposed to suggest attentional capture). These findings complement the existing body of literature on the beneficial impact of explicit shifts of spatial attention to the location of a crowded stimulus.

Introduction

When stimuli are presented far from fixation or central vision, it is more difficult to recognize the stimulus, indicating that the visual system is non-homogenous across the visual field (Jacobs, 1979; Poletti, Listorti, & Rucci, 2013; Weymouth, Hines, Acres, Raaf, & Wheeler, 1928). This becomes all the more demanding when the same stimulus is cluttered among other objects in space (flankers), due to a phenomenon termed as "visual crowding" (Bouma, 1970; Stuart & Burian, 1962; Toet & Levi, 1992), considered to be a major bottleneck affecting object recognition in the peripheral visual field (for reviews, see Levi, 2008; Pelli & Tillman, 2008). The effect of

crowding depends on the distance between the target and surrounding flankers. The distance between target and flankers where crowding is eliminated is called the "critical spacing" and determines the spatial extent of visual crowding. The spatial extent of visual crowding has been shown to be dependent on many factors, such as the eccentricity of the target from fixation (Bouma, 1970; Toet & Levi, 1992), properties of the target and flankers such as their contrast (Coates, Chin, & Chung, 2013) or configural relationships (Melnik, Coates, & Sayim, 2020), stimulus duration (Coates, Ludowici, & Chung, 2021), and, most relevant to the current study, the deployment of spatial attention.

Spatial attention helps an observer focus on specific objects or items that are present in the visual scene by increasing the priority for this object in the visual processing pathway (Yeshurun & Carrasco, 1999). One way this can be achieved is with spatial cueing, which indicates the future location of a target, leading to an enhancement in behavioral performance (Carrasco, Penpeci-Talgar, & Eckstein, 2000; Müller & Rabbitt, 1989; Nakayama & Mackeben, 1989; Posner, 1980; Yeshurun & Carrasco, 1999). More recently, studies have looked at whether explicitly drawing one's attention to a crowded target can minimize the detrimental effects of visual crowding (Chakravarthi & Cavanagh, 2009; Felisberti, Solomon, & Morgan, 2005; Freeman & Pelli, 2007; Huckauf & Heller, 2002; Morgan, Ward, & Castet, 1998; Scolari, Kohnen, Barton, & Awh, 2007; Strasburger, 2005; Yeshurun & Rashal, 2010). The results from these studies on the benefit of pre-cueing a crowded target have been mixed, with several finding a slight reduction in critical spacing (Chakravarthi & Cavanagh, 2009; Freeman & Pelli, 2007; Yeshurun & Rashal, 2010), whereas others have observed a reduction in critical spacing only at near eccentricities and not at farther eccentricities (Felisberti et al., 2005; Strasburger, 2005), whereas others found no reduction in critical spacing, but a general improvement in performance (Huckauf & Heller, 2002; Scolari et al., 2007). Finally, several older studies found no improvement in performance nor any reduction in

crowding from pre-cueing the target location (Nazir, 1992; Wilkinson, Wilson, & Ellemberg, 1997). Across these different studies, it is generally assumed that subjects shift their attention when pre-cued prior to target presentation. However, it remains unclear what proportion of trials the subjects did successfully shift their attention. Hence, a more objective means to understand when these attentional shifts occur is desirable. In the current study, we were particularly interested in understanding how naturally occurring shifts of attention affect behavioral performance in a crowded recognition task.

Traditionally, crowding experiments are performed by having subjects fixate on the center of the display while presenting stimuli at a different location that is far from the fixation point. However, even when subjects try to fixate upon a point in space, they make miniature eye movements unwittingly (Ditchburn, 1955; Engbert & Kliegl, 2003; Martinez-Conde, Macknik, & Hubel, 2004; Rucci & Poletti, 2015; Steinman, Haddad, Skavenski, & Wyman, 1973). These tiny eye movements or fixational eye movements (FEMs) are generally subdivided into tremors, drifts, and microsaccades (Ditchburn, Fender, & Mayne, 1959), with the latter being essential in exploring the stimulus while fixating (Ko, Poletti, & Rucci, 2010; Poletti et al., 2013). Additionally, over the last 2 decades, numerous studies have indicated the potential use of microsaccadic behavior to understand modulations in covert attention (Baumeler, Schönhammer, & Born, 2020; Engbert & Kliegl, 2003; Hafed, 2013; Hafed & Clark, 2002; Horowitz, Fencsik, Fine, Yurgenson, & Wolfe, 2007; Laubrock, Engbert, & Kliegl, 2005; Laubrock, Kliegl, Rolfs, & Engbert, 2010; Pastukhov & Braun, 2010). Covert attention comes into play when the orientation of gaze differs from the orientation of attention (Posner, 1980). The studies mentioned earlier have used paradigms comparable to the classic Posner cueing paradigm, where observers are required to respond to a peripheral target that is preceded by a central or peripheral cue, to understand the relationship between microsaccades and shifts of attention. Across these different studies, it has typically been observed that microsaccadic properties tend to modulate following stimulus onset. More specifically, microsaccadic rates initially show a rapid decrease (inhibition) following stimulus onset, followed by a window of increase (enhancement) where microsaccadic rates reach levels higher than that of baseline (Engbert & Kliegl, 2003; Hafed & Clark, 2002; Laubrock, et al., 2005; Laubrock, et al., 2010). Likewise, microsaccadic directions have generally shown a bias toward the stimulus location right around the time they are least likely to occur – the inhibition period (Hafed & Ignashchenkova, 2013; Pastukhov & Braun, 2010), indicating a dissociation between microsaccadic rate and direction. The strongest links between spatial attention and microsaccadic

direction have been shown to occur during the enhancement period of the microsaccadic rates (Laubrock et al., 2010), and involve microsaccades both toward and away from the target (see Engbert, 2006 for review). To further elucidate the role of spatial attention on visual crowding, we analyzed microsaccades (a possible attentional correlate) that occurred around the time a crowded stimulus was presented.

Thus, although explicit shifts in spatial attention have been shown to reduce the spatial extent of visual crowding (Chakravarthi & Cavanagh, 2009; Freeman & Pelli, 2007; Yeshurun & Rashal, 2010), it is still unclear whether an attentional correlate, such as microsaccade activity, can be used to predict the trial-by-trial performance in a crowded visual recognition task. In the current study, we had subjects identify a crowded target while monitoring fixational eye movements using a state-of-the-art retinal image tracker with unprecedented accuracy. To anticipate our results, in agreement with the literature using explicit cues, we found an enhancement in behavioral performance on trials with increased attentional capture suggested by microsaccadic characteristics, revealing a reduction in crowding even in the presence of implicit shifts of spatial attention.

Materials and methods

Participants

We recruited five normally sighted individuals that were free of any ocular pathology from the University of Houston College of Optometry (24 to 47 years, 1 female participant), to participate in the experiments reported here.
To ensure that the eye tracking was robust, subjects who required refractive corrections wore contact lenses while running the experiment. The experiment was performed monocularly; the non-tested eye was occluded with an eye patch. Visual acuity was confirmed to be 20/20 or better in the eye that was tested for all subjects, as assessed with an Early Treatment Diabetic Retinopathy (ETDRS) chart. The experimental protocol was approved by the internal review board at the University of Houston and written informed consent was obtained from the subjects before the start of the experiment.

Apparatus

A custom-built tracking scanning laser ophthalmoscope (TSLO; Sheehy, Yang, Arathorn, Tiruveedhula, de Boer, & Roorda, 2012), which offers high fidelity imaging of the retina, was used to monitor eye movements. The horizontal scanner runs at 16 KHz,
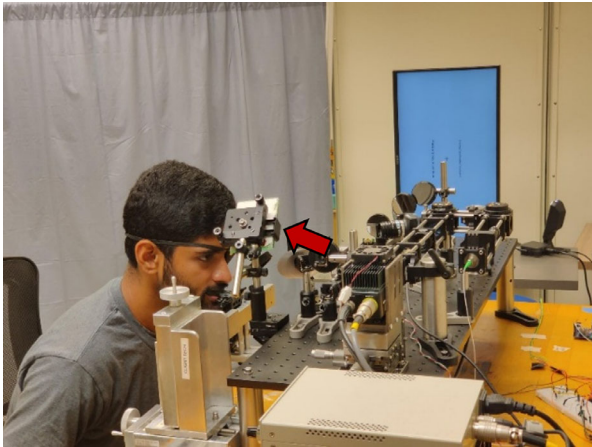
Figure 1. The experimental setup with the observer aligned in front of the system using a custom bite-bar. The observer gets to see both the raster from the Tracking Scanning Laser Ophthalmoscope and the external display simultaneously with the help of the periscope system consisting of a cold and regular flat mirror (indicated by the red arrow).

whereas the vertical scanner runs at 1/512 of that of the horizontal scanner, approximately 30 Hz, corresponding to the rate at which retinal videos were captured. An 840 nm super luminescent diode (SLD) with a 50 nm bandwidth, which was captured by a photomultiplier tube, was used to obtain retinal images. Stimuli were presented on an ASUS ROG StrixXg-8Q LCD monitor with a diagonal screen size of 62.5 cm and a refresh rate of 240 Hz. The resolution of the monitor was set to 1920 × 1080 pixels and was placed 180 cm from the subject, where each pixel on the display subtended 0.54 arc minutes of visual angle. The external display was made visible to the same eye that was being imaged by the TSLO using a cold mirror and a regular mirror that were set to be parallel to each other (i.e. a periscope; shown in Figure 1, indicated by the red arrow). The cold mirror transmits >95% of the infrared imaging raster while reflecting >90% of the visible light. This enabled observers to view both the imaging raster and the stimulus display when properly aligned. For our experiments, the field of view (FOV) of the system was set to 3 × 3 degrees, thus each pixel subtends 0.35 arc minutes of visual angle at its resolution of 512 × 512 pixels^2. A smaller FOV was chosen to increase the precision of eye tracking modality, however, this meant that relatively larger saccades would not be reliably detected by the system. To ensure that the raster did not interfere with the psychophysical task, we positioned it in the superior visual field, approximately 5 degrees above the fixation point. A custom bite-bar prepared for each subject using their dental impression that was then used to position the observer in front of the TSLO instrument, minimizing the overall head movement during the experiment.

To synchronize the timing between the imaging and stimulus channels, a photodiode was positioned in front of the stimulus display. Every time an event occurred on the stimulus computer a marker was drawn on the display which in turn triggered the photodiode. This was then detected by a 600 MHz microprocessor (Teensy 4.0; PJRC, Inc.), which momentarily attenuated the signal to the acousto-optic modulator (AOM). This reduction of the laser appeared as a black strip (due to the absence of the imaging beam) on the TSLO movie. The black line extended 512 pixels horizontally (the entire width of TSLO movie) and eight pixels vertically and served as an event marker which was later extracted from the retinal movies and used for precise temporal alignment of the retinal videos (and resultant eye traces) with the stimulus presentation times.

Retinal video analysis

The raw retinal videos were used to extract high temporal frequency eye position signals using cross-correlation techniques described elsewhere (Zhang, Gofas-Salas, Leonard, Rui, Snyder, Reecher, Mecê, & Rossi, 2021). The fast scanner of the TSLO system scans the retina one scan line at a time and each scanline is generally unaffected by eye motion, but it is possible for the eye to have moved between scan lines. To obtain retinal motion from SLO movies, a reference frame is initially formed using a subset of frames from the retinal video that were good in terms of image quality. Following this, each frame of the video was broken up into smaller horizontal strips, the shift of retinal features between the frame of interest and the reference frame was compared using a cross correlation procedure. We used a single global reference frame for each subject, this ensured that the absolute eye positions obtained would be comparable between trials. For this experiment, each frame of the video was divided into 18 strips, which was then used in the strip-based method, and we were able to reliably yield eye position signals at a temporal frequency of 480 Hz.

The eye position signal was then converted from pixels to degrees of visual angle based on the field size that was used to capture the retinal videos. Because the TSLO is primarily sensitive to small scale eye movements, any larger eye movements may be problematic for the strip-based method described above. One of the authors manually examined each video to ensure optimal tracking and trials with imperfect tracking either due to poor image quality or presence of larger saccades were repeated. The eye position signal extracted from the retinal videos was categorized into saccades and drifts using an adaptation of the median-based velocity threshold algorithm (Engbert & Kliegl, 2003) for microsaccade detection from the post processing modules of the Retinal Video Analysis Suite

(ReVAS; Ağaoğlu, Sit, Wan, & Chung, 2018) where the base algorithm is applied in a moving window to fine tune the boundaries of each event. Eye positions were classified as saccadic eye movements and drifts using the default parameters from ReVAS, when: (1) eye velocity exceeded 30 deg/second, (2) saccade duration was between 8 and 100 ms, and (3) the interval between one event and the next was greater than 20 ms. Following this, to ensure that the saccade events were not simply artifacts, we excluded events based on the following criterion: (1) saccade durations shorter than 15 ms or longer than 80 ms, (2) interval between one event and the next was greater than 200 ms, and (3) outlier events that did not follow the linear relationship between peak velocity and saccade amplitude. For the latter, we identified two separate trend lines upon conducting a Hough transform (Duda & Hart, 1972) on the relationship between peak velocity and saccade amplitude, we added a halfway line between the existing trend lines and used data that fell above the halfway line.

Experiment: Testing visual crowding using method of constant stimuli

The target consisted of custom-rendered numbers (0-9) designed with Sloan proportions that were either presented in isolation (uncrowded condition) or along with three other flankers (above, below, and on the outer side of the target). The inner flanker was removed as it would otherwise overlap with the fixation target at closer target eccentricities or at larger target-flanker spacings, following other crowding studies that have used a similar stimulus layout (Soo, Chakravarthi, & Andersen, 2018; Yeshurun, Rashal, & Tkacz-Domb, 2015; Yeshurun & Rashal, 2010). It has been shown that the outer flanker elicits a stronger effect when compared to the inner flanker (Petrov, Popple, & McKee, 2007), hence, we decided to remove the inner flanker for the current study. When flankers were present, they appeared at one of five different distances, specified as center-to-center distance from the target letter as multiples of letter height (nominal spacings): 1.6×, 2×, 3×, and 5×. The stimulus (target and flankers) was presented at five different eccentricities from fixation: 1°, 2°, 3°, 4°, and 5°. The post-mask consisted of a square patch of size 2.25° × 2.25°, where Sloan numbers (like those used in the main experiment) were randomly chosen and rotated before positioning them on a grid to fill up the square patch. The number of Sloan numbers used to fill the mask (density) was adjusted until they were found to just about touch each other. A set of five such masks were pre-rendered at the start of the experiment, on each trial, the set of masks that would be presented at locations corresponding to the target and flankers were randomly chosen.

The sequence of events that occur during the experiment are depicted in Figure 2. Each trial started with a fixation cross that was presented for 750 to 1250
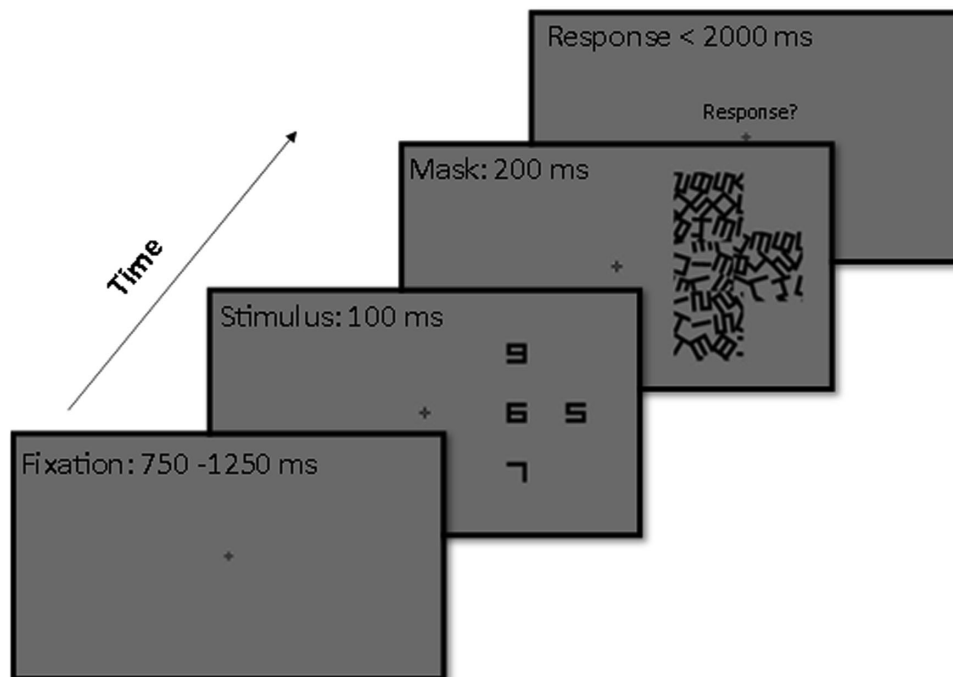

Figure 2. Trial sequence for the experiment. The subjects would report the central target of the stimulus following post-mask offset, by making a keypress.

milliseconds (ms), followed by the target and flankers (100% contrast) presented for a period of 100 ms. On a given trial, the stimulus (target and flankers) was randomly presented either to the right or left from the central fixation cross. The post-mask, centered on the target letter, was presented for 200 ms following stimulus offset. The subjects then reported the number that corresponded to the central target by pressing the corresponding number on a keyboard within a 2 second period following the post-mask offset. Auditory feedback was given at the end of each trial. Prior to the main experiment, the unflanked threshold letter size for each eccentricity was determined using the QUEST algorithm (Watson & Pelli, 1983) with 32 trials targeting a correct performance level of 75%. The letter size (target and flankers) used in the main experiment was set to 1.5× the threshold letter size for a given eccentricity obtained from the QUEST method (Toet & Levi, 1992). This was to ensure that performance would not be limited to the size of the target.

Each participant completed 20 blocks of trials, where each block consisted of trials at five different eccentricities (1°, 2°, 3°, 4°, and 5°). At a given eccentricity, a total of 50 trials (set) were performed: four nominal spacings (1.6×, 2×, 3×, and 5×) and unflanked times 10 repeats and retinal videos of 120 to 130 seconds were recorded simultaneously. Both the order of stimulus eccentricities and flanker spacings were randomized. Participants were given breaks between each set and at the end of each block.

Data analysis

Given the small field used to capture retinal images, detection of only smaller saccades was possible with the offline analysis method. Hence, we used all detected microsaccades for our analysis. Microsaccadic events detected from the eye position signal were independently grouped based on (1) their time course on each trial relative to stimulus onset and (2) based on their direction relative to stimulus location. For the latter, we first grouped microsaccades with an angle <45 degrees and >315 degrees as rightward microsaccades and ones with an angle >135 degrees and <225 degrees as leftward microsaccades. Finally, we categorized them as either congruent to the stimulus, if they were directed toward the stimulus location, or incongruent, if they were directed opposite to the stimulus location.

The time course of performance was computed by binning trials starting from –800 ms before stimulus onset to 800 ms following stimulus onset. The width of each bin or moving window was set to be 200 ms and the step size of the moving window was set to be 2 ms. Likewise, the microsaccadic metrics, such as the proportion of congruent microsaccades and saccade

amplitude, were computed in a similar fashion with the same bin parameters. The binning method was performed individually for each subject.

Statistical analysis

We conducted permutation tests to determine if the difference between the binned microsaccade metric was significantly different across the different conditions tested, where we used an alpha level of $p = 0.001$ (two-sided) to test for significance between conditions. The two conditions being tested were combined and shuffled, before drawing 10,000 samples without replacement and recalculating the statistical test. To compare behavioral performance, we used the difference of means between the groups. Prior to calculating the difference in means, we first converted the mean proportion correct from each group to their respective d' measure, considering the 10-Alternate Forced Choice (AFC) task used in the study, given by the following equation:

$$d' (sensitivity) = [0.86 - 0.085 \log (M - 1)]$$
$$\left[\log \left((M - 1) \frac{p}{1 - p} \right) \right] \quad (1)$$

where M is the number of alternatives in the task used and *p* is the proportion correct from the behavioral data (Smith, 1982).

The statistical measures between the conditions being compared and the shuffled data are compared to determine time points where significance levels were reached (Blair & Karniski, 1993). This procedure accounts for multiple comparisons and minimizes false positive errors (Groppe, Urbach, & Kutas, 2011). In addition, we took a more conservative approach when choosing an alpha level to minimize false positive errors.

To determine the critical spacing, we initially modeled the performance at each target flanker spacing using a cumulative distribution function (CDF) for the Weibull distribution, given by the following equation:

$$P_{Weibull} = 1 - \exp \left[-\left(\frac{x}{\lambda} \right)^{k} \right] \quad (2)$$

where x is the flanker spacing, λ is the scale parameter, and k is the shape parameter. This was used to calculate the level of model fit using the Weibull CDF.

$$Y_{val} = \gamma + (1 - \gamma - \lambda) \, P_{Weibull} \quad (3)$$

where γ is the guess rate, which was set to be 0.1 here (for a 10-AFC task), $P_{Weibull}$ is derived from Equation 2, and λ is the lapse rate. We repeated this for each of the conditions tested individually and the fitted

parameters was used to calculate the critical spacing, c. The critical spacing is defined as the distance between the target and flankers at which performance is similar to that without flankers. This was calculated using the following equation:

$$c = interpolate\ (0.75,\ Y,\ X) \quad (4)$$

where X is the x intercept and Y is the y intercept. The critical spacing was given by the point at which performance reached 75% of the asymptotic level.

Results

We analyzed a total of 500 retinal videos across subjects and extracted 35,806 microsaccades (2.2% of events were discarded, see the Methods section) between the time interval of 800 ms before stimulus onset and 800 ms after stimulus onset. It was only this subset of microsaccades that were used for subsequent analysis. The number of microsaccades detected were similar across eccentricities. To validate the microsaccade detection method, we looked at the relationship between the amplitudes and peak velocities of microsaccades (shown in Figure 3a), where we found a significant correlation between the two metrics (Pearson Correlation: $r = 0.96$, $p < 0.001$). This relationship is generally observed in the case of larger saccades (Zuber, Stark, & Cook, 1965) and we see a similar trend in the case of microsaccades. Microsaccades with smallest amplitudes but large peak velocities are possibly due to intrinsic noise of

the retinal image based tracking modality. Similarly, Figure 3b shows the microsaccade amplitude distribution, which has a median amplitude of 15 arc minutes.

Temporal modulation of microsaccadic activity

First, we wanted to look at whether microsaccades modulate when individuals were asked to perform a crowded recognition task. To this end, we looked at the relationship between the proportion of trials that had microsaccades in them at different time points relative to stimulus onset (shown in Figure 4). Like in previous studies (Engbert & Kliegl, 2003; Hafed & Ignashchenkova, 2013; Pastukhov & Braun, 2010), there was a sharp decrease or inhibition in microsaccadic frequency immediately after stimulus onset, with it reaching its lowest levels between approximately 125 and 150 ms following stimulus onset. Following this, there was an increase – rebound in microsaccadic frequency – with it surpassing baseline or pre-stimulus frequency levels. There was a second dip in microsaccadic frequency between approximately 350 and 400 ms following stimulus onset, which could be associated with the offset of the post-mask in our experimental sequence – previous researchers have found that any change in the display may trigger a microsaccadic rate change with inhibition followed by enhancement (Laubrock et al., 2005). A similar trend was observed across all spacings, and target eccentricities tested, with no major differences observed between them.
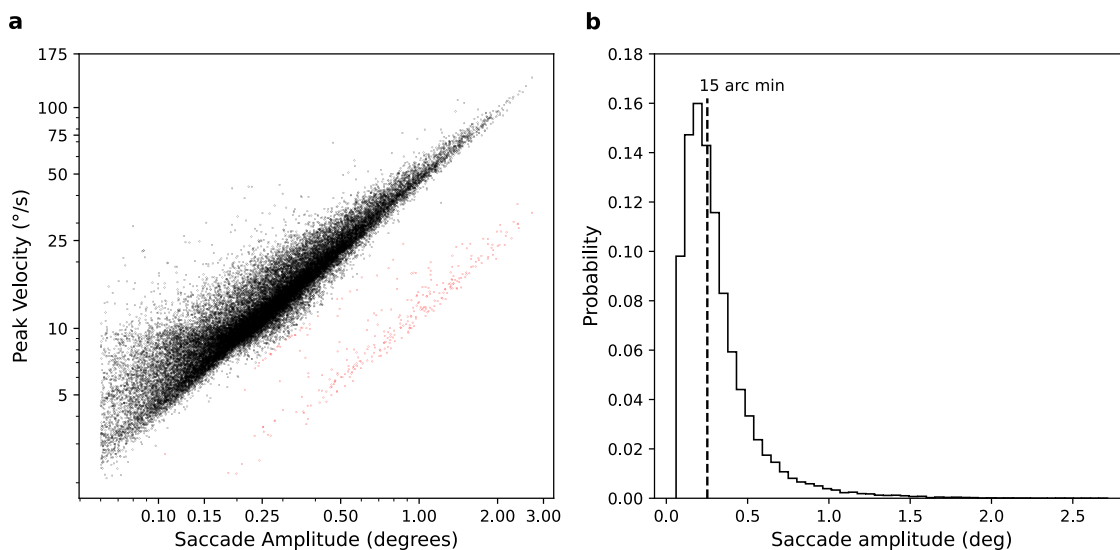

Figure 3. Saccade characteristics for events that occurred between 800 ms before and after stimulus onset. (**a**) Shows the relationship between the peak saccadic velocity and saccade amplitude. Data points rejected by the Hough Transform method are marked in red (2.2% of events were discarded). (**b**) Shows the saccade amplitude distribution.

Behavioral performance as a function of microsaccadic congruency

The primary goal of the study was to determine how microsaccadic behavior can predict the performance in a crowded recognition task. In Figure 5, we show behavioral performance when microsaccades occur at different time points relative to stimulus onset. Trials that did have microsaccades were subdivided based on whether they were directed towards (congruent) or away from (incongruent) the stimulus location (described in detail under data analysis section). Additionally, we looked at the performance on trials where microsaccades did not occur during the time window of interest around stimulus onset (between 800 ms before and after stimulus onset). To determine if there was an effect of microsaccadic congruency on behavioral performance, we conducted a two-tailed permutation test using the difference in means at each time point starting from 800 ms before stimulus onset to 800 ms after stimulus onset. Surprisingly, we observed that performance was better on trials with microsaccades that were incongruent to the stimulus location; this was seen at time intervals between 150 and 496 ms following stimulus onset ($p < 0.001$). This benefit in performance was observed on trials that had microsaccades that were directed away from the target location (incongruent) in the period of increased microsaccadic frequency or rebound interval occurring after stimulus onset, shown in Figure 4.

We then looked more closely at the relationship between behavioral performance and the different target-flanker spacings tested. More specifically, we looked at the time window where microsaccadic congruency was shown to affect performance or were found to be different (i.e. 150 to 496 ms after stimulus

onset, the shaded region on Figure 5). Figure 6 shows the behavioral performance across the different target-flanker spacings; each trace represents performance in each of the microsaccade conditions, each panel shows the data from the different target eccentricities that were tested. Additionally, we also added the performance from trials without microsaccades in the same time window between 150 and 496 ms following stimulus onset. As seen before, performance was best on trials where microsaccades were directed away from the stimulus location between 150 and 496 ms following stimulus onset. This difference was more pronounced for spacings 2×, 3×, and 5×, especially for eccentricities less than 5 degrees.

To account for the heterogeneity of the data, a bootstrapping technique was utilized (see reviews in Efron, 1982; Foster & Bischof, 1997) to determine confidence intervals. We resampled the performance measures with replacement, with the number of samples drawn each time matched to the size of the data subset for each condition. This was repeated 1000 times for each target eccentricity across all subjects and was computed independently for each microsaccade condition. This method helps to account for the small number of trials under each condition (Hinkley, 1988). To quantify this difference in the crowding effect, a CDF for the Weibull function (Equation 3) was fit to the resampled performance data across the different microsaccade conditions independently. The function was corrected for a guess rate of 10%, for a number recognition task with numbers ranging from 0 to 9. The resulting function fits and the corresponding bootstrapped 95% confidence intervals are shown in Figure 6 for trials with congruent microsaccades (green), trials with incongruent microsaccades (red), and trials with no microsaccades (black). The fitted functions were then used to compute the critical spacings for each of the microsaccade conditions independently. Previous studies have used similar approaches (Coates et al., 2021), by identifying the target-flanker spacing where performance dropped 1/e of the asymptotic level (Tripathy & Cavanagh, 2002) or by 10% of the amplitude of the model fit (Tripathy, Cavanagh, & Bedell, 2014). For the current experiment, the critical spacing was given by the spacing that yielded a 25% reduction in performance of the asymptotic level that was reached (Equation 4). The critical spacing varied between 2.5× and 3.86× for trials with incongruent microsaccades, 3.88× and 3.92× for trials with congruent microsaccades and 3.64× and 4.22× for trials without any microsaccades in the time interval between 150 and 496 ms following stimulus onset. Across eccentricities 1 degree through 4 degrees, we found a 26% decrease in critical spacing on average between trials that had incongruent microsaccades when compared to ones that had congruent microsaccades, whereas they were found
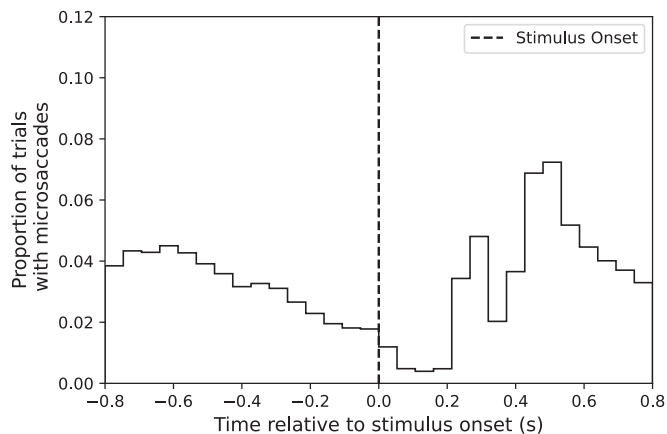

Figure 4. Proportion of trials with microsaccades is shown as a function of time from stimulus onset (indicated by the dashed line). Negative values on the x-axis represents events occurring prior to stimulus onset, whereas positive values indicate ones that occur after stimulus onset.
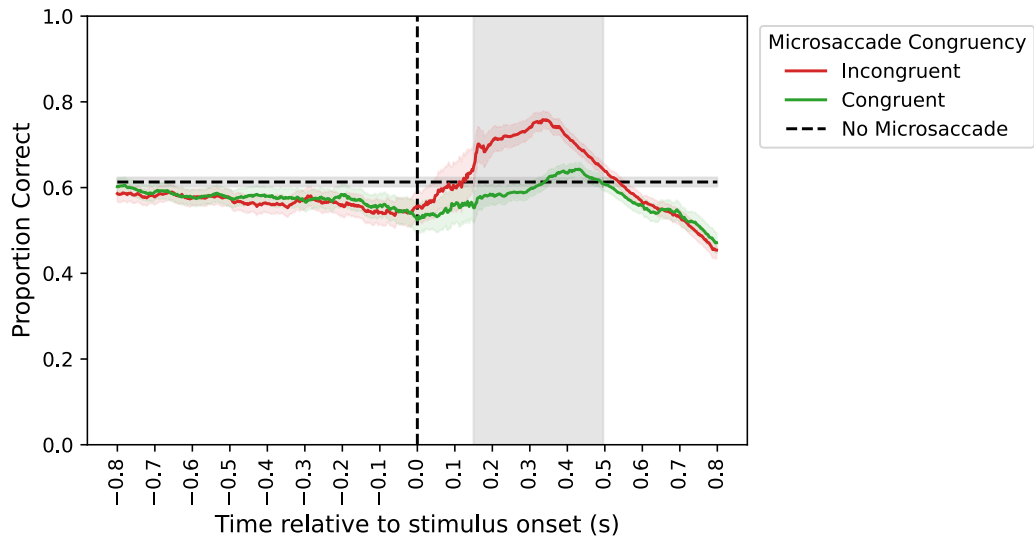
Figure 5. The mean performance is shown for trials that had microsaccadic events occurring at different time points relative to stimulus onset (indicated by the dashed line). The green trace represents the performance on trials with microsaccades occurring towards the stimulus location (congruent), whereas the red trace represents the performance on trials with microsaccades occurring away from the stimulus location (incongruent). The black line represents performance on trials without any microsaccades between 800 ms before and 800 ms after stimulus onset. The shaded region around each trace represents the 95% confidence intervals at each time point. The gray shaded regions represent time points where the traces were significantly different from each other.
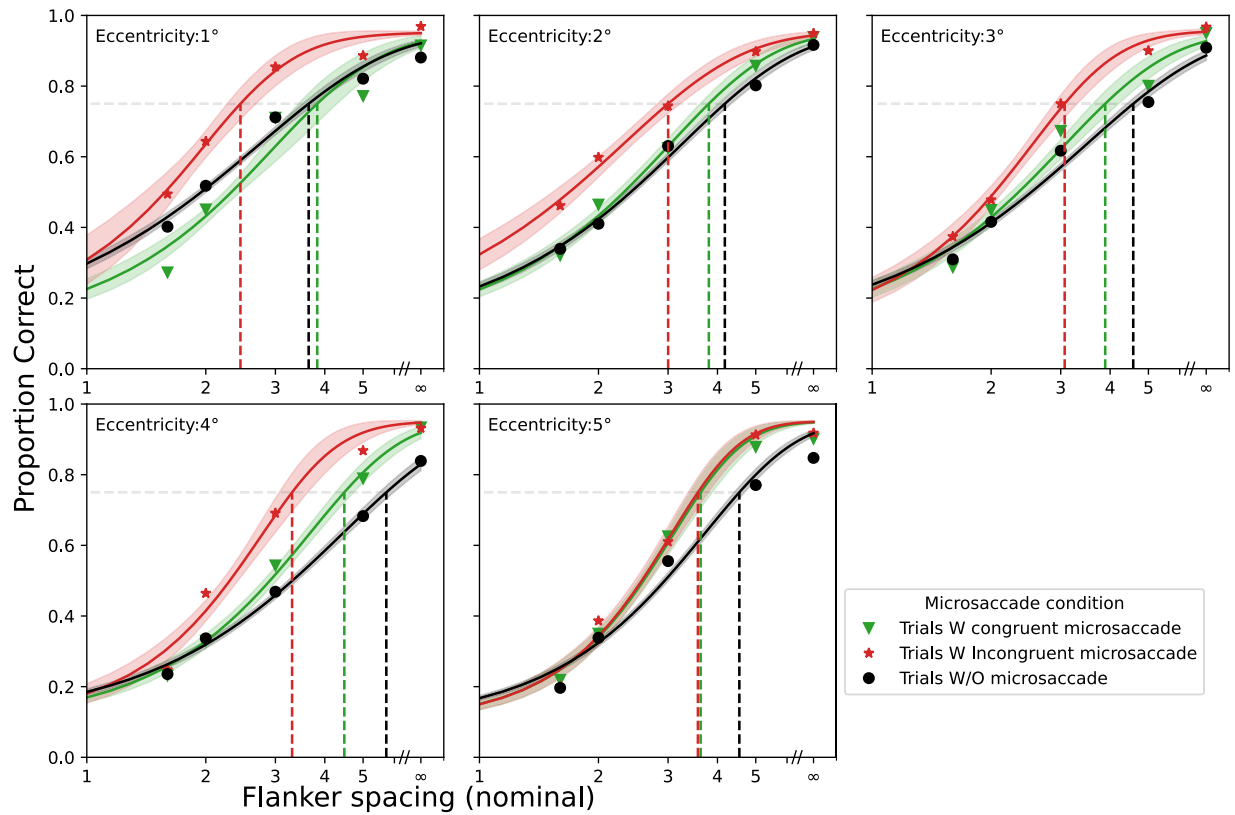

Figure 6. Behavioral performance vs. nominal flanker spacing, grouped by microsaccadic activity between 150 and 496 ms after stimulus onset. Each set of points represents the performance across each of the conditions tested (indicated by the legend), with each panel representing each target eccentricity tested. The curve was fitted for each condition using a cumulative distribution function for the Weibull distribution. The shaded region for each curve represents the bootstrapped 95% confidence interval.
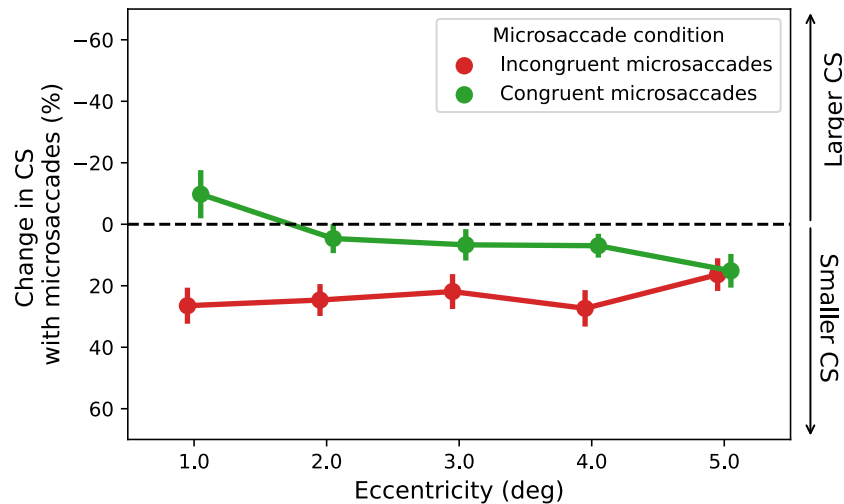
Figure 7. The percent change in critical spacing between the trials without microsaccades and the ones with microsaccades are plotted as a function of eccentricity (degrees). Trials were grouped based on microsaccadic activity between 150 and 496 ms following stimulus onset. Each individual trace represents the percent change in critical spacing for each individual microsaccade condition (indicated by the legend), with the green trace representing the difference between trials with no microsaccades and congruent microsaccades, while the red trace represents the difference between trials with no microsaccades and incongruent microsaccades. The error bars indicate the 95% confidence interval across subjects. The dashed horizontal line represents the condition where there is no change in critical spacing between the microsaccade and no microsaccade groups.

to be similar at 5 degrees. Trials with microsaccadic activity (congruent or incongruent) in the time interval between 150 and 496 ms were generally found to be better than trials that did not have microsaccadic activity, this was found to be consistent across all target eccentricities. Individual results from each subject at each target eccentricity were found to be consistent with the general trend, as shown in Figure A1.

Next, we investigated the benefit in performance for trials with incongruent microsaccades across the different eccentricities tested. We repeated a similar bootstrapping method as before to determine confidence intervals. We then fit the Weibull function (Equation 3) to the bootstrapped dataset and estimated the critical spacing using the same technique as before. Here, this was repeated 1000 times for each combination of subject and target eccentricity and was computed independently for each microsaccade condition. Figure 7 shows the percent change in critical spacings between the trials without microsaccades and each of the microsaccadic conditions (congruent and incongruent), indicated by each individual trace. We conducted a 2-way repeated measures ANOVA with the critical spacing as the dependent variable and two measures (eccentricity and microsaccade condition) as the within subject factors. We found a main effect of eccentricity ($F_{(df\ =\ 4)} = 90.3522$, $p < 0.001$) and microsaccade condition ($F_{(df\ =\ 2)} = 29.0460$, $p < 0.001$) on the critical spacing, with the interaction between the terms just barely reaching significance ($F_{(df\ =\ 8)} = 2.3719$, $p = 0.0395$). A Tukey HSD post hoc test revealed significant pairwise differences between congruent

and incongruent microsaccadic conditions, between congruent and no microsaccade conditions, and between incongruent and no microsaccade conditions (all three pairs: adjusted $p < 0.05$). Overall, we observed that the critical spacings were smaller for trials with incongruent microsaccades occurring between 150 and 496 ms following stimulus onset when compared with trials that either had congruent microsaccades or no microsaccadic activity in the same time interval.

Effect of stimulus onset on microsaccadic behavior

Previous studies have observed an increase in the proportion of microsaccades directed toward the stimulus, especially right after stimulus onset (Engbert & Kliegl, 2003; Hafed & Ignashchenkova, 2013; Pastukhov & Braun, 2010). Here, we determined whether these stimulus-driven microsaccadic directional biases could contribute to the effect observed. Figure 8 shows the relationship between the proportion of microsaccades that were directed toward the stimulus location (congruent) as a function of time from stimulus onset (indicated by the dashed vertical line). The horizontal line indicates the condition where both congruent and incongruent microsaccades are equally likely to occur (absence of directional bias). Additionally, the temporal modulation of microsaccadic frequency (see Figure 4) is shown by the histogram in the background. We observed an increase in congruency of microsaccadic direction right after
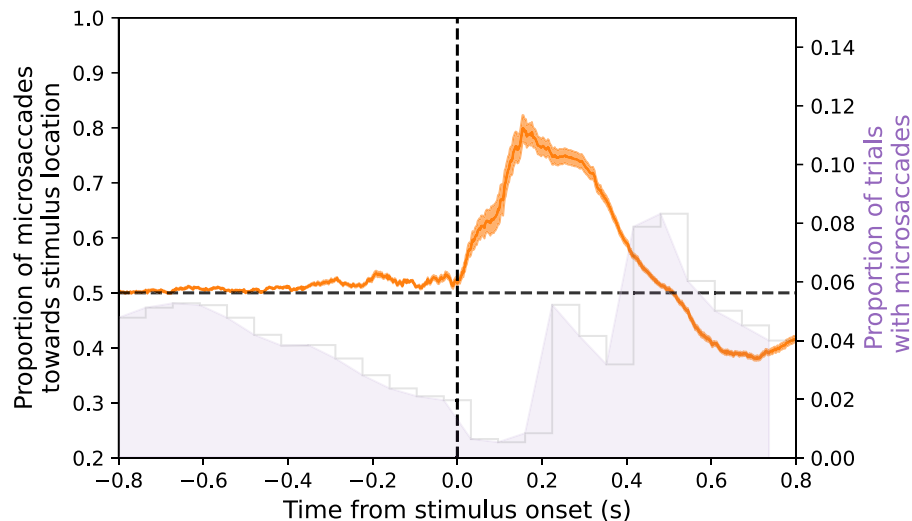
Figure 8. Proportion of microsaccades directed towards the stimulus location versus time from stimulus onset. The histogram representing the temporal modulation of microsaccades is shown in the back. The shaded region around each trace represents the 95% confidence intervals at each time point. The vertical dashed line indicates the stimulus onset time, and the horizontal solid line represents the condition where both leftward and rightward microsaccades are equally likely.

stimulus onset, followed by a slight bias away from the stimulus location 500 ms following stimulus onset. This indicates that a larger proportion of microsaccades occurring during this time interval are congruent to the stimulus location. Taken together, these findings indicate that incongruent microsaccades are less likely to occur during the rebound interval (150–496 ms after stimulus onset), but when they do occur, they are correlated with superior behavioral performance.

Discussion

In the current study, we used microsaccadic activity as a potential attentional marker, to monitor shifts in covert attention while having subjects perform a crowded recognition task. We used a state-of-the-art retinal image-based eye tracking system that offers high-precision eye tracking superior to standard pupil-based eye tracking techniques. To summarize our results, we have shown that the critical spacing — the extent of crowding — is reduced when incongruent microsaccades occur during the enhancement interval following stimulus onset. Given that the strongest association between spatial attention and microsaccadic direction was found for microsaccades that occur during the enhancement interval (Laubrock et al., 2010), this suggests that involuntary shifts in attention behave in agreement with the previously observed reduction in critical spacing when attention was pre-cued to the location of a crowded target (Chakravarthi & Cavanagh, 2009; Freeman & Pelli, 2007; Yeshurun & Rashal, 2010).

The microsaccadic rate was found to follow the frequently observed temporal rate signature, with an initial decrease (inhibition) following stimulus onset quickly followed by an increase (enhancement). Recently, it has been proposed that microsaccadic commands are executed following a rise-to-threshold mechanism which repeats after each microsaccade. But if the stimulus is presented during this time it is thought to alter the ongoing buildup in activity and in turn result in cancellation of microsaccades (Hafed & Ignashchenkova, 2013). In the current study, we observed that with each occurrence of a display change (target or mask onset) there was an associated drop in microsaccadic frequency approximately 100 ms later, causing dips around 150 ms and 350 ms, in agreement with previous observations (Laubrock et al., 2005).

The relationship of microsaccade direction and attention is less straightforward than the stimulus-driven effect on rate, varying based on microsaccade time, cue type (endogenous versus exogenous), response type and cue modality (Engbert & Kliegl, 2003; Laubrock et al., 2005; Rolfs, Engbert, & Kliegl, 2004; Rolfs, Engbert, & Kliegl, 2005). In previous studies that used the classical spatial-cueing paradigms (Posner, 1980), microsaccades were initially found to be more congruent to the cued location (Engbert & Kliegl, 2003; Laubrock et al., 2005; Rolfs et al., 2005) followed by a shift to more incongruent ones (Laubrock et al., 2005; Rolfs et al., 2004; Tian, Yoshida, & Hafed, 2016), with the timing of the effects driven by the cue type used. The former could be controlled by a more reflexive pathway with contributions from the visual and lateral intraparietal cortex (LIP; Engbert,

2006). The latter, however, could be driven by voluntary attempts to maintain fixation whereas exogenous cues are presented in the visual periphery, where subjects are required to inhibit automatic attempts to foveate or make large saccades in the direction of the cue (Rolfs et al., 2004). In other words, salient targets produce an "attend to me signal" that needs to be suppressed to be able to sustain fixation (Sawaki & Luck, 2010). This inhibition alters the activity among the population of neurons in the rostral pole of the superior colliculus (SC), which is responsible for maintaining fixation (Hafed, 2011; Hafed & Krauzlis, 2012), inducing a bias for making small movements in the opposite hemifield (Engbert, 2006; Rolfs et al., 2004). Here, we did not have an explicit cue to indicate the future target location, as we were more interested in understanding the spontaneous shifts in attention that occur following a briefly presented stimulus; we assumed that subjects would naturally shift their attention to the target location following onset. We found that microsaccades were directed toward the target location soon after target onset, whereas those that occurred late in the trial were found to be directed away from the target location. This direction signature is in good agreement with the effects of exogenous cues seen earlier (Hafed & Clark, 2002; Laubrock et al., 2005). On the other hand, behavioral performance was found to be better on trials that contained microsaccades directed away from the stimulus location during the enhancement interval, even though they were fewer in number. This finding, although counterintuitive at first, agrees with previous literature showing that microsaccade direction can be highly predictive of performance, shown previously in cued detection tasks, by proportion correct differences (Hafed & Clark, 2002), and by reaction time benefits (Laubrock et al., 2010). The implication is that microsaccadic direction suggests a stronger capture of attention by the target.

Although there is no consensus on the exact relationship between spatial attention and visual crowding, several theories have been proposed to explain why attentional cues can help diminish the effects of crowding. He, Cavanagh, and Intriligator (1996) proposed that crowding is due to the limited resolution of the attentional system. The reduction of crowding due to attentional cues could then be explained by increased resolution at attended locations (Intriligator & Cavanagh, 2001), or to a smaller "selection region" (Yeshurun & Rashal, 2010). Similarly, Strasburger (2005) suggested that crowding may be associated with unfocussed spatial attention, which can be transiently focused with a cue to enhance performance. In addition, several studies have examined the amount of information in crowded displays. Scolari et al. (2007) proposed that cueing may reduce the internal noise of the visual system. Finally, Dakin,

Bex, Cass, and Watt (2009) found that crowding adds noise to each target, whereas attention modulates sampling efficiency, and that the effects are separable. Because our study was not designed to investigate these effects, and used an implicit measure of spatial attention, it is difficult to speculate on the exact nature of the enhancement we observed, although the psychometric function shifts observed may imply an overall enhancement, rather than a benefit specific to crowding (the results from unflanked trials are shown in Figure B1).

Generally, visual information from a target is thought to accrue as a function of time following target onset to a maximal level or asymptote. However, this accrual of information would be interrupted if the target was quickly followed by a mask (backward mask) or when subjects are required to respond quickly, in which case the processing of information would stop at a pre-asymptotic level (Carrasco, Williams, & Yeshurun, 2002). Visual masking occurs when a mask occurs close to the target both in a spatial as well as temporal sense, resulting in a reduction of the visibility of the target (Bachmann, 1994; Breitmeyer, Ogmen, 2006). In the current experiment, to avoid ceiling effects in behavioral performance, we presented a visual mask immediately after stimulus offset at the same spatial location of the target. Several studies have shown an interaction between spatially non-overlapping backward masks (meta contrast masking) and spatial attention (Ramachandran & Cobb, 1995; Shelley-Tremblay & Mack, 1999; Tata, 2002), but their study design was found to be limited by ceiling and floor effects. However, studies that accounted for these methodological limitations found meta contrast masking and attention (endogenous and exogenous) to not have a strong interaction between them (Ağaoğlu, Öğmen, & Chung, 2016; Ağaoğlu, Breitmeyer, & Ogmen, 2018). Given the mixed findings on the interaction between attention and masking, we wanted to ensure that the attentional effects on visual crowding observed in the current study were not driven by backward masking. To this end, we conducted a control experiment with same set of subjects, the experiment sequence was similar to that of the main experiment, with the only difference being the absence of a post-mask following target offset. We tested a smaller, unmasked letter at a single eccentricity (3 degrees) and a single target – flanker spacing (2×) and found a benefit in performance similar to the main experiment on trials where incongruent microsaccades occurred between 150 to 496 ms following stimulus onset (as shown by Figure C1). This shows that even in the absence of a backward mask the effect of attention was found to emerge. Our findings from the control experiment agree with previous studies that found an attentional benefit in performance with and without a post-mask in the case of unflanked (Carrasco et al.,

2002) as well as flanked targets (Yeshurun & Rashal, 2010).

Because we conducted our eye movement analysis offline, we needed to ensure that subjects' gaze did not stray too close to the stimulus location prior to stimulus onset. Additionally, we aimed to ascertain if our results could be explained by a bias in eye position toward or away from the stimulus at the time of onset. To this end, we looked at the mean horizontal eye position at different time points between 200 ms before and 800 ms after stimulus onset. The eye positions were first normalized to the mean horizontal eye position for each subject prior to conducting our analysis. We only looked at trials that had microsaccades in the time window between 150 and 496 ms following stimulus onset where they were found to affect behavioral performance. We observed a slight bias of 0.07 degrees in eye position toward the target on trials that had incongruent microsaccades at the time of stimulus onset. Similarly, we found a bias of 0.05 degrees away from the target on trials with congruent microsaccades (as shown by Figure D1). Given that there was a slight bias in eye position, we conducted a subsequent analysis to determine whether the improved performance for the trials that had incongruent microsaccades could be driven by a bias toward the stimulus location. By binning each trial based on the actual retinal position at the time of stimulus onset, we found that behavioral performance was still better on trials with incongruent microsaccades compared to ones that had congruent microsaccades or no microsaccades, regardless of eye position (shown in Figure D2). This indicates that the benefit observed with incongruent microsaccades is independent of biases toward the stimulus location. Taken together, the slight shift toward or away from the stimulus location could result in incongruent or congruent microsaccades shifting gaze back to the center of the fixation target, but the bias itself was not causal of the enhancement in performance observed in the current study.

We observed eccentricity-dependent differences in the relationship of incongruent microsaccades and crowding, with the greatest effect at the near and intermediate eccentricities, whereas no effect was observed at the largest eccentricity (5 degrees) tested. First, the psychometric functions in Figure 6 show that the stimulus size scaling was effective in equating performance across eccentricities. The patterns in Figures 6 and 7 reveal that crowding increased similarly with eccentricity in the no microsaccade and incongruent microsaccade condition, such that the incongruent microsaccade condition had a critical spacing that was approximately 25% smaller than the no microsaccade condition. On the other hand, congruent microsaccades, while being similar to the no microsaccade condition at 1 degree, shifted toward the incongruent condition as eccentricity increased. One

proposal from Strasburger (2005) to explain his finding of a lack of a cueing benefit for crowding at 4 degrees (versus 1 degree and 2 degrees) was that cueing may have a steeper M-scaling function (slope of the effect strength versus eccentricity). Similarly, it is possible that the results we observed are due to differential effects of eccentricity on crowding and attention.

Peri-saccadic changes in visual perception are thought to occur primarily for briefly (<45 ms stimulus duration) presented stimuli and have been shown to be weak for longer (>100 ms) stimulus durations (Born, Krüger, Zimmerman, & Cavanagh, 2016). Reduced crowding for briefly presented saccade targets in the farther periphery has also been reported (Harrison, Mattingley, & Remington, 2013; Lin, Rizak, Ma, Yang, Chen, & Hu, 2014), although the findings remain in debate (Ağaoğlu, Breitmeyer, & Ogmen, 2016; Ağaoğlu & Chung, 2017; Buonocore, Fracasso, & Melcher, 2017). However, in the current study, the stimuli were presented for 100 ms, a duration in which the peri-saccadic shifts rarely occur. Had this phenomenon been present, there would have been a difference in behavioral performance in trials with congruent microsaccades within 50 ms before or after stimulus onset, an effect clearly absent (see Figure 5).

In summary, we used microsaccadic behavior following stimulus onset as a potential correlate of trial-to-trial shifts in covert attention and determined how these shifts affected behavioral performance. We found a decrease in critical spacing only on trials that had incongruent microsaccades during the enhancement interval following stimulus onset. This finding of reduced critical spacing when the target captures attention aligns with previous studies that explicitly cued the stimulus location with exogenous cues (Chakravarthi & Cavanagh, 2009; Freeman & Pelli, 2007; Strasburger, 2005; Yeshurun & Rashal, 2010).

Keywords: visual crowding, fixational eye movements, spatial attention

References

Ağaoğlu, M. N., & Chung, S. T. L. (2017). Interaction between stimulus contrast and pre-saccadic crowding. *Royal Society Open Science, 4*(2), 160559.

Ağaoğlu, M. N., Öğmen, H., & Chung, S. T. L. (2016). Unmasking saccadic uncrowding. *Vision Research, 127*, 152–164.

Ağaoğlu, M. N., Sit, M., Wan, D., & Chung, S. T. L. (2018). ReVAS: An open-source tool for eye motion extraction from retinal videos obtained with scanning laser ophthalmoscopy. *Investigative Ophthalmology & Visual Science, 59*(9), 2161–2161.

Ağaoğlu, S., Breitmeyer, B., & Ogmen, H. (2016). Metacontrast masking and attention do not interact. *Attention, Perception, & Psychophysics, 78*(5), 1363–1380.

Ağaoğlu, S., Breitmeyer, B., & Ogmen, H. (2018). Effects of Exogenous and Endogenous Attention on Metacontrast Masking. *Vision, 2*(4), 39.

Bachmann, T. (1994). *Psychophysiology of visual masking: The fine structure of conscious experience*. Hauppauge, NY: Nova Science Publishers Incorporated.

Baumeler, D., Schönhammer, J. G., & Born, S. (2020). Microsaccade dynamics in the attentional repulsion effect. *Vision Research, 170*, 46–52.

Blair, R. C., & Karniski, W. (1993). An alternative method for significance testing of waveform difference potentials. *Psychophysiology, 30*(5), 518–524.

Born, S., Krüger, H. M., Zimmermann, E., & Cavanagh, P. (2016). Compression of Space for Low Visibility Probes. *Frontiers in Systems Neuroscience, 10*, 10.

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature, 226*(5241), 177–178.

Breitmeyer, B., & Öğmen, H. (2006). *Visual Masking: Time Slices Through Conscious and Unconscious Vision*. Oxford Psychology Series. Oxford, UK: Oxford University Press.

Buonocore, A., Fracasso, A., & Melcher, D. (2017). Pre-saccadic perception: Separate time courses for enhancement and spatial pooling at the saccade target. *PLoS One, 12*(6), e0178902.

Carrasco, M., Penpeci-Talgar, C., & Eckstein, M. (2000). Spatial covert attention increases contrast sensitivity across the CSF: Support for signal enhancement. *Vision Research, 40*(10), 1203–1215.

Carrasco, M., Williams, P. E., & Yeshurun, Y. (2002). Covert attention increases spatial resolution with or without masks: Support for signal enhancement. *Journal of Vision, 2*(6), 4.

Chakravarthi, R., & Cavanagh, P. (2009). Bilateral field advantage in visual crowding. *Vision Research, 49*(13), 1638–1646.

Coates, D. R., Chin, J. M., & Chung, S. T. L. (2013). Factors Affecting Crowded Acuity: Eccentricity and Contrast. *Optometry and Vision Science : Official Publication of the American Academy of Optometry, 90*(7), 628.

Coates, D. R., Ludowici, C. J. H., & Chung, S. T. L. (2021). The generality of the critical spacing for crowded optotypes: From Bouma to the 21st century. *Journal of Vision, 21*(11), 18.

Dakin, S. C., Bex, P. J., Cass, J. R., & Watt, R. J. (2009). Dissociable effects of attention and crowding on orientation averaging. *Journal of Vision, 9*(11), 28–28.

Ditchburn, R. W. (1955). Eye-movements in relation to retinal action. *Optica Acta, 1*, 171–176.

Ditchburn, R. W., Fender, D. H., & Mayne, S. (1959). Vision with controlled movements of the retinal image. *The Journal of Physiology, 145*(1), 98–107.

Duda, R. O., & Hart, P. E. (1972). Use of the Hough transformation to detect lines and curves in pictures. *Communications of the ACM, 15*(1), 11–15.

Efron, B. (1982). *The jackknife, the bootstrap, and other resampling plans*. CBMS-NSF Regional Conference Series I Applied Mathematics. Philadelphia, PA: Society for Industrial and Applied Mathematics.

Engbert, R. (2006). Microsaccades: A microcosm for research on oculomotor control, attention, and visual perception. In *Progress in Brain Research* (Vol. *154*, pp. 177–192). New York, NY: Elsevier.

Engbert, R., & Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. *Vision Research, 43*(9), 1035–1045.

Felisberti, F. M., Solomon, J. A., & Morgan, M. J. (2005). The Role of Target Salience in Crowding. *Perception, 34*(7), 823–833.

Foster, D. H., & Bischof, W. E. (1997). Bootstrap estimates of the statistical accuracy of thresholds obtained from psychometric functions. *Spatial Vision, 11*(1), 135–139.

Freeman, J., & Pelli, D. G. (2007). An escape from crowding. *Journal of Vision, 7*(2), 22.

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology, 48*(12), 1711–1725.

Hafed, Z. M. (2011). Mechanisms for generating and compensating for the smallest possible saccades. *European Journal of Neuroscience, 33*(11), 2101–2113.

Hafed, Z. M. (2013). Alteration of Visual Perception prior to Microsaccades. *Neuron, 77*(4), 775–786.

Hafed, Z. M., & Clark, J. J. (2002). Microsaccades as an overt measure of covert attention shifts. *Vision Research, 42*(22), 2533–2545.

Hafed, Z. M., & Ignashchenkova, A. (2013). On the Dissociation between Microsaccade Rate and Direction after Peripheral Cues: Microsaccadic Inhibition Revisited. *Journal of Neuroscience, 33*(41), 16220–16235.

Hafed, Z. M., & Krauzlis, R. J. (2012). Similarity of superior colliculus involvement in microsaccade and saccade generation. *Journal of Neurophysiology, 107*(7), 1904–1916.

Harrison, W. J., Mattingley, J. B., & Remington, R. W. (2013). Eye Movement Targets Are Released from Visual Crowding. *Journal of Neuroscience, 33*(7), 2927–2933.

He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature, 383*(6598), 334–337.

Hinkley, D. V. (1988). Bootstrap Methods. *Journal of the Royal Statistical Society: Series B (Methodological), 50*(3), 321–337.

Horowitz, T. S., Fencsik, D. E., Fine, E. M., Yurgenson, S., & Wolfe, J. M. (2007). Microsaccades and Attention: Does a Weak Correlation Make an Index?: Reply to Laubrock, Engbert, Rolfs, and Kliegl (2007). *Psychological Science, 18*(4), 367–368.

Huckauf, A., & Heller, D. (2002). Spatial selection in peripheral letter recognition: In search of boundary conditions. *Acta Psychologica, 111*(1), 101–123.

Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology, 43*(3), 171–216.

Jacobs, R. J. (1979). Visual resolution and contour interaction in the fovea and periphery. *Vision Research, 19*(11), 1187–1195.

Ko, H., Poletti, M., & Rucci, M. (2010). Microsaccades precisely relocate gaze in a high visual acuity task. *Nature Neuroscience, 13*(12), 1549–1553.

Laubrock, J., Engbert, R., & Kliegl, R. (2005). Microsaccade dynamics during covert attention. *Vision Research, 45*, 721–730.

Laubrock, J., Kliegl, R., Rolfs, M., & Engbert, R. (2010). When do microsaccades follow spatial attention? *Attention, Perception, & Psychophysics, 72*(3), 683–694.

Levi, D. M. (2008). Crowding—An essential bottleneck for object recognition: A mini-review. *Vision Research, 48*(5), 635–654.

Lin, H., Rizak, J. D., Ma, Y., Yang, S., Chen, L., & Hu, X. (2014). Face Recognition Increases during Saccade Preparation. *PLoS One, 9*(3), e93112.

Martinez-Conde, S., Macknik, S. L., & Hubel, D. H. (2004). The role of fixational eye movements in visual perception. *Nature Reviews Neuroscience, 5*(3), 229–240.

Melnik, N., Coates, D. R., & Sayim, B. (2020). Emergent features break the rules of crowding. *Scientific Reports, 10*(1), 1–10.

Morgan, M. J., Ward, R. M., & Castet, E. (1998). Visual Search for a Tilted Target: Tests of Spatial Uncertainty Models. *The Quarterly Journal of Experimental Psychology Section A, 51*(2), 347–370.

Müller, H. J., & Rabbitt, P. M. (1989). Reflexive and voluntary orienting of visual attention: Time course of activation and resistance to interruption. *Journal of Experimental Psychology: Human Perception and Performance, 15*(2), 315–330.

Nakayama, K., & Mackeben, M. (1989). Sustained and transient components of focal visual attention. *Vision Research, 29*(11), 1631–1647.

Nazir, T. A. (1992). Effects of lateral masking and spatial precueing on gap-resolution in central and peripheral vision. *Vision Research, 32*(4), 771–777.

Pastukhov, A., & Braun, J. (2010). Rare but precious: Microsaccades are highly informative about attentional allocation. *Vision Research, 50*(12), 1173–1184.

Pelli, D. G., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature Neuroscience, 11*(10), 1129–1135.

Petrov, Y., Popple, A. V., & McKee, S. P. (2007). Crowding and surround suppression: Not to be confused. *Journal of Vision, 7*(2), 12.

Poletti, M., Listorti, C., & Rucci, M. (2013). Microscopic Eye Movements Compensate for Nonhomogeneous Vision within the Fovea. *Current Biology, 23*(17), 1691–1695.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*(1), 3–25.

Ramachandran, V. S., & Cobb, S. (1995). Visual attention modulates metacontrast masking. *Nature, 373*(6509), 66–68.

Rolfs, M., Engbert, R., & Kliegl, R. (2004). Microsaccade Orientation Supports Attentional Enhancement Opposite a Peripheral Cue: Commentary on Tse, Sheinberg, and Logothetis (2003). *Psychological Science, 15*(10), 705–707.

Rolfs, M., Engbert, R., & Kliegl, R. (2005). Crossmodal coupling of oculomotor control and spatial attention in vision and audition. *Experimental Brain Research, 166*(3–4), 427–439.

Rucci, M., & Poletti, M. (2015). Control and functions of fixational eye movements. *Annual Review of Vision Science, 1*, 499–518.

Sawaki, R., & Luck, S. J. (2010). Capture versus suppression of attention by salient singletons: Electrophysiological evidence for an automatic attend-to-me signal. *Attention, Perception, & Psychophysics, 72*(6), 1455–1470.

Scolari, M., Kohnen, A., Barton, B., & Awh, E. (2007). Spatial attention, preview, and popout: Which factors influence critical spacing in crowded displays? *Journal of Vision, 7*(2), 7.

Sheehy, C. K., Yang, Q., Arathorn, D. W., Tiruveedhula, P., de Boer, J. F., & Roorda, A. (2012). High-speed, image-based eye tracking with a scanning laser ophthalmoscope. *Biomedical Optics Express, 3*(10), 2611–2622.

Shelley-Tremblay, J., & Mack, A. (1999). Metacontrast Masking and Attention. *Psychological Science, 10*(6), 508–515.

Smith, J. E. K. (1982). Simple algorithms for M-alternative forced-choice calculations. *Perception & Psychophysics, 31*(1), 95–96.

Soo, L., Chakravarthi, R., & Andersen, S. K. (2018). Critical resolution: A superior measure of crowding. *Vision Research, 153*, 13–23.

Steinman, R. M., Haddad, G. M., Skavenski, A. A., & Wyman, D. (1973). Miniature Eye Movement. *Science, 181*(4102), 810–819. JSTOR.

Strasburger, H. (2005). Unfocussed spatial attention underlies the crowding effect in indirect form vision. *Journal of Vision, 5*(11), 8.

Stuart, J. A., & Burian, H. M. (1962). A Study of Separation Difficulty*: Its Relationship to Visual Acuity in Normal and Amblyopic Eyes. *American Journal of Ophthalmology, 53*(3), 471–477.

Tata, M. S. (2002). Attend to it now or lose it forever: Selective attention, metacontrast masking, and object substitution. *Perception & Psychophysics, 64*(7), 1028–1038.

Tian, X., Yoshida, M., & Hafed, Z. M. (2016). A Microsaccadic Account of Attentional Capture and Inhibition of Return in Posner Cueing. *Frontiers in Systems Neuroscience, 10*, 23.

Toet, A., & Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research, 32*(7), 1349–1357.

Tripathy, S. P., & Cavanagh, P. (2002). The extent of crowding in peripheral vision does not scale with target size. *Vision Research, 42*(20), 2357–2369.

Tripathy, S. P., Cavanagh, P., & Bedell, H. E. (2014). Large crowding zones in peripheral vision for briefly presented stimuli. *Journal of Vision, 14*(6), 11.

Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics, 33*(2), 113–120.

Weymouth, F. W., Hines, D. C., Acres, L. H., Raaf, J. E., & Wheeler, M. C. (1928). Visual acuity within the area centralis and its relation to eye movements and fixation. *American Journal of Ophthalmology, 11*, 947–960.

Wilkinson, F., Wilson, H. R., & Ellemberg, D. (1997). Lateral interactions in peripherally viewed texture arrays. *Journal of the Optical Society of America A, 14*(9), 2057.

Yeshurun, Y., & Carrasco, M. (1999). Spatial attention improves performance in spatial resolution tasks. *Vision Research, 39*(2), 293–306.

Yeshurun, Y., & Rashal, E. (2010). Precueing attention to the target location diminishes crowding and reduces the critical distance. *Journal of Vision, 10*(10), 16.

Yeshurun, Y., Rashal, E., & Tkacz-Domb, S. (2015). Temporal crowding and its interplay with spatial crowding. *Journal of Vision, 15*(3), 11.

Zhang, M., Gofas-Salas, E., Leonard, B. T., Rui, Y., Snyder, V. C., & Reecher, H. M. (2021). Strip-based digital image registration for distortion minimization and robust eye motion measurement from scanned ophthalmic imaging systems. *Biomedical Optics Express, 12*(4), 2353.

Zuber, B. L., Stark, L., & Cook, G. (1965). Microsaccades and the Velocity-Amplitude Relationship for Saccadic Eye Movements. *Science, 150*(3702), 1459–1460. JSTOR.

Appendix A: Individual psychometric function

Figure A1 shows the psychometric functions for each individual subject (shown in each row) at each of the different target eccentricities tested (shown in each column). We observed that despite differences in the number of microsaccades between observers, the general effect of incongruent microsaccades occurring with smaller critical spacing is robust across observers and eccentricities.
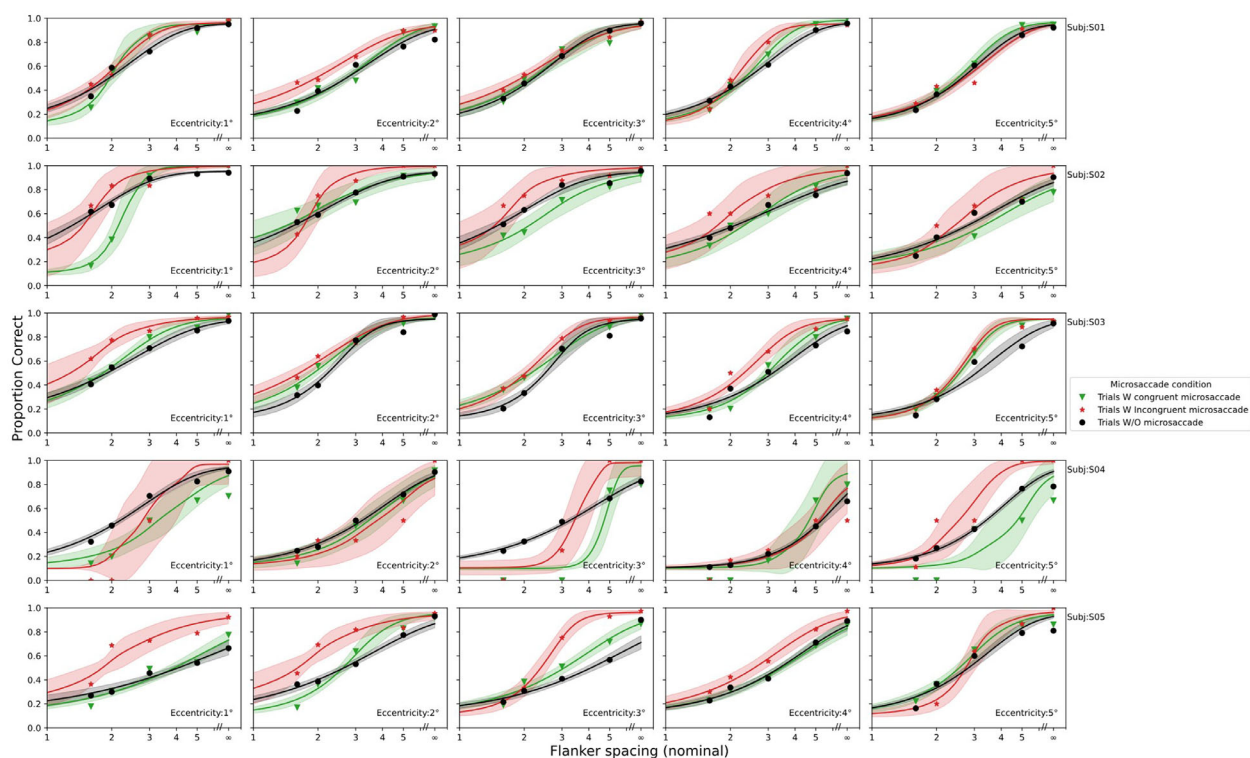
Figure A1. Behavioral performance vs nominal spacing. Each row shows performance as a function target flanker spacing for each subject, while each column represents data from the different target eccentricities tested. The shaded region shows the bootstrapped 95% confidence interval.

Appendix B: Results from unflanked trials

Figure B1 shows the same data as Figure 6, but here we show the results from only the unflanked trials for the
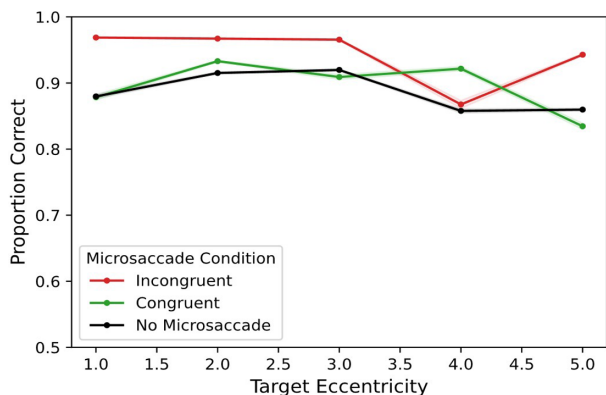

Figure B1. Bootstrapped proportion correct values are plotted against the target eccentricity for the unflanked target condition. Each individual trace represents the data from trials that had either congruent microsaccades (green), incongruent microsaccades (red) or no microsaccade (black) in the interval 150 to 496 ms following stimulus onset. The shaded region represents the bootstrapped 95 % confidence interval.

different conditions grouped based on microsaccadic activity and are plotted against eccentricity of the stimulus. We observed that performance was generally better for unflanked trials when microsaccades occurred 150 to 496 ms following stimulus onset, consistently across the different target eccentricities. Between the microsaccade conditions we found that trials with incongruent microsaccades tended to have better performance for the unflanked trials especially at eccentricities 1, 2, 3 & 5°, but we see them getting worse with incongruent microsaccades at 4°. Hence, we cannot entirely rule out an overall improvement in performance on trials with incongruent microsaccades occurring during the enhancement interval.

Appendix C: Control experiment without post-mask

Figure C1 shows the results from the control experiment with the same set of subjects as the main experiment, with the experiment sequence lacking a post mask following stimulus offset.

Here we tested at an intermediate eccentricity of 3 degrees from fixation with a single target-flanker spacing of 2x the target size. This ensured that we tested at an eccentricity where we observed a benefit
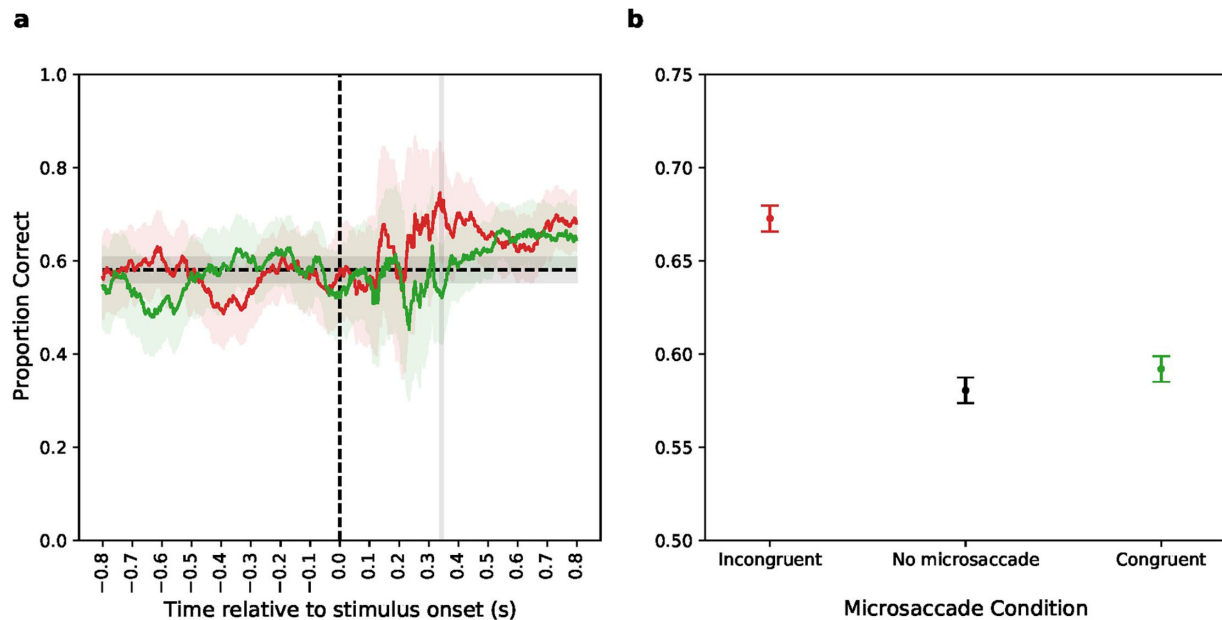
Figure C1. Behavioral performance from the control experiment lacking a post mask following stimulus offset. (a) Shows the mean performance for trials with microsaccades at different time points relative to stimulus onset. The green trace represents trials with microsaccades that were directed towards the stimulus location (congruent), whereas the red trace represents trials where microsaccades were directed away from the stimulus location (incongruent). The black line represents trials without any microsaccades in the interval between 800 ms before and after stimulus onset. (b) Shows the performance measures for a subset of trials that had microsaccades in the time interval 150 ms to 496 ms following stimulus onset, where we found a difference between microsaccade conditions in the main experiment.

from having incongruent microsaccades from the main experiment, similarly we used a spacing that would allow adequate crowding.

As expected, subjects tended be generally better in the absence of a post-mask, hence we reduced the target size to be at threshold size for each subject rather than 1.5 times the threshold size. We observed similar performance levels between the subjects.

We initially looked at behavioral performance on trials that had microsaccades at different time points relative to stimulus onset, as shown by Figure C1a. In agreement with the main experiment, we found a slight benefit in performance when microsaccades were incongruent to stimulus location few hundred ms following stimulus onset. We then looked at behavioral performance for trials that had microsaccades in the time interval of 150 to 496 ms, where we found a difference in the main experiment, as shown by Figure C1b. Here we looked at the performance based on the congruency of microsaccades relative to stimulus location. We can see clearly that performance is generally better when an incongruent microsaccade occurs in the interval 150 to 496 ms following stimulus onset, when compared to trials with congruent microsaccades in the same interval and the ones that did not have microsaccadic events, suggesting that the effect observed in the current study is independent of masking condition.

Appendix D: Effect of bias in eye position on performance

Figure D1 and Figure D2 show the results from the analysis done to ensure that biases in eye position towards or away from the stimulus location were not causal of the enhancement in performance observed in the current study.

Given that we used a global reference frame for each subject, this allowed us to compare eye position measures between trials. Prior to segregating eye positions by the different trial conditions, we normalized the horizontal eye position to the mean horizontal eye position for each subject, where the mean horizontal eye position is assumed to correspond the center of the fixation cross.

We then looked at eye positions only for trials that had microsaccades between 150 and 496 ms following stimulus onset, as shown by Figure D1. We did observe a slight bias of 0.07° in eye positions towards the target on trials that went on to have incongruent microsaccades in the time interval of 150 to 496 ms following stimulus onset. Similarly, we observed a bias of 0.05° away from the target in trials with congruent microsaccades at the time of stimulus onset.
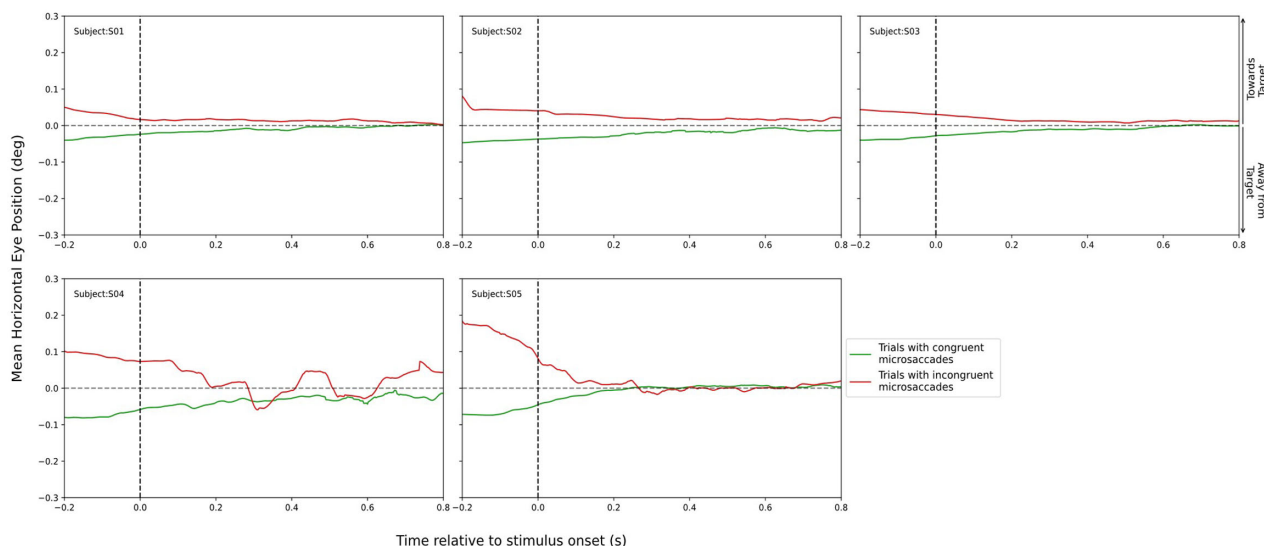
Figure D1. Mean horizontal eye position is plotted against time relative to stimulus onset. Each row represents the data from each subject. The individual traces represent the eye traces from trials that contained either congruent microsaccades (green line) or incongruent microsaccade during the time interval of 150 to 496 ms following stimulus onset.
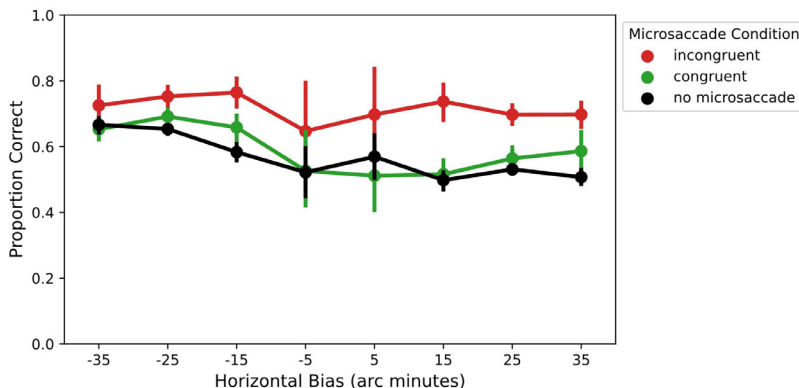

Figure D2. Behavioral performance as a function of bias in horizontal eye position during stimulus onset (0–100 ms).

The fixation cross in our study consisted of a '+' with circle at the center, with each arm of the '+' extending 0.18 deg from the central circle (0.05 deg diameter). The bias in eye position seen here could have originated from slight shifts in eye position when trying to fixate on our fixation target. More experienced observers (S01, S02) had a smaller bias or smaller shifts from the mean horizontal eye position, while less experienced observers (S03, S04, S05) had a larger bias or larger shifts from the mean horizontal eye position. We eventually see them returning to the center of the fixation cross following stimulus offset.

This shows that the shift in horizontal eye position prior to stimulus onset may not be large enough to cause a reduction in the stimulus eccentricity at the time of presentation. However, the slight shift away from the target and towards the target in trials with congruent and incongruent microsaccades could in fact be the result of microsaccades back to the center of the fixation target.

Given that a slight bias in eye position was observed at the time of stimulus onset, we needed to ensure that the improvement in performance for trials with incongruent microsaccades was not driven by the bias towards the stimulus location. By binning behavioral performance based on the eye position at the time of stimulus onset, we observed that the performance was still better for trials that had incongruent microsaccades compared to ones with congruent or no microsaccades. This benefit was found to be independent of any biases in eye position during stimulus presentation (as shown in Figure D2).