Temporal Dynamics of Divided Spatial Attention

Sirawaj Itthipuripat1*, Javier O. Garcia2, & John T. Serences1,2,*

1. Neurosciences Graduate Program, University of California San Diego, La Jolla, California 92093
2. Department of Psychology, University of California San Diego, La Jolla, California 92093

Conflict of Interest
The authors declare no competing financial interests.

Acknowledgements
This study was funded by RO1-MH092345 to JTS. J.O.G was supported by R01-MH68004 to Ramesh Srinivasan. We thank Candace Linscheid, Tony Abuyo, and Danna Lee for help with data collection.

*Correspondence
Neurosciences Graduate Program
Department of Psychology
University of California San Diego
La Jolla, California 92093
Email: Itthipuripat.Sirawaj@gmail.com or jserences@ucsd.edu

Copyright © 2013 by the American Physiological Society.
Abstract

In naturalistic settings, observers often have to monitor multiple objects dispersed throughout the visual scene. However, the degree to which spatial attention can be divided across spatially non-contiguous objects has long been debated, particularly when those objects are in close proximity. Moreover, the temporal dynamics of divided attention are unclear: is the process of dividing spatial attention gradual and continuous, or does it onset in a discrete manner? To address these issues, we recorded steady-state visual evoked potentials (SSVEPs) as subjects covertly monitored two flickering targets while ignoring an intervening distractor that flickered at a different frequency. All three stimuli were clustered within either the lower left or the lower right quadrant, and our dependent measure was SSVEP power at the target and distractor frequencies measured over time. In two experiments, we observed a temporally discrete increase in power for target- vs. distractor-evoked SSVEPs extending from approximately 350-150ms prior to correct (but not incorrect) responses. The divergence in SSVEP power immediately prior to a correct response suggests that spatial attention can be divided across non-contiguous locations, even when the targets are closely spaced within a single quadrant. In addition, the division of spatial attention appears to be relatively discrete, as opposed to slow and continuous. Finally, the predictive relationship between SSVEP power and behavior demonstrates that these neurophysiological measures of divided attention are meaningfully related to cognitive function.
Abbreviate title: Temporal Dynamics of divided spatial attention

Introduction

In everyday perception, organisms must often monitor multiple non-contiguous objects that are arrayed across the visual field. Early models proposed that spatial attention operates as a unitary spotlight that only covers one circumscribed region of the visual field and may switch rapidly among multiple relevant objects (Posner et al., 1980). Later models incorporated the notion of a variable-sized spotlight (zoom lens) that can be reshaped based on perceptual demands (Barriopedro and Botella, 1998; Erkisen and St. James, 1986; Eriksen and Yeh, 1985; Heinze et al., 1994; McCormick and Jolicoeur, 1994; Müller et al., 2003b). In contrast to these earlier models, flexible allocation models postulate that spatial attention can be allocated to non-contiguous regions of space (Awh and Pashler, 2000; Baldauf and Deubel, 2008; Bichot, Cave, and Pashler, 1999; Carlson et al., 2007; Castiello and Umiltà, 1992; Cavanagh and Alvarez, 2005; Dubois et al. 2009; Gobel et al., 2004; Godijn and Theeuwes, 2003; Hahn and Kramer, 1998; Howe et al., 2011; Kraft et al., 2005; Kramer and Hahn, 1995; Malinowski et al., 2007; McMains and Somers, 2004, 2005; Müller et al., 2003a; Niebergall et al., 2010, 2011). However, the extent to which attention can be divided across multiple locations is still controversial (Jans et al., 2010; Cave et al., 2010), and here we investigated three main theoretical issues related to this debate.

First, the extent to which attention can be divided between objects that are in close proximity is unclear, particularly when those items fall within the same hemifield/quadrant of space (Maertens and Pollmann, 2005; Malinowski et al., 2007; Pollmann, Zaidel, and von Cramon, 2003; Sereno and Kosslyn, 1991; McMains and Somers, 2004). Second, few studies have evaluated the temporal dynamics of divided spatial attention, so it is unclear whether the division of spatial attention arises slowly and gradually over time, or whether attention can be divided in a discrete manner analogous to a unitary spotlight that splits into multiple foci. Finally, predictive relationships between neurophysiological measures of divided spatial attention and behavior have not been clearly established.

Here, we evaluated these issues by monitoring neural activity associated with attended targets and ignored distractors using steady-state visual evoked potentials (SSVEPs). This method provides a temporally continuous measure of electrophysiological responses that oscillate at the same temporal frequency as a visual stimulus (Regan, 1989), and is thus useful for investigating the temporal dynamics of attentional phenomena (e.g. Andersen and Müller, 2010; Müller et al., 1998). In two experiments, subjects performed a multiple-object discrimination task in which two visual targets and an intervening distractor were presented in either the lower left or the lower right quadrant (Figure 1). The targets and the distractor were presented at different frequencies to elicit separable SSVEPs. In addition, we calibrated task difficulty for each subject by adjusting the contrast of the targets to yield approximately 66% correct responses. By setting accuracy at this below-ceiling level, we sought to ensure that we had enough power to compare stimulus- and response-locked SSVEP responses on correct and incorrect trials.

Using this approach, we did not find robust evidence that differences in stimulus-locked target and distractor SSVEP power predicted behavioral performance. However, when SSVEP power was time-locked to the behavioral response, we observed an abrupt divergence of target- and distractor-evoked SSVEP power immediately preceding a correct behavioral response. These results suggest that spatial attention can be divided between objects in the same visual quadrant, that this division happens in a relatively discrete manner, and that the degree of segregation between target and distractor evoked responses predicts behavioral performance.

Materials and Methods

Subjects
Abbreviate title: Temporal Dynamics of divided spatial attention

Forty-eight neurologically healthy volunteers with normal or corrected-to-normal vision were recruited from the University of California, San Diego (UCSD). All participants provided written informed consent in accordance with the human subjects Institutional Review Board at UCSD. Eleven subjects (4 females, mean age, 22) participated in SSVEP experiment 1 (E1), fifteen subjects (9 females, mean age, 21) participated in SSVEP E2, ten subjects (7 females, mean age, 23) participated in behavioral control experiment 1 (CE1), and thirteen subjects (7 females, mean age, 21) participated in behavioral CE2. Data from one subject in E2 were excluded from analysis due to excessive eye movement artifacts (see below). All subjects received either class credit or monetary compensation for participation ($15/hour for E1 and E2 and $10/hour for CE1 and CE2).

Stimuli and Experimental Design

The experiment was carried out using a PC running Windows XP. Stimulus presentation was controlled using MATLAB (Mathworks Inc., Natick, MA) with the Psychophysics Toolbox (version 3; Brainard, 1997; Pelli, 1997), and all stimuli were rendered on a 20” CRT monitor (hp p1203, Hewlett-Packard Company, Palo Alto, CA) running at 75Hz. Subjects were seated 60 cm from the computer screen and instructed to fixate at the center of the screen for the duration of each trial. Stimulus timing was directly measured via fiber-optic cables that were linked to photocells and attached to the stimulus monitor. The gain and sensitivity of this custom-built fiber optic photocell system (Javi’s Phototransmogrifier 2000, Electronic Development and Repair Facility, Department of Physics, UCSD) was adjusted to give accurate and fast sampling of stimulus parameters (i.e., precise timing of the flickering stimuli), and the output was fed directly into the electroencephalogram (EEG) data files for later use during data analysis.

An example of the stimulus presentation paradigm is illustrated in Figure 1. Before the start of each trial, a central fixation point was presented for 700ms, followed by the onset of three white squares (72 cd/m^2) that were presented at an isoecentric distance from fixation (9.5° visual angle) in either the lower left or the lower right quadrant (positions in the lower left and lower right quadrants, arranged from upper to lower eccentric locations, are hereby referred to as L1, L2, L3 and R1, R2, R3, respectively; see Figure 1). The background was middle grey (mean luminance of 21 cd/m^2). Each square stimulus subtended 2.4°x2.4° visual angle with equal spacing between each square (3.8° visual angle between the centers of adjacent squares). The two outer squares were always to-be-attended (L1/L3 or R1/R3), and each square contained a small light gray rectangle-shaped mark (0.1°x0.2° visual angle; 0.035% of the area of each square stimulus) on one of its edges. The contrast between the attended stimuli and the small markings was adjusted across a range from 10-30% depending on subjects’ performance on the block-by-block basis (see details below). The marked location on the two attended stimuli (L1/L3 or R1/R3) could either be the same, offset clockwise or counterclockwise by 90°, or offset by 180°; each of these target configurations was presented with equal probability. The middle square (L2 or R2) was located between the two targets and served as a distractor that contained no additional markings. During the experiment, subjects were asked to respond to the target configuration by pressing one of three buttons on a keyboard with their right index, middle, or ring finger indicating if the marked locations targets were in the same orientation, offset clockwise or counterclockwise by 90° or offset by 180° (yielding a three alternative forced choice design, or 3AFC). The small and low contrast markings on the attended stimuli were designed to make the task very difficult and to encourage subjects to continuously attend to both target objects for an extended period of time.

In E1, we separately assessed target- and distractor-related SSVEPs by flickering the two targets at the same frequency while the middle distractor was flickered at a different frequency (the targets and distractor were flickered at either 6.9Hz or 10.4Hz, counterbalanced within subjects across repeated
testing blocks). To discourage the possibility that divided attention was supported by grouping based on target frequency, and to ensure that our results were not specific to our exact choice of stimulus frequencies in E1 (6.9 and 10.4 Hz), we conducted E2 using a similar design except that all three squares were flickered at different frequencies (15 Hz, 18.76 Hz, and 25.02 Hz). The six possible frequency combinations were counterbalanced within observer across blocks of trials. In both experiments, the stimulus array was presented for 3200 ms on each trial with an inter-trial interval (ITI) of 1500 ms. Our main focus was to compare responses on correct and incorrect trials, as opposed to comparing responses to each stimulus when it was attended versus ignored. While this design deviates from most previous efforts, it allowed us to link EEG response modulation to behavioral performance while still maintaining perfectly matched sensory stimulation across conditions (since the stimuli were identical on correct and incorrect trials). To ensure that enough trials were acquired to support a comparison of SSVEP power on correct and incorrect trials, the luminance of the small square markings on each target stimulus was adjusted on a block-by-block basis to maintain performance at approximately 66% correct (which is 2 times chance given the 3AFC design). Subjects were instructed to respond as quickly and accurately as possible, and were free to respond before the termination of the stimulus array.

Before each recording session, subjects practiced for approximately 20 minutes to gain familiarity with the task and to calibrate initial difficulty thresholds. Each recording session consisted of 10 blocks of 60 trials, and contained an equal number of stimulus arrays presented in the lower left and right quadrants (in a pseudo-randomly selected order). Each block lasted approximately 5 minutes resulting in a 1.5-hour-long experiment session, including EEG preparation and short breaks in between each block.

In addition, we also ran two additional behavioral control experiments (CE1 and CE2) to evaluate the efficacy of the intermediate distractor. CE1 was identical to E1 and CE2 was identical to E2 except that only ½ of the trials contained the intermediate distractor and the other half did not.

**EEG Recording, Preprocessing, and SSVEP Analysis**

EEG data were recorded using a 128-channel Geodesic Sensor Net coupled with a NetAmps 200 amplifier (Electrical Geodesics Inc; EGI, Eugene, OR). EEG data were sampled at 1000 Hz and referenced to the central channel. Electrode impedances were kept below 50 kΩ, which is standard with high input impedance amplifiers like the EGI system. Blinks and vertical eye movements were monitored by four built-in electrodes placed above and below the left and right eyes. Horizontal eye movements were monitored by two built-in electrodes placed at the outer canthi of the eyes. Blink suppression and fixation control were verbally encouraged throughout the experiment.

We used customized MATLAB scripts for EEG preprocessing and SSVEP analysis. First, we applied a high-pass Butterworth filter with 2-dB attenuation at 2 Hz to remove slow drifts and a bandstop Butterworth filter with 30-dB attenuation between 58-62 Hz to attenuate line noise. Second, the continuous EEG recordings from each subject were segmented into single-trial epochs extending from 1000 ms before to 3500 ms after stimulus onset. Individual trials that exhibited prominent blink, electro-oculogram (EOG), or electromyogram (EMG) artifacts were discarded using threshold rejection (more than ±100 µV deviation from the mean) and visual inspection, which resulted in the removal of <15% of trials across subjects [with the exception of the single subject in E2 that was discarded from further analysis due to ~45% rejected trials based on these criteria]. Principal components were then computed and selected for removal based on visual analysis to attenuate any residual artifacts. Third, each of the trials was cropped to an integer number of cycles using the photocell traces that directly measured stimulus onset timing during each trial. Fourier coefficients were then computed on the epoched data
from each trial using the Fast-Fourier transform algorithm implemented in Matlab. Power across the entire trial interval at each target and distractor frequency was calculated as the squared real component of the Fourier coefficients.

The power of the SSVEP response over the entire trial at each target- and distractor-specific frequency was then binned based on spatial position (left or right) and on the subject’s performance (correct response or incorrect response; all trials in which a response was omitted were excluded from the SSVEP analysis, see Results section below). For the main analysis, we selected three EEG channels in the occipital-parietal area that exhibited the highest SSVEP power as electrodes of interest (EOIs); this selection process was carried out separately for each stimulus location and frequency assignment. Separate electrodes were used in each condition because of expected differences in distribution of SSVEP power across the scalp as a function of target position and frequency. For example, Figure 2 illustrates the topographical maps of averaged SSVEP power collapsed across all frequency assignments in E1 (top panel) and in E2 (bottom panel). SSVEP power was standardized across channels so that the positive and negative values indicate responses above and below the mean value, respectively (in units of z-scores). Note that the EOIs for both target- and distractor-evoked SSVEPs were clustered in the occipital-parietal area contralateral to the locus of visual stimulation, consistent with the known contralateral mapping of external spatial locations to internal cortical representations. To evaluate the reliability and the specificity of power at each stimulus presentation frequency, we compared the magnitude of the SSVEP response across the entire trial divided by the power in four surrounding frequency bins. This produces a measure of the signal-to-noise (SNR) ratio of the response at the specific flicker frequency of each stimulus (see e.g. Kim and Verghese; 2012, Sutoyo and Srinivasan, 2006; Srinivasan et al., 2007). Figure 3 illustrates these SNR measures at the flicker frequencies of 6.9Hz (a) and 10.4Hz (b) in E1 and at the flicker frequencies of 15Hz (c), 18.76Hz (d), and 25Hz (e) in E2. Across all flicker frequencies, there is a clear and sharply tuned increase in power.

Finally, we calculated the power at each stimulus frequency across time from each set of EOIs using an analytic Gabor basis function (Gaussian-weighted complex-valued sinusoid; see Bruns 2004; Canolty et al., 2007) with a fractional bandwidth of 0.5 Hz and a time-domain standard deviation that varied with the frequency of the stimulus (36ms and 54ms for the 10.4Hz and 6.9Hz stimuli in E1, and 15ms, 20ms, and 25 ms for the 25Hz, 18.76Hz, and 15Hz stimuli in E2; Canolty et al., 2007). The resulting SSVEP power timecourses were then averaged at each time point across the three electrodes of interest. Normalization was performed by dividing the power at each time point by the mean power from 500 to 0 ms before stimulus onset, and then this baseline was subtracted. We then collapsed the data across all flicker frequencies assigned to like conditions and across left and right stimulation epochs. The target and distractor responses for correct and incorrect trials were then time-locked to either stimulus onset or response onset, and averaged across all subjects.

Statistical Analysis

Average accuracy was calculated separately for stimulus arrays presented in the left and right visual fields and also for each combination of target and distractor frequency assignments (two combinations in E1, and six combinations in E2). Response times (RTs) were analyzed in a corresponding manner separately on correct and incorrect trials. In E1, two-way repeated-measures analysis of variance (ANOVA) tests were used to evaluate accuracy and RT data, with separate factors for stimulus location (left, right) and target frequency (6.9Hz, 10.4Hz). In E2, a similar statistical approach was performed, but there were six levels of the target frequency factor. False Discovery Rate (FDR) corrected post-hoc t-tests were also performed to further examine any significant effects revealed by the repeated-measures ANOVAs. For the analysis of stimulus-locked
SSVEP power and response-locked SSVEP power, we first performed repeated-measures *t*-tests to test for differences between target- and distractor-evoked responses at each timepoint on correct and incorrect trials. To correct for family-wise error rates, we then calculated the *p* values associated with the *t* scores for each time point. We sorted the *p* values in ascending order (*p₁*, *p₂*, *p₃*, …, *pₖ*, …, *p_M*), and determined a threshold according to the following: *pₖ > (K*α)/M*, where *pₖ* is the FDR-corrected *α* value (*p < 0.05*), *K* is the rank of the *p* value corresponding to the threshold, *α* is 0.05, and *M* is the total number of comparisons. All significant effects were defined as having a *p*-value lower than 0.05 after FDR correction. In addition, repeated-measures ANOVAs were performed on SSVEP power averaged across time windows identified as being significant based on the *t*-tests to evaluate possible interactions between accuracy (correct/incorrect) and stimulus type (target/distractor). FDR correction was also applied to the analysis of behavioral data.

**Results**

**Behavioral results**

In E1, the average percent correct, percent incorrect, and percent misses, collapsed across hemifields and flicker frequencies, were 65.4% (SD=8.3%), 25.8% (SD=11.0%), and 8.8% (SD=6.7%), respectively. A two-way repeated measures ANOVA revealed no effect of stimulus location (*F*(1,10) = 2.44, *p = 0.15), target frequency (*F*(1,10) = 0.004, *p = 0.95), and no interaction between these factors on accuracy (*F*(1,10) = 0.12, *p = 0.74). Average RTs on correct and incorrect trials were 1,980ms (SD=114.1ms) and 2092ms (SD=171.6ms), respectively. Two-way repeated measures ANOVAs revealed no main effect of stimulus location or flicker frequency on either correct or incorrect RTs, and no interaction between stimulus location and flicker frequency (all *F*-values ≤ 1.08 all *p*'s ≥ 0.32).

In E2, the average percent correct and percent incorrect, collapsed across hemifields and flicker frequencies, were 68.5% (SD=7.4%) and 31.5% (SD=7.4%), respectively (with no misses). A two-way repeated-measures ANOVA revealed a significant effect of target frequency on accuracy (*F*(5, 65) = 11.86, *p < 0.001), but no effect of stimulus location (*F*(1, 13) = 0.031, *p =0.826) and no interaction between target frequency and stimulus location (*F*(5, 65) = 1.024, *p = 0.41). Post-hoc *t*-tests revealed that the main effect of frequency resulted from higher performance when stimuli in positions 1, 2, and 3 were flickered at 25Hz, 18.76Hz, and 15Hz respectively (compared to the other 5 combinations), and higher performance when stimuli were flickered at 18.76Hz, 15Hz and 25Hz compared to when stimuli were flickered at either 18.76Hz, 25Hz, and 15Hz or 15Hz, 25Hz, and 18.76Hz (all *t*-values > 2.43, all *p*'s < 0.05 FDR-corrected). This slight impairment when the distractor frequency was higher than both target frequencies might be due to the fact that lower frequencies translate into fewer presentation cycles across a fixed stimulus duration. However, this scenario was rare (only 9.5% of all comparisons) and we did not observe this effect in E1. Therefore, it is difficult to determine the ultimate cause of this frequency assignment effect in E2. Average RTs for correct and incorrect trials were 1818ms (SD=134ms) and 1847ms (SD=231ms). Two-way repeated-measures ANOVAs on the RT data revealed no significant effects of either frequency or stimulus location, and no interaction between these factors (this was true for RTs on both correct and incorrect trials, all *F*-values ≤ 1.89, all *p*'s ≥ 0.107).

We also computed d-prime (d') on trials where the small target squares were offset by 0°, 90°, and 180° degrees (where hits were defined as reporting that the offset was X° when it was in fact X°, and false alarms were defined as reporting that the offset was X° when it was either Y° or Z°). No significant differences in the sensitivity were observed as a function of target offset (d' for E1: 1.02(SD=0.72), 1.13 (SD=0.47), 1.05 (SD=0.66) for 0°, 90°, 180° offsets, respectively, *F*(2, 20) = 0.698, *p = 0.509; d' for E2: 0.97(SD=0.49), 0.97 (SD=0.32), 0.85 (SD=0.41) for 0°, 90°, 180° offsets, respectively, *F*(2, 28) = 2.452, *p = 0.104). However, there was a modest but reliable difference in RTs...
on correct trials as a function of target offset (for E1, $F(2,20) = 7.169, p = 0.004$; for E2, $F(2,28) = 4.181, p = 0.029$). Post-hoc t-tests revealed that in E1, responses on $0^\circ$ offset trials ($1921.5\text{ms, } SD = 168.3\text{ms}$) were significantly shorter than responses on either $90^\circ$ ($2009.8\text{ms, } SD = 193.8\text{ms, } t(10) = 3.063, p = 0.012$) or $180^\circ$ trials ($2003.6\text{ms, } SD = 183.7\text{ms, } t(10) = 4.165, p = 0.002$). In E2, there was only a difference between responses on trials with a $0^\circ$ degree offset ($1783.3\text{ms, } SD = 155.1\text{ms}$) and trials with a $180^\circ$ offset ($1855.6\text{ms, } SD = 183.2\text{ms, } t(13) = 2.640, p = 0.02$).

SSVEP results

Figure 4 (top panel) shows stimulus-locked SSVEP power across the entire stimulus interval for correct and incorrect trials in E1. Power increased significantly above baseline from $\sim 50$-3200ms post-stimulus for both correct and incorrect target- and distractor-evoked SSVEPs (minimum t-value across this temporal window: $t(10) > 2.25, p < 0.05$). There were no significant differences in SSVEP power associated with targets and distractors across time on either correct (all $t(10) < 2.01$, n.s.) or incorrect trials (all $t(10) < 1.41$, n.s.). However, both targets had a different flicker frequency than the single distractor, and this sensory difference may have obscured any differences in the overall responses to targets and distractors. In contrast, comparisons between target (or distractor) responses on correct and incorrect trials are controlled for sensory differences. However, there were no power differences in the target-evoked SSVEPs on correct and incorrect trials (all $t(10) < 2.90$, n.s.) or in the distractor-evoked SSVEPs on correct and incorrect trials (all $t(10) < 2.49$, n.s.).

When the normalized power was locked to the response, however, significant differences emerged. On correct trials (Figure 4 bottom left), we found a significant increase in the power of the target-evoked SSVEP compared to the distractor-evoked SSVEP extending from 350ms to 150ms before the response (minimum $t(10) > 3.12$, all p’s < 0.05). However, no differences in target- and distractor-evoked SSVEP power were observed on incorrect trials (maximum $t(10) < 3.06$, n.s.; Figure 4 bottom right). A repeated measures ANOVA with factors for accuracy (correct/incorrect) and stimulus type (target/distractor) revealed no main effect of accuracy ($F(1, 10) = 0.55$, n.s.) but a significant main effect of stimulus type ($F(1, 10) = 6.82, p = 0.026$) and a significant interaction between accuracy and stimulus type on SSVEP power averaged across a response window extending from 350 to 150 ms before the behavioral response ($F(1, 10) = 13.39, p = 0.004$).

To address the possibility that attention might be efficiently split only in the trials in which subjects identified the targets quickly, we conducted an auxiliary analysis on the SSVEP data in E1, specifically sorting the data into trials with short and long RTs via a median split. Figure 5 shows the stimulus-locked (left) and response-locked SSVEP data (right) sorted into short RT trials (top) and long RT trials (bottom). As in the main analysis (Figure 4), we observed no significant difference between target- and distractor-evoked SSVEP responses on correct trials at any timepoint ($t(10) < 2.55$, n.s.). However, response-locked SSVEP responses significantly increase right before a correct response is made, and this is true for trials with short RTs ($-350$ to $-150$ms: $t(10) > 3.17, p < 0.05$) and long RTs ($-330$ to $-260$ms: $t(10) > 3.14, p < 0.05$). Importantly, no response-locked differences were found on incorrect trials ($t(10) < 2.31$ for short RTs and $t(10) < 2.08$ for long RTs). This pattern gave rise to a significant interaction between stimulus type (target/distractor) and accuracy on short RT trials ($F(1, 10) = 5.101, p = 0.047$), and a marginally significant interaction on long RT trials ($F(1, 10) = 3.887, p = 0.077$).

To evaluate the possibility that the divergence in the SSVEP power associated with targets and distractors on correct trials was driven by motor activity preceding the response, which was presumably focused contralateral to the right hand, we also tested SSVEP power separately when the stimuli were in the left and right visual fields. On correct trials, we found a significant divergence in target- and
distractor-evoked SSVEP power for stimuli presented in both the left hemifield ($t(10) = 2.34, p = 0.04$, averaged across a window 350 to 150ms before the response) and the right hemifield ($t(10) = 2.89, p = 0.05$, averaged across a window 350 to 150ms before the response). The observation of a qualitatively similar effect in both hemifields argues against motor preparation as the main factor in driving the divergence of SSVEP power associated with targets and distractors on correct trials.

Furthermore, to evaluate the possibility that the increased target-evoked power on correct trials was facilitated by grouping the targets based on a common flicker frequency, we used separate flicker frequencies for all three stimuli in E2 (Figure 6). This experiment also speaks to the ‘motor preparation’ account referred to above, as the flicker rates used in E2 were well above the theta range typically associated with motor preparation (e.g., Luu and Tucker, 2001; Makeig et al., 2004). In addition, assigning (and counterbalancing) a unique flicker frequency to each stimulus allowed for a more balanced comparison between the SSVEP responses associated with targets and distractors (compared to in E1, where both targets shared a common flicker frequency).

The SSVEP results in E2 were highly consistent with those in E1 (Figure 6). For stimulus-locked data (Figure 6 top panel), power increased significantly above baseline from ~50-3200ms post-stimulus for both target- and distractor evoked SSVEPs (minimum t-value: $t(13)> 2.20, p < 0.05$). However, just as in E1, there were no significant differences in SSVEP power between target- and distractor-evoked SSVEPs at any time point for either correct (all $t(13) < 3.30, n.s.$) or for incorrect trials (all $t(13) < 3.36, n.s.$) and no significant differences in power across correct and incorrect trials for either target (all $t(13) < 2.60, n.s.$) or distractor-evoked SSVEP (all $t(13) < 2.49, n.s.$).

When the data were aligned to response onset (Figure 6, bottom left), we found a significant increase in the target-related SSVEP compared to the distractor-related SSVEP across a window extending from 350 to 150ms before the response (collapsed across left and right stimulus presentations: minimum t-value, $t(13) > 2.56$; left stimuli only all $t(13) > 2.49$; right stimuli only all $t(13) > 2.70$; all $p’s<0.05$). No power differences were observed on incorrect trials (maximum t-value, $t(13)< 2.22$, all n.s., Figure 5 bottom right). A repeated measures ANOVA with factors for accuracy and stimulus type revealed no main effect of accuracy ($F(1, 13) = 3.10, n.s.$), but a significant main effect of stimulus type ($F(1, 13) = 7.32, p=0.018$) and a significant interaction between accuracy and stimulus type on SSVEP power averaged across a window from 350 to 150 ms before the response ($F(1, 13) = 9.744, p=0.008$).

Just as in E1, we also analyzed SSVEPs separately on trials with the fastest and slowest RTs. Figure 7 shows the stimulus-locked (left) and response-locked SSVEP data (right) sorted into short RT trials (top) and long RT trials (bottom). When time-locked to stimulus-onset, there were no significant differences between target- and distractor-evoked SSVEP responses on correct trials ($t(13)< 2.837, n.s.$). However, the response-locked SSVEPs were significantly different right before a correct response for both short RT (-400 to -200ms: $t(10)> 3.17, p < 0.05$) and long RT (-300 to -150ms: $t(10)> 3.14, p < 0.05$) trials. Importantly, no response-locked differences were found on incorrect trials ($t(10) < 2.31$ for short RTs and $t(10) < 2.08$ for long RTs). This pattern gave rise to a significant interaction between stimulus type (target/distractor) and accuracy on short RT trials ($F(1, 13) = 6.036, p = 0.029$) and on long RT trials ($F(1, 13) = 11.464, p =0.005$).

Finally, we performed two control experiments (CE1 and CE2) to determine if the interleaved distractors used in E1 and E2 actually interfered with target processing. We found that in both control studies, subjects were significantly slower when an intervening distractor was presented compared to when no distractor was present (CE1: mean RT with a distractor = 1674ms; mean RT without a distractor = 1598ms; $t(9) = 3.606, p=0.0057$. CE2: mean RT with a distractor = 1697ms; mean RT without a distractor = 1611ms; $t(12) = 7.086; p<0.0001$). Subjects were also less accurate when a distractor was present versus absent (CE1: percent correct with a distractor = 69.6%; percent correct
without a distractor = 72.83%; \( t(9) = 2.457, p=0.036 \). CE2: percent correct with a distractor = 71.88%; percent correct without a distractor = 74.62%; \( t(12) = 2.317; p=0.039 \). Together, these two control experiments suggest that the intervening distractors did compete with target processing, and are thus likely to encourage split attention foci.

**Discussion**

To study the temporal dynamics of divided attention, we employed a multiple-object discrimination task, where two visual targets and an intervening distractor were presented in a single quadrant at different flicker frequencies. An analysis of SSVEPs was used to continuously monitor the responses associated with each stimulus. In both experiments, we failed to observe a difference between target- and distractor-evoked responses when the normalized SSVEP power was timelocked to the onset of the stimulus. Such effects may have been obscured in E1, as a direct comparison between target and distractor evoked responses may have been confounded due to the common flicker frequency of the target stimuli. However, a comparison of target and distractor evoked responses on correct and incorrect trials, which controls for this sensory difference, also revealed no stimulus locked effects. Moreover, each stimulus in E2 had a unique frequency, and similar null results were observed. That said, we did not measure responses to the targets and distractors when they were attended and when they were unattended, so this null result should be interpreted with caution. In contrast, however, when the data were locked to the onset of the behavioral response, target-evoked SSVEP power was selectively larger compared to distractor-evoked SSVEP power approximately 150-350 ms before correct (but not incorrect) responses. Taken together, this evidence suggests that (1) attention can be divided across non-contiguous regions of space even when the stimuli are presented in close proximity within the same visual quadrant, (2) the division of spatial attention can occur in a temporally discrete manner and (3) there is a strong predictive relationship between the magnitude of responses tied to spatially distinct stimuli and the success of perceptual decision making.

Over the past several decades, evidence supporting the flexible division of attention has been accumulating based on tasks such as multiple object tracking (reviewed by Cavanagh and Alvarez, 2005). However, studies that have reported neuronal evidence in favor of divided attention typically employed visual stimuli that were spaced quite far from one another, often crossing the vertical meridian or spanning the entire visual field (e.g., McMains and Somers, 2004, 2005; Müller et al., 2003a). In one notable exception, McMains and Somers (2004) found evidence for divided attention between two locations in a single quadrant. However, in their study, one target stimulus was always located at fixation and the other target was located peripherally. Therefore, the extent to which attention can be efficiently divided between objects that are in close spatial proximity is still unclear, particularly when more than one relevant stimulus falls within the same hemifield/quadrant of space (Maertens and Pollmann, 2005; Pollmann, Zaidel, and von Cramon, 2003; Sereno and Kosslyn, 1991). More recently, Malinowski et al. (2007) also tested whether attention could be divided into multiple spotlights when the stimuli were spaced closely together (i.e., within the same hemifield) using an SSVEP design. In contrast to the present study, they did not find consistent evidence for divided attention. Instead, they observed divided attention effects only when subjects attended to stimuli in the upper left but not in the lower left quadrant. The authors suggested that these mixed results may have been caused by the higher perceived saliency of the intermediate distractor in the lower quadrant. However, computing the average SSVEP power over the entire stimulus period (~2-3s) [as done by Malinowski et al. (2007)] may have reduced their ability to detect evidence for divided attention, particularly if the effect has a discrete onset that immediately precedes the behavioral response as in the current study.
Consistent with some aspects of Malinowski et al. (2007), our stimulus-locked analysis yielded no significant power difference between target- and distractor-evoked SSVEPs. At first glance, the results from this stimulus-locked analysis are consistent with spotlight models in which attention is always unitary (Posner et al., 1980), or zoomed/reshaped based on task demands (Barriopedro and Botella, 1998; Erkisen and St. James, 1986; Erkisen and Yeh, 1985; Heine et al., 1994; McCormick and Jolicoeur, 1994; Müller et al., 2003b). However, our data argue against this account based on two key pieces of evidence. First, the unitary spotlight models predict that the overall level of stimulus-locked SSVEP power should predict discrimination success. Instead, we observed that overall stimulus-locked SSVEP power did not predict successful performance on the discrimination task. Second, a unitary spotlight model cannot account for a response-locked increase in SSVEP power that immediately precedes successful discriminations, an observation that strongly favors the divided attention hypothesis (Awh and Pashler, 2000; Baldauf and Deubel, 2008; Bichot, Cave, and Pashler, 1999; Carlson et al., 2007; Castiello and Umiltà, 1992; Cavanagh and Alvarez, 2005; Dubois et al. 2009; Gobel et al., 2004; Godijn and Theeuwes, 2003; Hahn and Kramer, 1998; Howe et al., 2011; Kraft et al., 2005; Kramer & Hahn, 1995; Malinowski et al., 2007; McMains and Somers, 2004, 2005; Müller et al., 2003; Niebergall et al., 2010; 2011). This response-locked increase in power that predicted successful performance was replicated in two experiments, and the temporal window of the divergence between target- and distractor-evoked SSVEP power was consistently found to occur ~350-150ms preceding correct responses. The significant main effect of stimulus type (target vs. distractor) and interaction between stimulus type and accuracy on SSVEP power averaged across a 350-150ms window prior to the response further suggests that successful discrimination was achieved by a temporally discrete segregation of neuronal responses tied to targets and distractors, respectively.

Although we cannot directly assess the degree to which our results were driven by active neuronal suppression, our findings complement recent single-unit electrophysiological data recorded from the middle temporal visual area (MT) of macaques. The monkeys were trained to attend to two translating objects while ignoring an intermediate distractor located inside the receptive field of the neuron that was being monitored (Niebergall et al. 2011). Niebergall et al. (2011) found that when the monkey attended to the two translating objects, neuronal responses associated to the distractor were either suppressed or they remained unaltered, providing strong evidence in support of divided spatial attention. Taken together with the present observations, it is possible that the division of spatial attention is achieved not only by enhancing responses to attended targets but also by actively suppressing responses to ignored distractors. Future research in humans using SSVEP could further explore this issue by using paradigms in which a clear baseline SSVEP response was maintained throughout a block of trials, thereby providing a benchmark for evaluating relative degrees of enhancement and suppression (see e.g. Andersen and Müller, 2010 for an example of this approach).

Since the difference between target- and distractor-evoked SSVEP power was observed only when the data were locked to behavioral responses, it is possible that this difference simply reflects a broadband increase in power related to the preparation of a motor response. However, this is unlikely because we obtained sharply tuned SSVEP signals peaking only at the stimulus frequencies in both E1 and E2 (Figure 3). Furthermore, there are several other factors that argue against a broadband frequency modulation. First, if the effect reflects a broadband increase in power related to response preparation, there should have been an additive effect for both target- and distractor-evoked SSVEP responses. Instead, we observed a significant divergence of target- and distractor-evoked SSVEP responses, which rules out a broadband source of our main effect of interest. Second, the observed divergence in power on correct and incorrect trials was consistent across the two experiments even though stimulus frequencies ranged widely from 6.9Hz to 25Hz. Thus the effect is unlikely to be specific to any intrinsic oscillation related
Abbreviate title: Temporal Dynamics of divided spatial attention

to response preparation.

Among past psychophysical studies that reported evidence in support of divided attention, some assumed that attention can be split into multiple spotlights only in a rapid and discrete manner and thus limited their cue-to-stimulus intervals (CSIs) to below 200ms (e.g., Baldauf and Deubel, 2008; Bichot et al., 1999; Kramer & Hahn, 1995). On the other hand, others have assumed that the division of attention can be sustained over a longer period of time and thus employed longer CSIs (~500ms to 10s, e.g., Awh and Pashler, 2000; Gobell et al., 2004; Kraft et al., 2005). To more directly address this issue, Dubois et al. (2009) conducted a psychophysical experiment in which CSIs were varied and found that divided attention could occur only over very short timescales (CSIs < ~200ms). In contrast, Müller et al. (2003a) observed divided attention effects on SSVEP responses averaged over the entire stimulus interval (~2-3s) and thus proposed that divided spatial attention could be sustained over several seconds. However, an effect that is based on the averaged SSVEP power over a long temporal window could be driven either by a truly sustained process, or by a temporally discrete process that has a large impact on the overall average power of the SSVEP response. In the present study, we examined changes of target- and distractor-evoked SSVEP power across time to characterize the temporal properties of divided attention, and the data support a temporally discrete splitting process that occurs immediately preceding correct responses and is sustained for only several hundred milliseconds (which is broadly consistent with the temporal estimates based on the behavioral data reported by Dubois et al. 2009). Nevertheless, our conclusion that attention can be divided in a temporally discrete manner must be qualified as our stimuli were close together and always presented in the same visual quadrant. It is possible, for example, that sustained divided attention might be more feasible when the targets are in different hemifields (Maertens and Pollmann, 2005; Pollmann, Zaidel, and von Cramon, 2003; Sereno and Kosslyn, 1991), and it is also possible that cross-hemispheric division might operate on a more graded timescale due to anatomical constraints governing cross-hemispheric communication (e.g., McMains and Somers, 2004, 2005; Müller et al., 2003). Thus, future studies will be needed to further explore the relationship between the time-course of divided attention and the spatial properties of the visual display.

Finally, given that the focus of spatial attention can be alternated rapidly, it is possible that subjects were switching between the two target stimuli as opposed to continuously dividing attention. In the present studies, we observed a divergence in SSVEP power right before a correct response across a relatively brief temporal window that lasted for approximately 200ms. While estimates of switching-time vary considerably across studies, if the lower limit of 100ms proposed by Jans et al. (2010) is adopted, then subjects in our task may in fact have been switching attention rapidly between the two target stimuli (and this would also pose a challenge for many previous studies as well, e.g. Baldauf et al., 2006; Gobell et al., 2004; McMains and Somers, 2004; 2005; Müller et al, 2003a). However, the lower bound of 100ms is based on studies that could not rule out parallel encoding accounts, so this may be an underestimate of the actual switch time (e.g., Czerwinski, Lightfoot, and Shiffrin, 1992; Kramer and Hahn, 1995). On the other hand, if a more conservative estimate of switching time in the range of 200-250ms (or more) is adopted (e.g., Duncan, Ward, and Shapiro, 1994; Kröse and Julesz, 1989; Weichselgartner and Sperling, 1987), then it is less likely that a rapid-switching account can explain the present data or the results reported by other labs. That said, the temporal scale at which attention can be switched is still an issue of debate, and it is difficult to conclusively rule out all forms of this account. However, even in light of this ambiguity, the present data provide a high temporal resolution metric that we think plausibly reflects the division of attention immediately preceding a successful perceptual discrimination.

In summary, our data suggest that dividing attention across non-contiguous locations is possible even when the stimuli are in close proximity and located within the same visual quadrant. Moreover, the
multifocal division of attention within a quadrant onsets in a discrete manner and is primarily time-locked to the onset of a correct behavioral response. Thus, these data suggest a tight temporal relationship between the transient division of spatial attention and successful perceptual decision making.
Abbreviate title: Temporal Dynamics of divided spatial attention

**Figure legends**

**Figure 1.** Depiction of the stimulus display. On a given trial, subjects maintained central fixation and viewed three stimuli presented in either the lower left or the lower right quadrant. Their task was to make a discrimination on the upper and the lower stimuli (either R1 and R3 or L1 and L3, see text for details), while ignoring the intervening distractor (either R2 or L2). A trial in which the stimuli were presented in the lower right quadrant is shown, and the dashed lines surrounding each potential stimulus location were not visible in the actual experiment.

**Figure 2.** Posterior-occipital view of topographical maps depicting SSVEP power averaged across the 3.2-sec stimulus duration (z-scores; the first row of each panel) and EOI probability maps (the second row of each panel; color bars indicate the probability with which each electrode was included in an EOI across subjects), corresponding to each of the possible target (blue dotted frame) and distractor locations (red dotted frame) in E1 (N=11; top panel) and E2 (N = 14; bottom panel). Note that in E1, the two targets flickered at the same frequency so only two topographical maps were produced, one for trials in the lower left quadrant and one for trials in the lower right quadrant. In contrast, the two targets in E2 flickered at different frequencies, yielding six distinct topographical maps corresponding the lower left and lower right targets. Also, note that SSVEP power associated with left target was most prominent over right occipital/parietal cortex, and SSVEP power associated with right targets was most pronounced over left visual cortex. This pattern is expected based on the contralateral mapping of external visual space to internals cortical representations.

**Figure 3.** The signal-to-noise ratio (SNR) at different frequencies, computed across the entire trial. The sharp peaks at each stimulation frequency demonstrate the specificity of the SSVEPs at 6.9Hz and 10.4Hz in E1 and at 15.0Hz, 18.7Hz, and 25.0Hz in E2. SNR was calculated by dividing power at central frequency by the power in four surrounding frequency bins.

**Figure 4.** Normalized power of target- and distractor-evoked SSVEPs on correct and incorrect trials in E1, either time-locked to the stimulus onset (top panel) or to the onset of the behavioral response (bottom panels). Note that in the bottom panels, responses on correct and incorrect trials are plotted separately for clarity. In all panels, the zero point on the x-axis represents the time of either stimulus onset or response onset for stimulus-locked and response-locked data, respectively. Red horizontal bar in lower left panel indicates the temporal window over which the target-evoked SSVEP power was significantly greater than the distractor evoked SSVEP power on correct trials (p<0.05 FDR corrected). Shaded areas in all plots represent ± 1 S.E.M. across subjects.

**Figure 5.** Stimulus-locked (left column) and response-locked (right column) normalized power of target- and distractor-evoked SSVEPs in E1, sorted into trials with short (top row) and long RTs (bottom row). The small top panel above each stimulus-locked plot shows the RT histograms (data from all subjects), with the mean trial count on the y-axis. In all panels, the zero point on the x-axis represents the time of either stimulus onset or response onset for stimulus-locked and response-locked data, respectively. Red horizontal bar in the response-locked plot indicates the temporal window over which the target-evoked SSVEP power was significantly greater than the distractor evoked SSVEP power on correct trials (p<0.05 FDR corrected). Shaded areas in all plots represent ± 1 S.E.M. across subjects.
Figure 6. Normalized power of target- and distractor-evoked SSVEPs on correct and incorrect trials in E2, either time-locked to the stimulus onset (top panel) or to the onset of the behavioral response (bottom panels). Note that in the bottom panels, responses on correct and incorrect trials are plotted separately for clarity. In all panels, the zero point on the x-axis represents the time of either stimulus onset or response onset for stimulus-locked and response-locked data, respectively. Red horizontal bar in lower left panel indicates the temporal window over which the target-evoked SSVEP power was significantly greater than the distractor evoked SSVEP power on correct trials (p<0.05 FDR corrected). Shaded areas in all plots represent ± 1 S.E.M. across subjects.

Figure 7. Stimulus-locked (left column) and response-locked (right column) normalized power of target- and distractor-evoked SSVEPs in E2, sorted into trials with short (top row) and long RTs (bottom row). The small top panel above each stimulus-locked plot shows the RT histograms (data from all subjects), with the mean trial count on the y-axis. In all panels, the zero point on the x-axis represents the time of either stimulus onset or response onset for stimulus-locked and response-locked data, respectively. Red horizontal bar in the response-locked plot indicates the temporal window over which the target-evoked SSVEP power was significantly greater than the distractor evoked SSVEP power on correct trials (p<0.05 FDR corrected). Shaded areas in all plots represent ± 1 S.E.M. across subjects.
Abbreviate title: Temporal Dynamics of divided spatial attention

References


Abbreviate title: Temporal Dynamics of divided spatial attention


Abbreviate title: Temporal Dynamics of divided spatial attention


18


zoom in
Experiment 1 (E1)

- **target-evoked SSVEP** (correct trials)
- **distractor-evoked SSVEP** (correct trials)
- **target-evoked SSVEP** (incorrect trials)
- **distractor-evoked SSVEP** (incorrect trials)
1. Stimulus-locked Response-locked

**Short RTs**

- **Normalized power**
  - RTs
- **Stimulus-locked**
  - Target-evoked SSVEP (correct trials)
  - Distractor-evoked SSVEP (correct trials)
- **Response-locked**
  - Target-evoked SSVEP (incorrect trials)
  - Distractor-evoked SSVEP (incorrect trials)

**Long RTs**

- **Normalized power**
  - RTs
  - n.s.

**Time (sec)**

- 0 1 2 3 4 5

**Normalized power**

- 0 0.5 1.0 1.5 2.0

**Stimulus-locked**

- Target-evoked SSVEP (correct trials)
- Distractor-evoked SSVEP (correct trials)

**Response-locked**

- Target-evoked SSVEP (incorrect trials)
- Distractor-evoked SSVEP (incorrect trials)

**n.s.**
Experiment 2 (E2)

- **target-evoked SSVEP (correct trials)**
- **distractor-evoked SSVEP (correct trials)**
- **target-evoked SSVEP (incorrect trials)**
- **distractor-evoked SSVEP (incorrect trials)**

Stimulus-locked normalized power

Response-locked normalized power
Stimulus-locked

Short RTs

Response-locked

Long RTs

n.s.

*