

# Sustained Multifocal Attentional Enhancement of Stimulus Processing in Early Visual Areas Predicts Tracking Performance

Viola S. Störmer,<sup>1,2</sup> Gesche N. Winther,<sup>2,3</sup> Shu-Chen Li,<sup>2,4</sup> and Søren K. Andersen<sup>5</sup>

<sup>1</sup>Harvard University, Cambridge, Massachusetts 02138, <sup>2</sup>Max Planck Institute for Human Development, 14195 Berlin, Germany, <sup>3</sup>Free University of Berlin, 14195 Berlin, Germany, <sup>4</sup>TU Dresden, 01062 Dresden, Germany, and <sup>5</sup>University of California at San Diego, San Diego, California 92093

Keeping track of multiple moving objects is an essential ability of visual perception. However, the mechanisms underlying this ability are not well understood. We instructed human observers to track five or seven independent randomly moving target objects amid identical nontargets and recorded steady-state visual evoked potentials (SSVEPs) elicited by these stimuli. Visual processing of moving targets, as assessed by SSVEP amplitudes, was continuously facilitated relative to the processing of identical but irrelevant nontargets. The cortical sources of this enhancement were located to areas including early visual cortex V1–V3 and motion-sensitive area MT, suggesting that the sustained multifocal attentional enhancement during multiple object tracking already operates at hierarchically early stages of visual processing. Consistent with this interpretation, the magnitude of attentional facilitation during tracking in a single trial predicted the speed of target identification at the end of the trial. Together, these findings demonstrate that attention can flexibly and dynamically facilitate the processing of multiple independent object locations in early visual areas and thereby allow for tracking of these objects.

## Introduction

Adaptive behavior in many situations, such as driving a car or watching team sports, requires the ability to continuously monitor multiple independently moving objects at different locations in the visual field. Previous behavioral multiple object-tracking studies (Pylyshyn and Storm, 1988; for a review, see Cavanagh and Alvarez, 2005) have precisely quantified this ability; however, the neural mechanisms underlying multiple object tracking are still much debated. A likely explanation is that tracking of multiple objects is afforded by parallel attentional selection of these objects. However, the flexibility required to select four or more moving objects in parallel exceeds the hitherto demonstrated capacity of selective attentional enhancement of visual stimulus processing. Classically, selective attention has been thought to operate on a single location (Posner, 1980; LaBerge, 1995), a view that has only more recently been challenged by demonstrations of concurrent selection of two noncontiguous locations (Müller et al., 2003; McMains and Somers, 2004). However, to date, there is no direct evidence showing that multiple objects at separate dy-

namically changing locations can be selected by demonstrating enhanced visual processing of these objects.

Recent functional magnetic resonance imaging (fMRI) studies of attentional tracking showed the involvement of parietal and frontal brain areas, but not early visual areas, during tracking (Culham et al., 1998, 2001; Jovicich et al., 2001; Howe et al., 2009). This suggests that multiple target selection occurs at later processing stages and not by early enhancement of visual stimulus processing. However, these studies were not able to separately assess the processing of targets and nontargets and hence might have missed activities related to the differential processing of these objects.

Here we directly compared selective processing of targets and nontargets during a tracking task. Participants were asked to track either five or seven objects among identical nontarget objects. Visual processing of targets and nontargets was assessed by measuring steady-state visual evoked potentials (SSVEPs) elicited by these stimuli. The SSVEP is the oscillatory potential field generated by the visual cortex in response to flickering stimuli. It has the same fundamental frequency as the driving stimulus (Regan, 1989), and its amplitude can be enhanced by attention (Morgan et al., 1996; for a recent review, see Andersen et al., 2011b). Targets and nontargets flickered at different frequencies during the tracking period (8.5 vs 9.4 Hz), thereby allowing us to examine the SSVEP to each stimulus type separately. If processing of multiple targets is indeed enhanced continuously in the visual cortex, this should lead to larger SSVEP amplitudes for targets. Furthermore, if limited attentional resources must be shared among tracked target stimuli, one would expect the relative magnitude of the attentional enhancement of SSVEP amplitudes to

Received Aug. 21, 2012; revised Jan. 31, 2013; accepted Feb. 8, 2013.

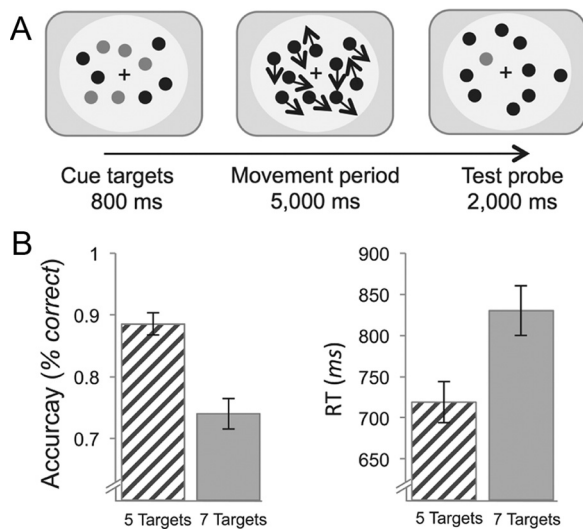
Author contributions: V.S.S., G.N.W., S.-C.L., and S.K.A. designed research; V.S.S. and G.N.W. performed research; V.S.S., G.N.W., and S.K.A. analyzed data; V.S.S., G.N.W., S.-C.L., and S.K.A. wrote the paper.

V.S.S. has been supported by the International Max Planck Research School: The Life Course: Evolutionary and Ontogenetic Dynamics. S.K.A. was supported by Deutsche Forschungsgemeinschaft Grant AN 841/1-1. We thank Steven A. Hillyard for helpful discussions on using SSVEPs to study multiple object tracking.

Correspondence should be addressed to Viola Störmer, Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge, MA 02138. E-mail: vstormer@fas.harvard.edu.

DOI:10.1523/JNEUROSCI.4015-12.2013

Copyright © 2013 the authors 0270-6474/13/335346-06\$15.00/0



**Figure 1.** Experimental procedure and behavioral results. **A**, Illustration of a trial with five target and five nontarget disks. Target disks and test probe at the end of the trial are denoted in light gray here, and the movement of the disks is indicated by the arrows. During the movement period, target and nontarget disks flickered at separate frequencies (for details, see Materials and Methods). **B**, Accuracy rates (percentage correct) and correct RTs show a clear effect of set size, i.e., performance was lower when seven targets were tracked. Error bars denote 1 SEM.

decrease with an increasing number of to-be-tracked targets. Finally, if enhancement of visual processing is functionally relevant for behavioral tracking performance, the magnitude of this enhancement should predict tracking performance.

## Materials and Methods

### Participants

Eighteen healthy right-handed volunteers participated in the study. All participants had normal or corrected-to-normal visual acuity and normal color vision, gave written informed consent, and received monetary compensation for their participation (€17). Data from two subjects were excluded from the analysis because they reported noticing the different flicker frequencies for targets and nontargets (for details, see below, Stimuli and procedure); data from one other subject were excluded because >40% of the trials were rejected as a result of artifacts in the EEG. Thus, the final sample contained data from 15 young adults (eight female; 23–31 years of age; mean  $\pm$  SD age, 26.7  $\pm$  2.2 years). All participants were part of a larger study and had participated in behavioral assessments (Störmer et al., 2011) and one EEG assessment (Störmer et al., 2013) of a different multiple object-tracking task before. The ethics committee of the Max Planck Institute for Human Development approved the study.

### Stimuli and procedure

The study took place in a sound-attenuated and electrically shielded chamber that contained a 19-inch cathode ray tube monitor (1024  $\times$  768 pixels; refresh rate: 85 Hz). Participants were seated 70 cm in front of the display. The background color of the screen was gray (10.5 cd/m<sup>2</sup>). The stimuli were presented in a circular white viewing field (diameter: 15.03°; 124 cd/m<sup>2</sup>) in the center of the screen. The chamber was dark throughout the experiment, and the participants were instructed to maintain their gaze at the fixation cross (0.45°  $\times$  0.45°) in the center of the screen throughout each trial.

The experiment consisted of 14 blocks of 32 trials each. At the beginning of a trial, a number of black disks (0.9° diameter) were presented at random locations within the viewing area. On half of the trials, 10 disks appeared (5-target condition), and on the other half of the trials, 14 disks appeared (7-target condition). After 800 ms, half of the disks were briefly marked in red (30 cd/m<sup>2</sup>), designating them as targets. The target disks turned back to black after 900 ms, and all disks started moving randomly in linear trajectories across the viewing area (Fig. 1A). The disks moved at a constant speed of  $\sim$ 3.2°/s and changed trajectory only when they made

contact with the outer barrier of the viewing area or with each other (no occlusion). The movement of the disks stopped after 5 s, and one of the disks turned red, marking it as a test probe. On half of the trials, one of the target disks was probed, and on the remaining half of the trials, one of the nontarget disks was probed. Participants were instructed to memorize the target disks that were highlighted at the beginning of each trial and keep track of them throughout the movement period. At the end of the trial, participants had to indicate whether the test probe was one of the targets or not by pressing a left or right button on the keyboard (left and right Ctrl key) within a 2-s response window. Response buttons were counterbalanced between participants. In 75% of all trials, targets and nontargets flickered at different frequencies during the tracking period. These different-frequency trials allowed the clear separation of target and nontarget processing in the SSVEP responses. On half of the different-frequency trials (12 trials per block), targets flickered at  $\sim$ 8.5 Hz (five frames on, five frames off), and nontargets flickered at  $\sim$ 9.4 Hz (five frames on, four frames off). This assignment was reversed for the remaining half of the different-frequency trials. On the remaining 25% of all trials, both targets and nontargets flickered at the same frequency. These same-frequency trials served as a behavioral control for potential frequency effects on performance and were not part of the SSVEP analysis. On half of the same-frequency trials (four trials per block) all disks flickered at  $\sim$ 8.5 Hz, and on the remaining half of the trials, all disks flickered at  $\sim$ 9.4 Hz. All trial types were randomly intermixed within each block. After the experiment, tracking strategies and the perception of the flicker were assessed with a standardized questionnaire with open questions.

### Data analyses

**Behavioral data analyses.** Accuracy rates (percentage correct) and correct reaction times (RTs) were analyzed separately using repeated-measures ANOVA, with factors set size (five vs seven), frequency condition (same vs different), and target frequency (8.5 vs 9.4 Hz).

**Electrophysiological recordings and analyses.** Brain electrical activity was recorded from 64 Ag/AgCl electrodes arranged according to the 10–10 system in an elastic scalp cap (BrainAmp DC amplifiers; Brain Products). The right mastoid served as a recording reference. The horizontal electrooculogram was recorded bipolarly using two electrodes that were positioned lateral to the external canthi, and the vertical electrooculogram was measured through one electrode below the left eye and FP1 (above the left eye). Electrode impedances were kept below 5 k $\Omega$ . EEG data were filtered with a bandpass of 0.1–100 Hz and digitized with a sampling rate of 500 Hz. Signal processing for the SSVEP analysis was performed with MATLAB (MathWorks) using the EEGLAB toolbox (Delorme and Makeig, 2004) and custom-written scripts. The main SSVEP analysis was performed only for trials with correct responses. Epochs containing horizontal eye movements, blinks, and muscle artifacts were rejected by visual inspection. In addition, myographic artifacts at single channels and noisy channels were interpolated using spherical splines. Artifact-free data were re-referenced to the average reference. The averaging epochs extended from 500 to 4800 ms after movement onset. SSVEP amplitudes were quantified as the absolute of the complex Fourier coefficients for each frequency, participant, and condition separately. The resulting SSVEP amplitudes from three occipital electrodes (O1, Oz, O2) exhibiting the overall highest SSVEP amplitudes were analyzed by repeated-measures ANOVAs with factors disk type (targets vs nontargets) and set size (five targets vs seven targets). We then normalized the SSVEP signal for each set size condition and frequency by dividing the amplitude by the mean amplitude of targets and nontargets for that particular frequency (Andersen et al., 2011a). The normalized amplitudes were collapsed over frequencies and analyzed by two-way repeated-measures ANOVAs with factors disk type (targets vs nontargets) and set size (five targets vs seven targets).

Multiple object-tracking performance may be capacity limited in a way that a total selective attention capacity is shared between all attended stimuli (Pylyshyn and Storm, 1988; Oksama and Hyönä, 2004). If this were the case, attending to more targets should reduce the SSVEP attention effect, which here reflects the overall differences between processing of all target and nontarget stimuli. This assumption is supported by studies that found reduced SSVEP amplitudes to attended stimuli when

two stimuli were attended compared with when only one stimulus was attended (Andersen et al., 2009; Toffanin et al., 2009). The potential of the SSVEP to sensitively reflect the allocation of attention in cases with higher numbers of stimuli has been demonstrated in studies showing differential effects of feature-based and spatial attention on as many as four concurrently presented stimuli (Andersen et al., 2008, 2011a).

In the present experiment, an attenuation effect would occur regardless of whether attention was distributed concurrently to all targets, in which case each target would receive less attention with more targets, or whether attention switched between targets, in which case the amount of time each target could be attended would decrease as target numbers increases. To directly assess whether our data are consistent with a limited capacity account, we devised a test as follows: assuming strict capacity limitation, the magnitude of the attention effect  $E$  (attended – unattended) is equal to the capacity  $C$  divided by the number ( $n$ ) of targets (i.e., the set size):

$$E_n = C/n.$$

Hence, attentional capacity  $C$  can be quantified independently for each set size by multiplying the magnitude of the attention effect  $E_n$  with the corresponding set size  $n$ . We calculated the attentional capacities  $C_5$  and  $C_7$  and compared the two values by means of a paired  $t$  test. Under the assumption of a strict capacity limit, both set sizes should yield comparable estimates for  $C$ .

**Topographical mapping and source analyses.** To characterize the scalp topography of SSVEP amplitudes, isopotential contour maps were created using ERPSS (University of California, San Diego). Topographical voltage maps were plotted for the time intervals that were analyzed above and were created separately for the different set size conditions (five vs seven targets) and disk types (targets vs nontargets). To gain information on the cortical sources giving rise to the scalp-measured SSVEP attention effects, the complex Fourier coefficients of the difference of SSVEPs elicited by targets and nontargets, collapsed over set size, were subjected to variable resolution electromagnetic tomography (VARETA; Bosch-Bayard et al., 2001) for each frequency separately. The resulting three-dimensional complex vector, representing the direction, phase, and amplitude for each voxel, was statistically tested against zero by means of Hotelling's  $t^2$  test using Bonferroni's correction for multiple comparisons.

**Single-trial analyses.** All trials (correct and incorrect) were included in the single-trial analysis. SSVEP amplitudes were computed separately for each frequency, participant, and trial. The resulting amplitude values were normalized separately for each subject, condition, and frequency by dividing the amplitudes for each trial by the absolute amplitude of the mean of all trials. Next, for each subject, we computed the differences of the normalized target-minus-nontarget amplitudes separately on each trial. This yielded the magnitude of the attention effect (target-minus-nontarget difference) for each trial and subject. For each participant, the resulting amplitude differences were sorted from small to large separately for each set size condition. The sorted trials were then split in half depending on the magnitude of the attention effect (i.e., trials with smaller attention effect vs trials with larger attention effect, separately for each participant). Finally, for each participant and set size condition, mean accuracy rates and mean RTs were calculated for each dataset, resulting in an accuracy and RT value for each participant for trials with large and small attentional modulation, respectively. An ANOVA with factors amplitude difference (small vs large) and set size (5 vs 7) was conducted.

## Results

As expected, behavioral performance was lower when participants were asked to track seven compared with five targets (Fig. 1B). Tracking accuracy was reduced in the seven-target condition ( $74.0 \pm 0.025\%$  correct, mean  $\pm$  SEM) compared with the five-target condition ( $88.5 \pm 0.018\%$  correct, mean  $\pm$  SEM;  $F_{(1,14)} = 152.54$ ,  $p < 0.0001$ ,  $\eta^2 = 0.623$ ), and RTs were slower for the seven-target condition ( $830 \pm 30$  ms, mean  $\pm$  SEM) relative to the five-target condition ( $719 \pm 25$  ms, mean  $\pm$  SEM;  $F_{(1,14)} =$

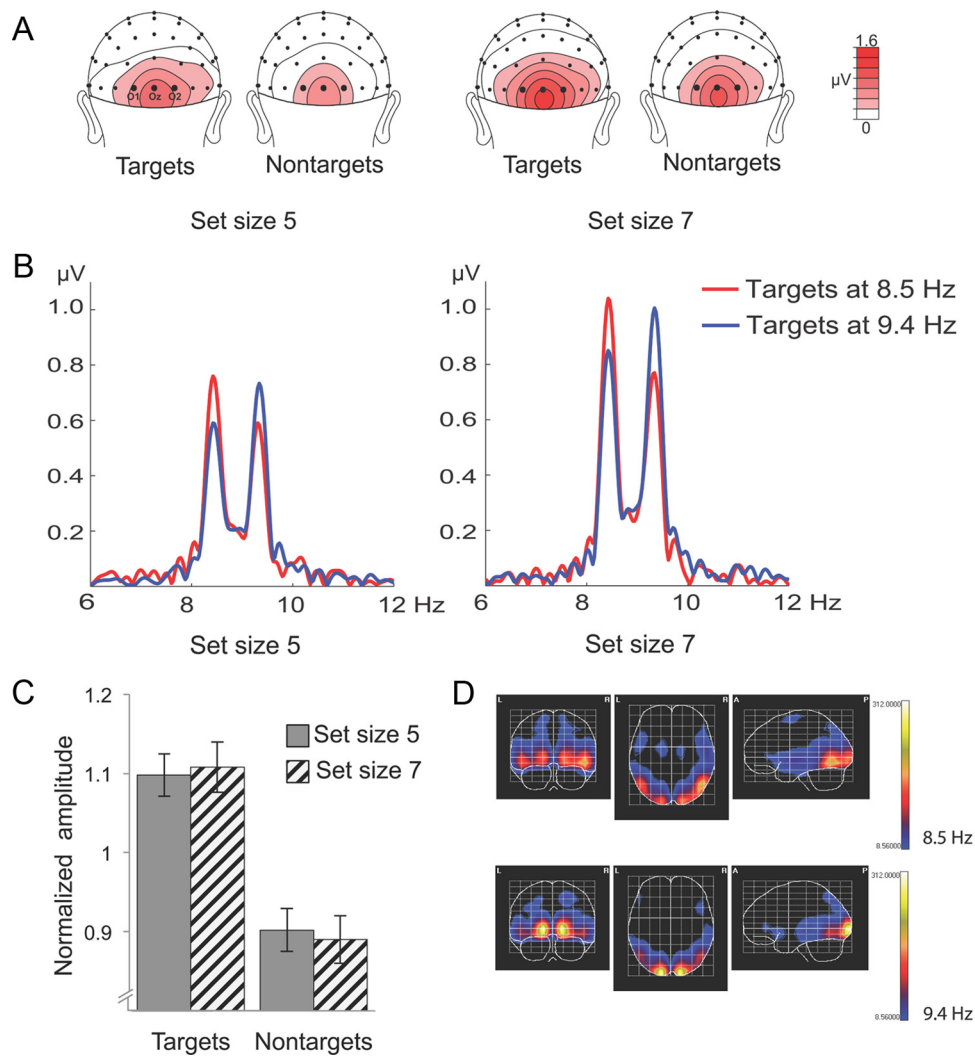
$59.87$ ,  $p < 0.0001$ ,  $\eta^2 = 0.570$ ). To examine whether tracking performance benefited from the different flicker frequencies of target and nontarget objects, both types of stimuli flickered at the same frequency on 25% of the trials. Performance did not differ between these frequency conditions (same vs different; accuracy,  $F_{(1,14)} = 1.01$ ,  $p = 0.312$ ,  $\eta^2 = 0.009$ ; RT,  $F_{(1,14)} = 2.09$ ,  $p = 0.17$ ,  $\eta^2 = 0.008$ ) or for trials in which targets were presented at 8.5 versus 9.4 Hz (accuracy,  $F_{(1,14)} = 0.06$ ,  $p = 0.814$ ,  $\eta^2 = 0.0001$ ; RT,  $F_{(1,14)} = 2.28$ ,  $p = 0.153$ ,  $\eta^2 = 0.017$ ).

The amplitude spectrum obtained by Fourier transformation when observers tracked five (set size five) or seven targets (set size seven) shows clear amplitude peaks at the frequencies of the flickering stimuli (8.5 and 9.4 Hz; Fig. 2B). Target processing was facilitated during tracking as indicated by larger SSVEP amplitudes elicited by targets compared with nontargets ( $F_{(1,14)} = 12.86$ ,  $p = 0.003$ ,  $\eta^2 = 0.231$ ). Additionally, SSVEP amplitudes were larger for the seven-target condition compared with the five-target condition ( $F_{(1,14)} = 59.45$ ,  $p < 0.0001$ ,  $\eta^2 = 0.372$ ). To directly compare attentional modulation between the five-target and seven-target conditions, SSVEP amplitudes were normalized separately for each frequency and set size condition and then collapsed over frequencies (Andersen et al., 2011a). The resulting normalized amplitudes were enhanced for targets compared with nontargets ( $F_{(1,14)} = 14.69$ ,  $p = 0.002$ ,  $\eta^2 = 0.453$ ). Surprisingly, the magnitude of this effect did not depend on the number of tracked targets (Fig. 2C; set size  $\times$  attention,  $F_{(1,14)} = 0.13$ ,  $p = 0.71$ ,  $\eta^2 = 0.0007$ ), indicating that on average each target was equally enhanced regardless of set size. Strict accounts of limited capacity would have assumed that the relative attentional enhancement of all targets, as assessed by the normalized SSVEPs, should decrease with increasing set size. To rule out that our failure to observe such an effect was not merely attributable to lack of statistical power, we devised a second test in which the null hypothesis was derived from the assumption of strict capacity limitation (see equation above). The magnitude of SSVEP attention effects multiplied by set size was significantly larger for the condition with seven targets relative to the condition with five targets ( $t_{(14)} = 2.15$ ,  $p = 0.004$ ), indicating that a hypothesis of strict capacity limitation cannot account for the magnitudes of SSVEP attention effects in our data.

The topographical voltage maps of SSVEP amplitudes show focal maxima around central occipital electrodes in all conditions (Fig. 2A), suggesting that these signals arose from a common neural generator in the visual cortex. The locations of the neural sources of the difference between SSVEPs elicited by targets and nontargets were estimated by means of VARETA (Bosch-Bayard et al., 2001). Two separate peaks of modulation were observed in each hemisphere for both frequencies (Fig. 2D). The Montreal Neurological Institute coordinates of these peaks were as follows: for 8.5 Hz, ( $-14$ ,  $-98$ ,  $-2$ ) and ( $-50$ ,  $-62$ ,  $-10$ ) for left and ( $21$ ,  $-91$ ,  $-10$ ) and ( $50$ ,  $-62$ ,  $-10$ ) for right; and for 9.4 Hz, ( $-14$ ,  $-98$ ,  $-2$ ) and ( $-50$ ,  $-62$ ,  $-10$ ) for left and ( $14$ ,  $-98$ ,  $-2$ ) and ( $50$ ,  $-69$ ,  $-10$ ) for right. Thus, in good agreement with previous findings (Di Russo et al., 2007; Andersen et al., 2008), peak modulation of SSVEP amplitudes was observed in regions containing early visual areas V1–V3 as well as motion-sensitive area MT.

If tracking performance depends on sustained attentional modulation of early visual processing, we would expect that attentional modulation of SSVEP amplitudes in a single trial predicts behavioral performance at the end of this trial. To test for this, attentional modulation was calculated as the difference of normalized amplitudes between targets and nontargets on each trial. As depicted in Figure 3, trials in which participants showed





**Figure 2.** Topographical distribution of the SSVEP response, SSVEP signal in the power spectrum, normalized SSVEP amplitudes for the two set size conditions, and estimated neural sources of the SSVEP attention effect. **A**, Spline-interpolated voltage maps of the SSVEP amplitudes for targets and nontargets for the two set size conditions, collapsed across both frequencies. Peak amplitudes are found at occipital sites (O1, Oz, O2). Maps display back views of the scalp. **B**, Grand-average amplitude spectrum obtained by fast Fourier analysis of the SSVEP waveforms for targets and nontargets at the different frequency rates and for the five-target condition (left) and the seven-target condition (right), collapsed across three occipital electrodes (O1, Oz, O2). Two clear peaks at the stimulation frequencies (8.5 and 9.4 Hz) are visible. **C**, Normalized amplitudes for targets and nontargets for the five-target (solid gray) and seven-target (dashed black) conditions show equal attentional enhancement of targets versus nontargets. **D**, Statistical parametric maps of cortical current-density distributions giving rise to the SSVEP amplitude difference between targets and nontargets for both frequencies separately. The scale represents  $t^2$  values, and the threshold of 8.56 corresponds to  $p < 0.001$  corrected for multiple comparisons.

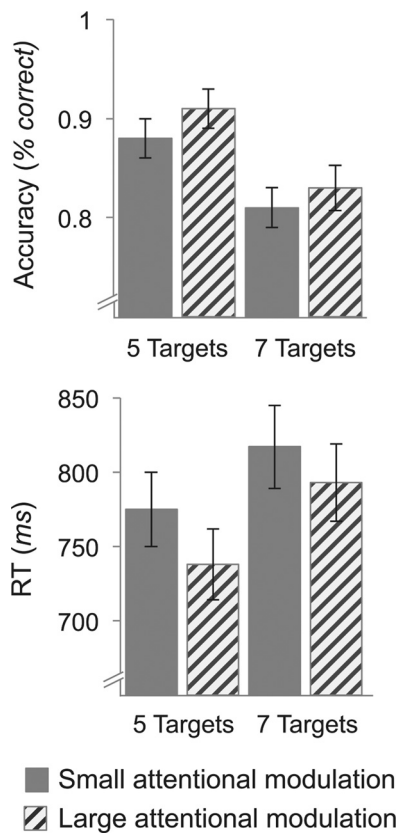
a stronger attentional modulation yielded faster RTs ( $F_{(1,14)} = 21.58, p = 0.0004, \eta^2 = 0.192$ ), although the magnitude of attentional modulation was not reliably related to tracking accuracy ( $F_{(1,14)} = 1.88, p = 0.192, \eta^2 = 0.022$ ). Attentional facilitation was measured during the tracking period itself and hence preceded the behavioral responses. Therefore, sustained attentional facilitation in visual processing pathways predicted RT during tracking.

## Discussion

The present study demonstrates that keeping track of multiple moving objects enhances visual processing of these objects continuously. By concurrently tagging moving targets and nontargets with different frequencies, we separately quantified visual processing of simultaneously presented stimuli while participants tracked a subset of them. SSVEP amplitudes elicited by moving targets were larger than SSVEP amplitudes elicited by

identical moving nontargets. This amplitude modulation was localized to early and mid-levels of processing in visual cortex. This demonstrates that inputs from multiple separate locations are selected and continuously facilitated beginning at an early sensory processing level.

Observers responded faster on those trials that showed larger attentional modulation of SSVEP amplitudes (Fig. 3). This was the case for both set size conditions, pointing to a reliable relationship between neural enhancement and RT. Accuracy rates tended to be higher for those trials that showed larger attentional modulations (Fig. 3); however, there was no reliable relationship between SSVEP amplitude modulations and accuracy, most likely because accuracy was a less sensitive performance measure in the present task. Whereas accuracy is a dichotomous variable (correct vs incorrect), RTs are continuous and thus provide a more sensitive measure, particularly on the single-trial level (Jensen, 2006). The association be-



**Figure 3.** Mean accuracy rates (percentage correct; top) and correct RTs (bottom) for trials with small or large SSVEP amplitude modulations for the two set size conditions. Trials with larger attentional modulation of SSVEPs showed better performance. Error bars denote SEM.

tween SSVEP amplitudes and RT implies that selective processing in early visual areas critically determines tracking performance.

Overall, SSVEP amplitudes were larger for the seven-target compared with the five-target condition (Fig. 2*B*) because the magnitude of SSVEP amplitudes depends on the size of the flickering stimulus, such that more (or larger) stimuli elicit larger SSVEP responses (Regan, 1989). However, the observed relative attention effect was similar in magnitude for the five-target and seven-target conditions (Fig. 2*C*). It should be noted that the main SSVEP analysis only included trials in which targets were tracked correctly (see Materials and Methods). So although more tracking errors were made in the seven-target condition, the magnitude of facilitation was comparable for the five-target and seven-target conditions for trials with correct tracking performance. This result appears to be inconsistent with accounts that assume strict capacity limitations such as serial switching (Howard and Holcombe, 2008; Holcombe et al., 2011). Assuming serial switching, the total time each target is attended, and hence the magnitude of attention effects, would be inversely proportional to the number of targets. An additional test confirmed that SSVEP amplitudes in our experiment significantly deviated from such a pattern. Note that other studies found reduced attentional enhancement of SSVEP amplitudes when attention was divided between stimuli (Andersen et al., 2009; Toffanin et al., 2009). Therefore, the present results suggest that an equal amount of attentional resources was allocated to each of the tracked targets, regardless of the total number of targets (five or seven).

The magnitude of the attention effects in the present study is comparable with the magnitude of attention effects in unifocal

spatial attention tasks: on average, modulations of SSVEP responses were enhanced by 37.3% (Keitel et al., 2012), 15.6% (Andersen et al., 2011a), and 37.9% (Quigley et al., 2012) when one location was attended. Tasks and stimulus displays varied considerably between these studies, but in no case was the reported attention effect even close to being five or seven times larger than in the present study (five targets: 21.7%; seven targets: 24.5%), as would have been consistent with a serial switching account.

The nature of the mechanism that allows parallel selection of multiple moving targets remains to be determined. A likely explanation appears to be that observers group targets into one object (e.g., a virtual polygon) and constantly update this internal representation as the targets move (Yantis, 1992). Accordingly, attention would need to keep track of one object representation only, regardless of the number of targets. This would be consistent with the amount of attentional modulation in our experiment being equal for both set sizes and comparable with previous experiments using unifocal attention tasks. According to this explanation, attentional modulation of visual target processing may be essential to maintain a representation of such a virtual polygon. This is similar to the proposal that visual working memory representations are actively maintained by selective modulations of early visual areas (Awh and Jonides, 2001). Tracking errors would then mainly result from failures to form or update this virtual object representation (Yantis, 1992). However, the increasing number of errors at higher set sizes must be attributable to other factors than a limitation of selective modulation of early visual processing, because this was found to be constant across set sizes.

The lack of set size by attention interaction certainly does not rule out the possibility that the number of tracked targets influences attention effects at other processing stages. Previous fMRI studies reported clear set size effects in parietal brain regions during tracking (Culham et al., 2001; Jovicich et al., 2001; Howe et al., 2009). Accordingly, multiple object tracking might follow a hybrid model in which attention operates in a capacity limited, possibly even serial manner, only at later processing stages, whereas early processing stages—indexed by the SSVEP—exhibit parallel facilitation of tracked targets. Recordings from neurons in the lateral intraparietal area of rhesus monkeys, an area implicated in the control of attention, have provided evidence for serial selection by suggesting that attention prioritizes only one location at the time (Bisley and Goldberg, 2003). The very short stimulus presentations used in that study leave open the intriguing possibility that attention very rapidly oscillates at shorter timescales, thereby becoming effectively divided at longer timescales.

We localized the attentional modulations recorded over occipital scalp sites to regions containing early visual areas V1–V3 and motion-sensitive area MT. Previous fMRI studies did not report modulations of the early visual areas (Culham et al., 1998, 2001; Jovicich et al., 2001; Howe et al., 2009). These studies were unable to separately assess the processing of targets and nontargets and thus are likely to have overlooked any differential attentional modulations of targets relative to nontargets. Here, by separating processing of targets and nontargets by means of SSVEP, we were able to show that selective processing of multiple targets occurs in early and mid-levels of the visual cortex.

Our results are in good agreement with recent electrophysiological data that showed transient attentional modulations of early visual processing during attentive tracking (Drew et al., 2009; Doran and Hoffman, 2010; Störmer et al., 2013). These

previous studies found that processing of salient probe flashes superimposed on the targets was enhanced from 100 to 175 ms after probe onset, providing important information on the time course of the attention effect. However, it remains questionable how well these results indexed processes of continuous multifocal attention for two reasons. First, the measurement was indirect in that the processing of the salient probe flashes, rather than the targets themselves, was assessed. Second, the salient flashes might have attracted exogenous attention and thereby introduced an unwanted influence on the results.

To record separable SSVEP signals from targets and nontargets, these stimuli flickered at different frequencies. Two nearby frequencies (8.5 and 9.4 Hz) were chosen to minimize the possibility that participants would use the differences in flicker frequencies to select targets from nontargets. Neither accuracy nor RTs differed between trials in which targets and nontargets were presented at the same or different frequencies. This speaks against the possibility that the different flicker frequencies might have mediated attentional selection. Such an explanation would also be inconsistent with previous SSVEP studies in which feature-selective attentional enhancement spread to stimuli presented at different frequencies when they shared features (e.g., color) with the attended stimulus (Andersen et al., 2008, 2011a). Note that the frequencies used in those studies (Müller et al., 2006; Andersen et al., 2008, 2011a) were farther apart than those in the present study.

We demonstrated the feasibility of SSVEPs in multiple object-tracking paradigms and its potential for gaining new insights about the mechanisms underlying the ability of tracking multiple moving objects. Our experimental design is limited in that we only examined two set sizes, both of which were toward the upper limit of most participants' tracking ability. Future investigations with a wider range of set sizes will be important for a more complete understanding of attentional modulations during tracking and will allow more direct comparisons with previous SSVEP studies of spatial attention to stationary objects.

The present findings suggest that selection operates similarly in static unifocal and dynamic multifocal attention situations. They indicate that the human brain selects multiple relevant objects by enhancing processing in visual areas concurrently, even when these objects move. Importantly, we found that attention modulated processing of sensory signals in early visual cortex, a brain region that has mostly been overlooked in previous neuroimaging studies of dynamic multifocal attention. The magnitude of this modulation of selective visual processing directly predicts tracking performance. Such an attention system that can flexibly select multiple independent locations dynamically and continuously in crowded visual scenes appears to be highly effective.

## References

- Andersen SK, Hillyard SA, Müller MM (2008) Attention facilitates multiple stimulus features in parallel in human visual cortex. *Curr Biol* 18:1006–1009. [CrossRef Medline](#)
- Andersen SK, Müller MM, Hillyard SA (2009) Color-selective attention need not be mediated by spatial attention. *J Vision* 9(6):2.1–7. [CrossRef Medline](#)
- Andersen SK, Fuchs S, Müller MM (2011a) Effects of feature-selective and spatial attention at different stages of visual processing. *J Cogn Neurosci* 23:238–246. [CrossRef Medline](#)
- Andersen SK, Müller MM, Hillyard SA (2011b) Tracking the allocation of attention in visual scenes with steady-state evoked potentials. In: *Cognitive neuroscience of attention* (Posner MI, ed), Ed 2. New York: Guilford.
- Awh E, Jonides J (2001) Overlapping mechanisms of attention and spatial working memory. *Trends Cogn Sci* 5:119–126. [CrossRef Medline](#)
- Bisley JW, Goldberg ME (2003) Neuronal activity in the lateral intraparietal area and spatial attention. *Science* 299:81–86. [CrossRef Medline](#)
- Bosch-Bayard J, Valdés-Sosa P, Virues-Alba E, Aubert-Vázquez E, John ER, Harmony T, Riera-Díaz J, Trujillo-Barreto N (2001) 3D statistical parametric mapping of EEG source spectra by means of variable resolution electromagnetic tomography (VARETA). *Clin Electroencephalogr* 32:47–61. [Medline](#)
- Cavanagh P, Alvarez GA (2005) Tracking multiple targets with multifocal attention. *Trends Cogn Sci* 9:349–354. [CrossRef Medline](#)
- Culham JC, Brandt SA, Cavanagh P, Kanwisher NG, Dale AM, Tootell RBH (1998) Cortical fMRI activation produced by attentive tracking of moving targets. *J Neurophysiol* 80:2657–2670. [Medline](#)
- Culham JC, Cavanagh P, Kanwisher NG (2001) Attention response functions: characterizing brain areas using fMRI activation during parametric variations of attentional load. *Neuron* 32:737–745. [CrossRef Medline](#)
- Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *J Neurosci Methods* 134:9–21. [CrossRef Medline](#)
- Di Russo F, Pitzalis S, Aprile T, Spitori G, Patria F, Stella A, Spinelli D, Hillyard SA (2007) Spatiotemporal analysis of the cortical sources of the steady-state visual evoked potential. *Hum Brain Mapp* 24:323–334. [CrossRef Medline](#)
- Doran MM, Hoffman JE (2010) The role of visual attention in multiple object tracking: evidence from ERPs. *Atten Percept Psychophys* 72:33–52. [CrossRef Medline](#)
- Drew T, McCollough AW, Horowitz TS, Vogel EK (2009) Attentional enhancement during multiple-object tracking. *Psychon Bull Rev* 16:411–417. [CrossRef Medline](#)
- Holcombe AO, Linares D, Vaziri-Pashkam M (2011) Perceiving spatial relations via attentional tracking and shifting. *Curr Biol* 21:1135–1139. [CrossRef Medline](#)
- Howard CJ, Holcombe AO (2008) Tracking the changing features of multiple objects: progressively poorer perceptual precision and progressively greater perceptual lag. *Vision Res* 48:1164–1180. [CrossRef Medline](#)
- Howe PD, Horowitz TS, Morocz IA, Wolfe J, Livingstone MS (2009) Using fMRI to distinguish components of the multiple object tracking task. *J Vision* 9(4):10.1–11. [CrossRef](#)
- Jensen AR (2006) *Clocking the mind: mental chronometry and individual differences*. Amsterdam: Elsevier.
- Jovicich J, Peters RJ, Koch C, Braun J, Chang L, Ernst T (2001) Brain areas specific for attentional load in a motion-tracking task. *J Cogn Neurosci* 13:1048–1058. [CrossRef Medline](#)
- Keitel C, Andersen SK, Quigley C, Müller MM (2012) Independent effects of attentional gain control and competitive interactions on visual stimulus processing. *Cereb Cortex*. Advance online publication. Retrieved February 19, 2013. doi:10.1093/cercor/bhs084. [CrossRef Medline](#)
- LaBerge D (1995) *Attentional processing: the brain's art of mindfulness*. Cambridge, MA: Harvard UP.
- McMains SA, Somers DC (2004) Multiple spotlights of attentional selection in human visual cortex. *Neuron* 42:677–686. [CrossRef Medline](#)
- Morgan ST, Hansen JC, Hillyard SA (1996) Selective attention to stimulus location modulates the steady-state visual evoked potential. *Proc Natl Acad Sci U S A* 93:4770–4774. [Medline](#)
- Müller MM, Malinowski P, Gruber T, Hillyard SA (2003) Sustained division of the attentional spotlight. *Nature* 424:309–312. [CrossRef Medline](#)
- Müller MM, Andersen S, Trujillo P, Valdés-Sosa P, Malinowski P, Hillyard SA (2006) Feature-selective attention enhances color signals in early visual areas of the human brain. *Proc Natl Acad Sci U S A* 103:14250–14254. [CrossRef Medline](#)
- Oksama L, Hyönä J (2004) Is multiple object tracking carried out automatically by an early vision mechanism independent of higher-order cognition? An individual difference approach. *Vis Cogn* 11:631–671. [CrossRef](#)
- Posner MI (1980) Orienting of attention. *Q J Exp Psychol* 32:3–25. [CrossRef Medline](#)
- Pylyshyn ZW, Storm RW (1988) Tracking multiple independent targets evidence for a parallel tracking mechanism. *Spat Vis* 3:179–197. [CrossRef Medline](#)
- Quigley C, Andersen SK, Müller MM (2012) Keeping focused: sustained spatial selective visual attention is maintained in healthy old age. *Brain Res* 1469:24–34. [CrossRef Medline](#)
- Regan D (1989) *Human brain electrophysiology: evoked potentials and evoked magnetic fields in science and medicine*. New York: Elsevier.
- Störmer VS, Li SC, Heekeren HR, Lindenberger U (2011) Feature-based interfer-

- ence from unattended visual field during attentional tracking in younger and older adults. *J Vis* 11(2):pii:1. [CrossRef Medline](#)
- Störmer VS, Li SC, Heekeren HR, Lindenberger U (2013) Normal aging delays and compromises early visual attentional selection: evidence from event-related potentials during multiple-object tracking. *J Cogn Neurosci* 25:188–202. [CrossRef Medline](#)
- Toffanin P, de Jong R, Johnson A, Martens S (2009) Using frequency tagging to quantify attentional deployment in a visual divided attention task. *Int J Psychophysiol* 72:289–298. [Medline](#)
- Yantis S (1992) Multielement visual tracking: attention and perceptual organization. *Cognit Psychol* 24:295–340. [CrossRef Medline](#)